University of Szeged

Doctoral School of Geosciences

Ph.D. Thesis

Abdelrhim Eltijani

Szeged, 2023

Doctoral School of Geosciences

Sedimentary records of abandoned channel alluviation: insights from compositional data analysis, endmember modeling, and wavelet transform.

PhD. Thesis

Author

Abdelrhim Eltijani
BSc(hons), MSc

Supervisor

Dr. Dávid Molnár
Assistant professor, Department of Geology and Paleontology, University of Szeged

Consultant

Dr. János Geiger
Associate professor, Department of Geology and Paleontology, University of Szeged, scientific advisor of GEOCHEM Ltd
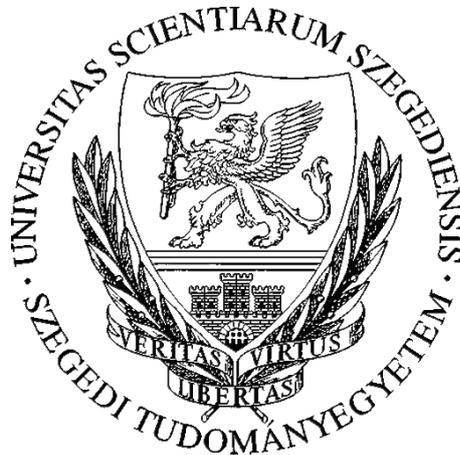
Department of Geology and Paleontology Faculty of Science and Informatics University of Szeged

December 2023
Szeged

TABLE OF CONTENTS

LIST OF TABLES

1. Introduction

1.1 Oxbow lakes and alluviation processes

The sediment deposition and accumulation process occurs within specific environmental conditions involving physical, chemical, and biological factors (Nichols, 2009). These factors play an essential role in shaping the characteristics of the deposited sediment. Depositional environments vary, from fluvial to marine to aeolian and glacial environments. Within the quaternary fluvial system, meandering and their oxbow lakes and abandoned channels sub-environments are common. They result from complex meandering channel migration in alluvial plains. Lakes are susceptible to environmental change, and their response to significant climate fluctuations is extremely complex (Woolway & Merchant, 2019; Zhang et al., 2020). Therefore, oxbow lakes are critical paleoenvironmental indicators for elucidating the interplay between environmental change and fluvial dynamics (Pawłowski et al., 2015; Wang et al., 2023). their sedimentary infill contains proxies. This study uses the terms oxbow lakes and abandoned channels interchangeably. The physical imprint of the environment can be inferred by studying the grain size distributions (GSDs) of the infills because of the grain size sensitivity to transport and depositional mechanism associated with channel alluviation. Other parameters can help decipher the biochemical interactions of the environment, i.e., magnetic susceptibility (MS) and Loss-On-Ignition (LOI). The temporal domain of the interplay of these processes can be established based on radiocarbon dating.

Description and understanding of the depositional processes, pattern, and sediment transport using grain-size parameters are well established. They have become a standard routine in sedimentology, employing various approaches, including several probability density function (PDF) approaches (e.g., log-normal). Utilizing log-normal distribution coefficients can circumvent the need for a detailed account of the various proportions of individual components in GSD introduced by the first property (Folk and Ward, 1957; Evans and Benn, 2014). However, using log-normal distribution is challenging, primarily when GSDs are non-normally distributed (Fieller et al., 1992; Fredlund et al., 2000; Beierle et al., 2002). Additionally, PDF approaches may be problematic when the distribution is multimodal, as this could result in an inaccurate determination of the mean and variance using the model distribution coefficients (Roberson and Weltje 2014). Multivariate statistics e.g., principal component analysis (PCA) and cluster

analysis (CA) can help utilize the entire grain size data; however, challenges posed by the nature of grain size data may arise. Utilizing dimensionality reduction on the PDF (e.g., log-normal and log skew-Laplace coefficients) is redundant because the dataset has already undergone decomposition into a set of model coefficients, with the data being masked through the model fit. Meanwhile, the direct application of the multivariate techniques on the grain size data is not straightforward (Aitchison, 1986; Aitchison et al., 2000; Buccianti et al., 2006; Filzmoser et al., 2009).

For instance, several problems are associated with using PCA in CoDA: 1. Compositional data has a sum constraint *(sections 1.2 and 2)*, *meaning that the components must add up to a constant sum*. However, PCA assumes that the variables are independent, which violates the sum constraint. "*2. Compositional data can exhibit spurious correlations due to the sum constraint"*. PCA can amplify these spurious correlations, leading to misleading results. "*3. PCA assumes that the variables are measured on the same scale, but compositional data is represented on different scales, e.g., weight percentages"*. This can lead to biased results in PCA. "*4. PCA can be difficult to interpret in compositional data analysis"*, as the principal components may not correspond to meaningful physical or chemical processes. To overcome these problems, the CoDA approach of Aitchison 1986 transforms the compositional data into a log-ratio space where PCA can be applied.

The compositional variation of a sample can be explained by a combination of a fixed set of compositions known as endmembers (EMs) (Weltje, 1997). Linear unmixing models, e.g., end-member modeling (EMMA), are commonly used to analyze the proportional contributions of these EMs in the compositional data (Weltje, 1997). EMMA serves as a statistical method that decomposes mixed constituents based on the contributions of their respective EMs (Dietze et al., 2012).

1.2   Research Question

How can abandoned channel alluviation history be comprehensively studied and characterized through the analysis of grain size distributions (GSDs) of their infills, incorporating advanced techniques such as compositional data analysis (CoDA) (section 2), endmember modeling (EMM) (section 3), and continuous wavelet transform (CWT) (section 4), and how can these approaches provide insights into the interplay between environmental change and fluvial dynamics in quaternary fluvial systems?

1.3 Objectives

1. To assess the applicability of compositional data analysis (CoDA) and Euclidean data analysis (EDA) combined with principal component analysis (PCA) and cluster analysis (CA) in characterizing grain size distributions (GSDs) of abandoned channel infills, with a focus on addressing the challenges posed by non-normally distributed and multimodal GSDs.

2. To apply multivariate statistics, including principal component analysis (PCA) and linear discriminant analysis (LDA), to analyze the relationships between different grain size fractions and parameters, aiding in the interpretation of sedimentary processes.

3. To employ endmember modeling (EMM) techniques, specifically utilizing the Weibull function, to statistically unmix GSDs and identify distinct grain size modes, elucidating sedimentation processes in abandoned channels.

4. To model the cyclicity of sediment deposition in abandoned channels through continuous wavelet transform (CWT), integrated with CoDA and EMM models, to provide insights into the temporal dynamics of sediment accumulation.

5. To validate the sedimentological interpretability of the statistical results using Passage's CM diagram, establishing relationships between sediment textures and deposition processes.

6. To characterize the Tövises bed in the eastern Great Hungarian Plain, including its sedimentary sequences, geochronology, and additional parameters such as loess-on-ignition (LOI) and magnetic susceptibility (MS).

7. To integrate radiocarbon dating and Bayesian age-depth modeling to establish the temporal context of sediment deposition in the Tövises bed.

8. To provide a comprehensive quantitative framework for the study of abandoned channel alluviation history by combining advanced techniques, contributing to a better understanding of the interplay between environmental change and fluvial dynamics within quaternary fluvial systems.

1.4 Compositional data and grain size distribution (GSD) (section 2- Eltijani et al. 2022a)

Compositional data represents the relative proportions or percentages of parts or components that make up a whole (Greenacre 2019). Compositional data has a unique

structure that requires special handling in analysis because the sum of all components is fixed at 100%. This means that the values of one component are dependent on the values of the other components, and therefore, the data cannot be analyzed in the same way as non-compositional data. All these properties apply to the grain size data.

(Firstly. GSDs contain relative information about the proportions of different particle sizes in a sediment sample); This means any modification to one proportion of the distribution induces a reciprocal alteration to all remaining proportions. (*Secondly. Grain-size fractions are presented as percentages, containing nonnegative values that sum to constant or 100%.).*

It is imperative to apply specialized techniques to circumvent the misleading interpretations caused by spurious correlations of the grain size fractions in the GSDs. Aitchison (1986) proposed compositional data analysis (CoDA); this approach is based on the principle of *"sub-compositional"* coherence, *i.e., the parts of a composition are constrained to a simplex or a space of constant sum and therefore are not independent of each other* (Greenacre et al., 2019). In multimodal GSD, each grain-size fraction or mode (e.g., clay) represents a sub-composition. CoDA uses log-ratio transformations to overcome this constraint. The fundamental log ratios are obtained by taking the logarithm of the ratio of two parts of a composition (Greenacre 2020). Various types have been suggested, such as the additive log-ratio (alr), the centered log-ratio (clr), and the isometric log-ratio (ilr), which have been employed since Aitchison's earliest research (Aitchison et al. 2000; Greenacre 2019). For instance, if p variables (i.e., grain size fractions) are in the compositional data, the log-ratio transformation can function in (p – 1) dimensional space (Aitchison 1986; Pawlowsky-Glahn and Buccianti 2011). These can remove the compositional constraint of *grain size data*, allowing unconstrained multivariate analysis hence a thorough characterization and obtaining comprehensive information on the GSDs. Moreover, this approach ensures sub-compositional coherence, meaning that the results and conclusions related to a subset of parts remain unchanged when other parts are added or removed. A prerequisite for this method is that all data values are positive. Therefore, ratios are transformed logarithmically and are compared multiplicatively rather than additively.

1.5 MATERIALS: The Tövises bed

Tövises bed is located in the Pocsaj "gate" and Érmellék region in the eastern Great Hungarian Plain (Figure 1.1). The topmost Holocene sequence of this region is characterized by loess sequences overlaid by alluvial fan deposits (Szöőr et al., 1991; Sümegi and Vissi, 1991; Sümegi, 1993).

1.6 Methods

1.6.1 Grain size Analysis *(sections 2-4- Eltijani et al. 2022a and b; Eltijani et al. 2023)*

Undisturbed 346 cm core sediments were stored at a constant temperature of 4ºC and sectioned into 346 discrete 1 cm subsamples. Grain size analysis was completed at 1 cm intervals, and the measurements procedure was followed is Konert and Vandenberghe (1997). First, the samples were dried at 55 ºC. Then 30 ml of $Na_2P_6O_{18}$ solution was added to 0.6 g of the sample to disperse the particles. The grain size analysis was carried out in the Department of Geology and Paleontology at the University of Szeged, Hungary, using the Easysizer20 laser particle sizer instrument of OMEC company; with a measuring range: of 0.1 to 500 μm and a repeatability error of less than 3%. The device uses 54 built-in detectors based on the Mie scattering. After measurements, the GSDs were decomposed into grain size fractions following the Udden Wentworth scale (Udden, 1914; Wentworth, 1922). The (D5), (D50), and (D95) parameters were determined to emphasize the finest, average, and coarsest sizes.

1.6.2 Sediment Cores and Geochronology *(section 3- Eltijani et al. 2022b)*

The core from the Tövises bed core exhibits bimodal and polymodal GSDs. The entire sequence is characterized by periodic intercalation of the sediments, albeit the core is composed of fine materials (clay to very fine sand) (Figure 1.2). The upper part of the core contains alternations, of course, and medium with fine silts in the middle. The second section contains medium silt intercalated with coarse and fine silt at the bottom and the top, respectively. The middle part comprises fine, and very fine silty medium silt followed by the alternating medium and fine silts at the bottom. The lower part consists of alternating coarse, medium, and fine silts. The very fine sand and the medium silty coarse silt fractions show a parallel vertical change in the compositional diagram, while the very fine sand fraction forms five peaks in the sequence.
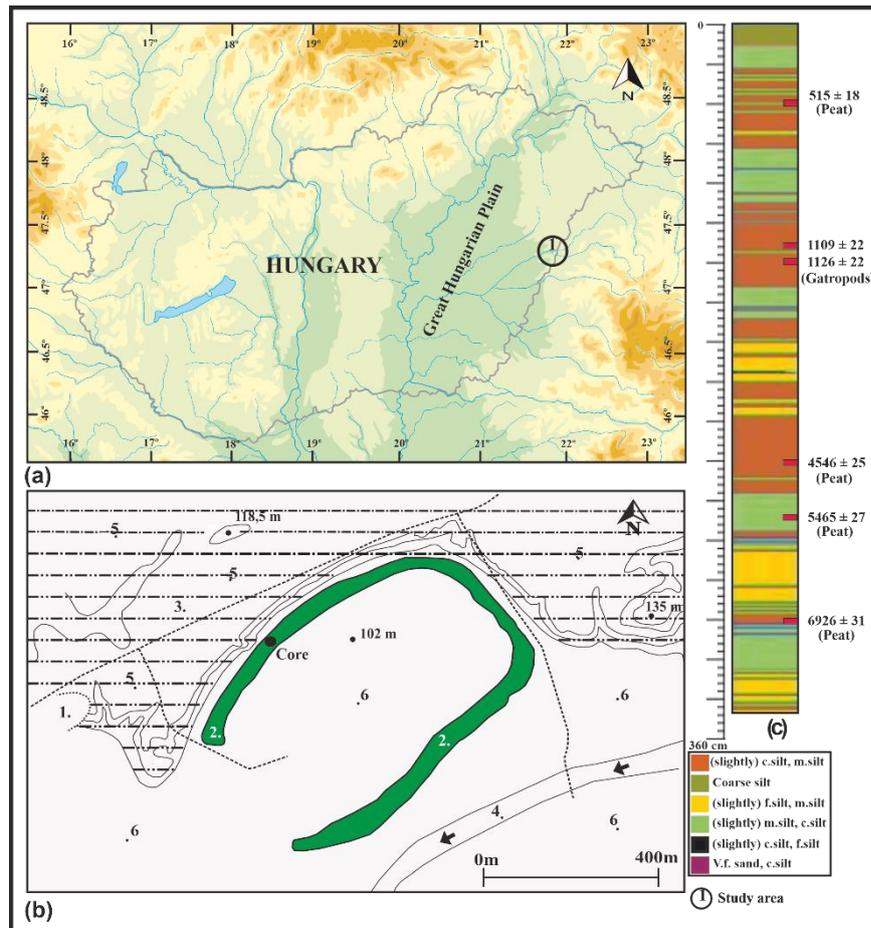
Figure 0.1 Map of the study area. (a) Location of the Tövises paleochannel in GHP; (b) the core location and geology paleochannel of the vicinity: 1. sandpit, 2. Tövises paleochannel, 3. loess-covered Pleistocene alluvial fan, 4. canalized bed of Ér creek, 5. loess, and 6. alluvia sediments; (c) Lithological and vertical GSDs in the sediments, the red boxes indicate the chronology intervals
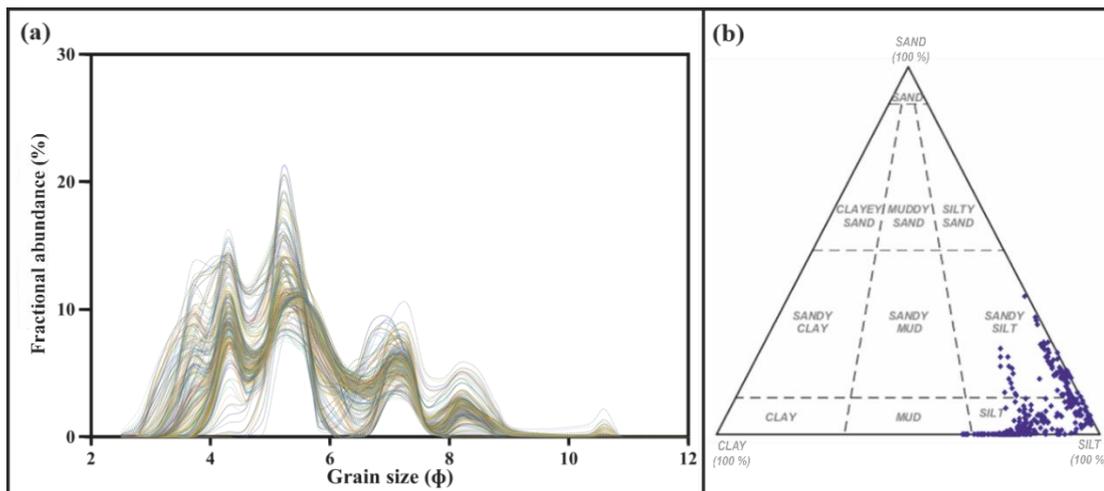
Figure 0.2 GSD characteristics of the Tövises bed core. (a) Overlay plot of the GSDs of all samples (n = 345), illustrating the polymodality of Tövises bed core sediments; (b) classification of grain-size composition in Tövises bed core, fine ternary plot (Folk 1954).

13

The ^{14}C chronology was based on AMS dates from six samples analyzed at the International Radiocarbon AMS Competence and Training Center (INTERACT), Institute for Nuclear Research, Loránd Eötvös Research Network, Debrecen, Hungary. The radiocarbon dates were calibrated to calendar ages in CALIB 8.1 using the IntCal20 calibration curve. The age-depth modeling was performed in R language version 4.1.3 using the R bacon 2.5.8 age-depth modeling package (Blaauw and Christeny 2011) and was based on Bayesian statistics.

1.6.3 Loess-on-Ignition (LOI) and Magnetic Susceptibility (MS) (section 3-Eltijani et al. 2022b)

The sediment (LOI) parameters of this study, i.e., organic material (OM) and carbonate content, were measured at 1 cm resolution following the general LOI procedure (Santisteban et al. 2004; Vereş 2002). One gram of each sample was subject to sequential heating for 24 h at 550 °C to combust the (OM) (LOI_{550}) and at 950 °C to evacuate the CaCO3 from carbonate content (LOI_{950}) and the measured parameter expressed in weight percent.

MS is a bulk measurement representing all the individual susceptibility contributions from the different magnetic mineral grains in the sediment sample. Therefore, it indicates the composition changes linked to climatic conditions that influenced the sediments by measuring magnetizable mineral grains contents. In this study, the MS is measured for homogenous bulk samples using the Bartington MS2 Meter at a 2.7 MHz frequency (Dearing 1999); each sample was measured three times, and the obtained values were averaged. The Cumulative sum (CUSUM) chart detected small shifts and change points in the magnetic susceptibility values. CUSUM uses the cumulative sum of deviations from a given data or subgroup means, and this method is sensitive to small shifts in the process mean (Scouse 1985).

1.6.4 Multivariate statistics *(section 2-Eltijani et al. 2022a)*

1.6.4.1 *Principal component analysis (PCA)*

PCA is a dimensionality reduction technique that summarizes information from a large dataset into small variables that retain the most information from large datasets. PCA identifies the principal components of the data, linear combinations of the original variables that capture the most variation in the data. Detailed descriptions of PCA can be found in several textbooks (e.g., Davis, 2002). PCA can characterize grain size distribution by analyzing the relationships between different grain size fractions and

parameters, such as mean grain size, standard deviation, skewness, and kurtosis). These measures can be treated as variables.

The principal components (PCs) are given by the eigenvectors of the correlation matrix between the different grain size fractions and the statistical parameters; the variances are derived from the associated eigenvalues in PCA. Each PC describes a percentage of the total variance in the GSD; the first PC accounts for the most variation, and all subsequent PCs capture the remaining variance hierarchically. The PCs are orthogonal to each other and thus uncorrelated by definition.

The critical point of PCA is that the PCs can be linked with sedimentary processes, revealing groups of variables that would not be visible otherwise. The interpretation of PCA results seeks to uncover processes (in this case, transportation processes) that cause correlations between the PCs and the original variables (Davis 2002; Szilágyi and Geiger 2012). Using PCA to analyze grain size distribution, one can identify patterns of variation that may be related to different transport and depositional processes or environmental conditions. This information can help understand the samples' history and evolution and predict their behavior under different transport and deposition conditions.

1.6.4.2 *Cluster analysis (CA)*

The clustering entails the segregation of a specified dataset into distinct groupings or clusters, where the members of each cluster display more significant similarity to one another compared to those in other clusters (Halkidi, Batistakis, and Vazirgiannis 2001). The clustering algorithms are categorized as *hierarchical* and *nonhierarchical* (optimal partitioning) (Gulagiz and Suhap 2017). The hierarchical clustering algorithms select the most alike pair of objects within a cluster; however, this approach incurs a high computational cost as all objects are compared before each clustering iteration (Gulagiz and Suhap 2017). Hierarchical clustering comprises agglomerative and divisive approaches. The agglomerative or bottom-up approach combines data points into clusters (Ma and Chow 2004). The divisive or top-down approach operates in the opposite direction, with the computational load required by divisive and agglomerative approaches comparable (Kuo et al., 2002). Cluster analysis can play a role in characterizing grain size distribution by identifying patterns or groups of grain sizes within a sample. For GSD, different clusters or groups of grain sizes can represent

different modes or components of the distribution. For example, a bimodal distribution might have two distinct clusters of grain sizes, each representing a different mode of transportation and deposition.

This study applies the hierarchical agglomerative algorithm because the dendrogram could depict the hierarchy of the different sediment transport and deposition modes. Ward's method defined the objective function algorithm (Ward 1963), where the sum of squared distances as similarity coefficient is minimized. By resampling an equal number of samples from each cluster, it is possible to filter out typical samples; this technique can also be employed to objectively reduce the size of large datasets without the risk of losing crucial data points (Liu and Sun 2021). However, validating the clustering results poses some difficulties as the clustering itself (Pfitzner, et al., 2009); there are 'internal' and 'external' approaches to the evaluations, but they have much uncertainty (Feldman and Sanger 2007). Therefore, (Passage's CM diagram) is applied to validate the sedimentological interpretability of the statistical results.

1.6.4.3   Linear Discriminant Analysis (LDA)

LDA is a technique that distinguishes between two or more data groups based on observed quantitative variables. It was developed for categorizing objects from a set of independent variables in one or more sets of mutually exclusive groups. This model is robust, easy to use, and has high predictive accuracy; however, it needs prior knowledge of groups.

The larger the standardized coefficient indicates a more significant contribution of the corresponding variable to the discrimination between groups. Therefore, the variable with the highest (regression) coefficient contributes most to predicting group membership. The analysis calculates one discriminant function for each group, and these functions are independent by construction, so the discrimination between groups does not overlap. The classification table shows the result of assigning observed and new cases to a group using derived classification rules.

1.6.5   The CM Diagram section (2- Eltijani et al. 2022a)

The CM diagram (Passega 1957) (Passega 1964) is used to establish the relationships between the sediment textures and processes of deposition. Passega (1957) defined M and C as the cumulative GSDs' median and one percentile. These values can readily be obtained as the grain diameters (in mm or μm) belonging to the 50 and 95 percentiles of

the cumulative distribution functions. Sometimes it is hard to determine the diameter of one percentile based on the laboratory analysis; therefore, the D95 percentile is used instead.

1.6.6 Endmember Modeling (EMM) *(section 3-Eltijani et al. 2022b)*

Compositional variation of a particular sample may be attributed to a mixture of a fixed number of compositions, denoted as endmembers (EMs) (Weltje 1997). EMM is used to analyze the proportionate contributions of EMs of compositional data (Weltje 1997). Grain-size EMM has become widely recognized for unmixing the GSDs, which helps reconstruct the sedimentation process. The grain-size EMM represents a statistical unmixing of robust polymodal grain-size population. Each resulting EM represents a particular grain-size mode of particles of the same sedimentation history. Fluvial produce characteristic GSDs by sorting sediments; however, the syn and postdepositional processes obscure and mix the process signatures preserved in the sedimentary archive. EMM is a viable strategy for unmixing wholesale grain size distribution (GSD) as it decomposes the overlapping components concerning the contribution of its EMs. Thus, it provides a process-based explanation of the GSD structure (Elisabeth Dietze et al. 2012a).

Many researchers have developed various algorithms following the first EMM algorithm introduced by Weltje (1997). The EMM of the GSDs was performed using the AnalySize v.1.2.0 EMM algorithm (Paterson and Heslop 2015), a graphic user interface (GUI) based in MATLAB that offers a wide range of distribution functions that can be used to unmix the original distribution into endmembers (EMs) (Meyer et al. 2020). The function used here is the Weibull distribution function (Wu et al. 2020) which provides an accurate explanation of the dynamics of the EMs sedimentation process by considering the geological background of the GSD while unmixing the GSD. The optimum number of the EMs is determined based on the determination coefficient (R^2) and the angular deviation (θ) among the EMS and the GSDs; This procedure enhances the fitting accuracy and avoids data overfitting (Kong et al. 2021).

1.6.7 Time series analysis and wavelet transformation *(section 4)*(Eltijani, et al., 2023)

Time series analysis is commonly used to discover a steady pattern in regularly measured datasets for identifying and quantifying processes. One conventional time series approach to extract cyclicity uses a Fourier transform (FT). FT captures the

frequencies over the entire signal "i.e., the analysis window cannot capture features in the signals that are either longer or shorter than the window" (Prokoph and Patterson 2004); besides, many Fourier components are required for the analysis. The wavelet transforms (WT) is beyond FT (Saadatinejad and Hassani 2013) that can characterize shorter time series, multiscale, non-steady processes and extract local and temporal information (e.g., periods cyclic events) simultaneously in spatial, time, and domains. Examples of WT applications in geology and geophysics include (Álvarez et al. 2003; Li, Deng, and Dai 2007; Sinha, Routh, and Anno 2009; Coconi-Morales, et al 2010). In sedimentary time series, the WT can detect gradual and abrupt GSD variations and characterize the pattern of the related transport and deposition processes.

In WT, the primary function is wavelike (wavelet), with a frequency and an amplitude decaying to zero at the two ends. The focus on the spatial frequency perspective as the analysis is for depth series and consider t to represent depth. A variable such as s(t) is the signal under consideration (i.e., the sand fraction datasets as a function of the core depth t). WT decomposes a signal (t) in terms of some elementary functions $\psi a, b\ (t)$ derived from the "mother wavelet" $\psi(t)$ by dilation (stretching or compressing) and translation:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\ \psi\left(\frac{t-b}{a}\right) \quad (1)$$

Where b represents the position (translation), a (> 0) the scale (dilation) of the wavelet, and $\psi\_(a,b)\ (t)$ is the daughter wavelet ($1/\sqrt{a}$) is an energy normalization factor for maintaining the same energy between daughter and mother wavelets (Lau and Hengyi Weng 1995). The wavelet transform value is the transform of a real signal (t) concerning the analyzing wavelet $\psi(t)$ defined as a convolution integral:

$$w_{\psi}(a,b) = \frac{1}{\sqrt{a}} \int \psi^{*}\left(\frac{t-b}{a}\right)\ s(t)dt \quad (2)$$

Where a represents the complex conjugate of the mother wavelet $(t)\psi$, a is the location parameter, and b is the scaling dilation parameter. The CWT compares the signal to shifted and dilated wavelets by comparing the wavelet signals at various scales and positions b. The result is a function of two variables (Kadkhodaie and Rezaee 2017).

2 Applying Grain-size and Compositional Data Analysis for Interpretation of the Quaternary Oxbow Lake Sedimentation Processes: Eastern Great Hungarian Plain

Abdelrhim Eltijani[*1] ; Dávid Molnár[1]; László Makó[1]; János Geiger[1,2]; Pál Sümegi[1]

Abstract

Grain size distribution is one of the paleoenvironmental proxies that provide insight to the statistical distribution of size fractions within the sediments. Multivariate statistics have been used to investigate the depositional process from the grain size distribution. Still, the direct application of the standard multivariate methods is not straightforward and can yield misleading interpretations due to the compositional nature of the raw grain size data. This paper is a methodological framework for grain size data characterization through the centered log-ratio transformation and Euclidean data, coupled with principal component analysis, cluster analysis, and linear discriminant analysis to examine Quaternary sediments from Tövises bed in the southeast Great Hungarian Plain. These approaches provide statistically significant and sedimentologically interpretable results for both datasets. However, the details by which they supplemented the conceptual model were significantly different, and this discrepancy resulted in a different temporal model of the depositional history.

2.1   Introduction

Grain Size Distribution (GSD) is a paleoenvironmental proxy that provides information on depositional environments and processes (He et al., 2015; Zhang et al., 2018). GSD first evolves during transport and deposition (Reading, 1996; McLaren et al., 2007). Numerous works have interpreted the depositional conditions based on GSDs by applying univariate statistical parameters. For instance, the median, mean, sorting, and skewness of the distribution (e.g., Folk and Ward, 1957; Visher, 1969; Blott and Pye, 2001; Fournier et al., 2014). The CM patterns where (C) is one percentile and (M) is the median grain size can help analyze the ancient and recent depositional processes (Passega, 1957, 1964, 1977). One of the methods to determine the grain size fractions susceptible to environmental changes is classifying the GSDs using the standard deviation of the distribution (Boulay et al., 2003). These methods, however, depend on median diameter instead of the whole distribution, revealing relative and restricted information on the distribution (Zhang et al., 2018).

The polymodal GSDs indicate the presence of individual subpopulations (Folk and Ward, 1957; Ashley, 1978; Flemming, 1988). Moreover, the polymodality also suggests that the sediments were not well mixed in the suspension. Multivariate approaches to distinguish the subpopulations of GSDs include, among others, cluster analysis (CA) and principal component analysis (PCA) (Sarnthein et al., 1981; East, 1985, 1987). However, the application of cluster analysis on GSDs focused on provenance studies and stratigraphic analysis. Furthermore, these studies used a few parameters, e.g., mean and standard deviation. These parameters provide limited information on the depositional conditions (Donato et al., 2009; Zhang et al., 2018). Although statistical analysis is useful for GSDs description, PCA represents a crucial dimensionality reduction method (Palazón and Navas, 2017; Katra and Yizhaq, 2017). The direct application of such statistical techniques for analyzing GSDs is challenging because the grain size data is a typical compositional set (e.g., Aitchison, 1982; Flood et al., 2015). The sum of weight percentages of the size fractions is 100%. Consequently, they do not form an independent system (Aitchison, 1986; Flood et al., 2015). The definition of compositional data has gradually evolved from vectors of positive components adding to a given constant (e.g., Aitchison, 1982) to a new general definition based on equivalence classes (Egozcue et al., 2018). A composition is a set of multivariate vectors that vary by a scalar factor and have nonnegative components. Accordingly, the

composition can be regarded as a vector of proportions with nonnegative components constrained to a K constant.

One of the crucial questions in statistical analysis is elucidating the compositional constraint. A simple approach ignores compositional constraints (i.e., Euclidean data analysis approach) and treats the data as Euclidean (Tsagris et al., 2016). On this topic, there is another school following the work of Aitchison (1982, 1983, 1992). The followers of this school have suggested several transformations using the logarithms of ratios, or log ratios, to get solutions for the compositional constraints (e.g., Aitchison, 1986; Aitchinson et al., 2002; Pawlowsky-Glahn and Egozcue, 2001; Egozcue et al., 2018).

This paper aims to study the utility and geological interpretability of the Euclidean and centered log ratio (clr) transformation approaches coupled with cluster analysis, principal component analysis, and Linear Discriminant Analysis through investigating the GSD of oxbow lake sediments of Tövises bed from eastern Great Hungarian Plain.

2.2 The workflow of the analysis

The workflow involves the calculation of the D50, D95, and D5 percentiles from the cumulative grain size distributions, then the decomposition of GSD into clay, very fine silt, fine silt, medium silt, coarse silt, and very fine sand fractions. These fractions are then transformed by applying the additive log-ratio transformation (clr) in the open-source CoDaPack, version 2.02.21. (Aitchison, 1986; Egozcue and Pawlowsky-Glahn, 2005). Subsequently, two datasets were compiled for further use. The first dataset contained the frequency percentages of the different grain size fractions and the diameters belonging to the D5, D50, and D95 cumulative percentages of the grain size distributions.

The second contained the clr-transformed percentages of the grain size fractions and the D5, D50, and D95 diameters. The PCA was performed on both datasets. The transporting processes were interpreted using the high loadings of the most important components. The optimum number of PCs to be considered can be determined based on the Kaiser criterion (i.e., retention of PCs whose eigenvalue is greater than 1) and scree plot criteria. The first and second PC scores of each sample were used as variables in the hierarchical cluster analysis (HCA) applying Ward's method (Ward, 1963). Since the clustering method relied on the component scores, the resulting clusters are

supposed to correspond to the different transport processes of the fluvial system. This thought was checked by depicting the points of the clusters in the CM diagram (Figure 5). Parallelly, the statistical reality of the clusters was tested by discriminant analysis.

2.3   Result and Interpretation

2.3.1   Result

2.3.1.1   Result of principal component analysis (PCA)

The PCA for clr-transformed data has resulted in two PCs. The first component described 60.038% of the total variance, while the second accounted for 23.182%. In the case of the non-transformed data, the first two components could describe 87.198% of the total variance (Table 2.1). The variance explained by the first two PCs is large enough for both datasets to base the interpretations on them.

In the case of the non-transformed dataset, the grain size fractions belonging to the very fine suspension have high positive loadings in PC1. All the coarser grain size fractions and the M and C parameters appeared with strong negative component loadings in the first PC. PC2, uncorrelated with PC1, has one variable (i.e., medium silt fraction (Table 2.1). The clr-transformed dataset showed that the grain sizes from clay to medium silt showed high positive loadings. In contrast, the coarsest grain size fractions showed significantly negative PC loading. The M and C change parallel with the coarse silt fraction in the PC2 (Table 2.1).

Table 2.1 The results of the PCA for the clr-transformed and non-transformed datasets. Principle component loadings larger than 0.7 are highlighted in gray.

			clr-Transformed		Non-Transformed	
			PC 1	PC 2	PC 1	PC 2
Fractions		Clay	0.929	-0.397	0.955	0.073
		Very Fine Silt	0.746	0.288	0.924	0.005
		Fine Silt	0.941	0.003	0.997	0.045
		Medium Silt	0.290	0.943	0.897	0.236
		Coarse Silt	-0.934	0.199	-0.448	0.739
		Very Fine Sand	-0.902	-0.269	-0.898	-0.222
Parameters		D5	-0.847	0.418	-0.781	-0.268
		M (=D50)	-0.982	0.049	-0.533	0.794
		C (=D95)	-0.910	-0.185	-0.006	0.914
% of variance			60.038	23.182	71.786	15.412
Cumulative % of the variance			60.038	83.22	71.786	87.198

2.3.1.2   Result of cluster analysis (CA)

Both datasets could be subdivided into four clusters. Their average compositions are summarized in Table 2.2. The average compositional data showed that the fine silt medium silty prevails in cluster 4 in both sample sets; the medium silt in cluster 3 in the clr-transformed set and cluster 2 in the non-transformed set; the coarse silt in cluster 2 of the clr-transformed and cluster 1 of the non-transformed sets; and the cluster 1 group of the clr-transformed data and cluster 3 subsets of the non-transformed sets medium and coarse silts (Table 2.2).

According to the CM pattern, the studied sediments were deposited partly from the RS (Figure 2.1a, b) and partly from the QR (Figure 2.1a, b). Within the QR, three parts can be described as fine to coarse: fine-grained, medium-grained, and coarse-grained QR. In Figure 2.1, four colors code is used to assign the four clusters of the clr-transformed and non-transformed sets. This method was effective in the identification of transport processes. The results showed that the deposits of RS belonged to cluster 4 in the case of clr-transformed and cluster 3 in the case of non-transformed datasets. The sediments of the fine-grained graded suspension were represented by cluster 3 and cluster 2 in the clr-transformed and the non-transformed samples, respectively. Cluster 1 of both datasets contained the medium-grained QR sediments, while the coarse-grained QR was shown by cluster 2 of the clr-transformed and cluster 4 of the non-transformed datasets (Figure 2.1a, b). It showed that similar suspensions could be identified in the clr-transformed and the non-transformed sets. Cluster 1 contained the medium-grained QR (Figure 2.1).

2.3.1.3   Result of discriminant analysis (LDA)

The relationship between the two classifications and the reliability of the clustering is further explained by discriminant analysis results (Table 2.3). The success of classifications in both the clr-transformed and non-transformed sets was high, 95% and 94%, respectively. That is, both classifications were valid from the statistical point of view.
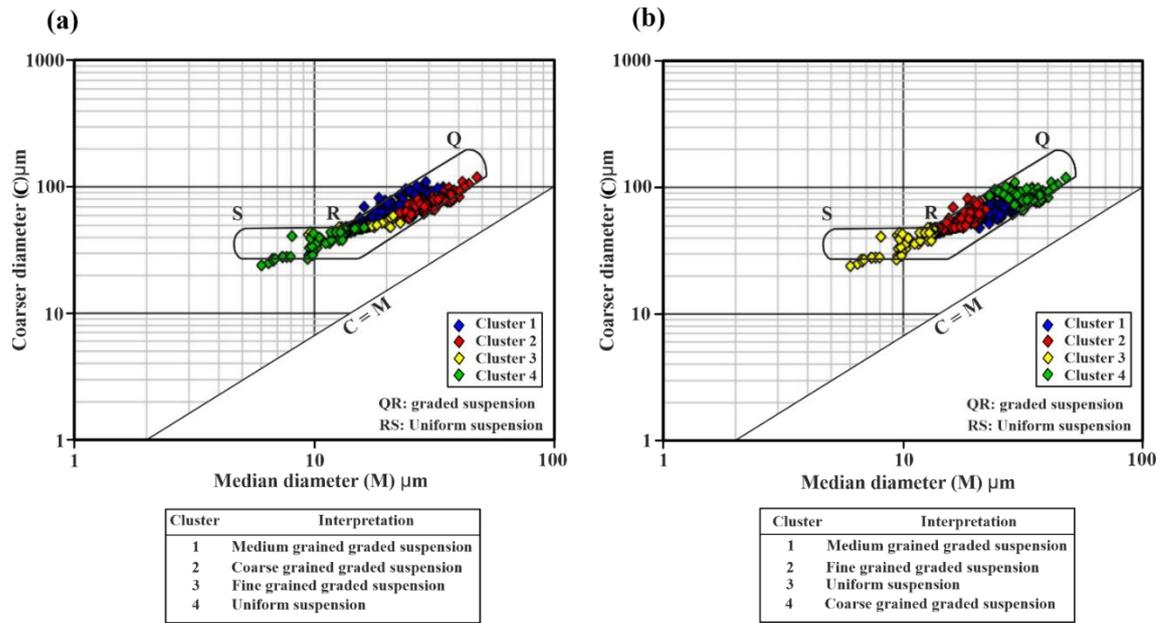
Figure 2.1 Relations between the clusters and the CM diagrams in (a) the clr-transformed and (b) the non-transformed datasets.

Table 2.2 The average composition of the clusters in the clr-transformed and the non-transformed datasets. The dominant fractions are highlighted.

Data	Cluster	No. of Samples	Clay	very Fine Silt	Fine Silt	Medium Silt	Coarse Silt	Very Fine Sand	Sediment Classification
clr-Transformed	Cluster 1	114	11.7	6.9	16.62	29.22	27.01	8.4	Coarse Silty Medium Silt
	Cluster 2	81	1.47	6.74	7.69	29.04	41.95	13	Coarse Silt
	Cluster 3	98	18.52	8.19	20.11	30.92	21.87	0.38	Medium Silt
	Cluster 4	53	25.13	11	22.33	29.22	12.3	0.02	Fine Silty Medium Silt
Non-Transformed	Cluster 1	60	1.38	6.53	7.27	28.89	43.58	12.33	Coarse Silt
	Cluster 2	79	9.52	7.6	17.18	32.28	27.64	5.77	Medium Silt
	Cluster 3	39	5.65	5.63	10.52	24.38	32.84	20.32	Medium Silty Coarse Silt
	Cluster 4	168	21.1	8.97	20.64	29.93	19	0.36	Fine Silty Medium Silt

Table 2.3 The discriminant analysis results of the clr-transformed and non-transformed datasets

Class	clr-Transformed					Non-Transformed				
	Percent Correct	Cluster 1 p=.1850	Cluster 2 p=.4509	Cluster 3 p=.1503	Cluster 4 p=.2139	Percent Correct	Cluster 1 p=.1850	Cluster 2 p=.4509	Cluster 3 p=.1503	Cluster 4 p=.2139
Cluster 1	99.1228	113	1	0	0	95.3125	61	3	0	0
Cluster 2	100	0	81	0	0	100	0	156	0	0
Cluster 3	100	0	0	98	0	73	0	14	38	0
Cluster 4	69.8113	0	0	16	37	95	2	2	0	70
Total	95.0867	113	82	114	37	94	63	175	38	70

2.3.1.4 Vertical sequence of the identified types of suspensions

In Figure 2.2, each sample is assigned to the suspension type suggested by Figure 2.1 and put back to the actual stratigraphical order. In that way, the vertical pattern of the temporal change of the suspension type is established. In Figure 2.1, the pattern of QR (Figure 2.1a, b) is subdivided into; fine, medium, and coarse-grained QR. So, the fining upward (FU), and coarsening upward (CU) (Figure 2.2) are implied. In that way, both vertical sets described a seven-step temporal evolution. In the case of clr-transformed data, the CU sequences suggested flooding conditions with gradually increasing transport energy. In contrast, these flooding sequences were described with relatively thick beds of coarse-grained graded suspension (Figure 2.2).

2.3.2 Interpretation

The most striking feature of the Tövises bed core is the periodic intercalation of the sediments, albeit the core is composed of fine grains (clay to very fine silt). This situation is well expressed in the compositional chart (Figure 2.2). The CM pattern of the samples suggests that the cyclicity can be connected to the intermittently increasing transport energy when the coarser-grained QR can also appear. In these periods, sediments of traction carpet origin were deposited. So, an oxbow lake environment is

probably the depositional site. This oxbow lake intermittently could get sediment influx from the nearby main channel in the form of QR.
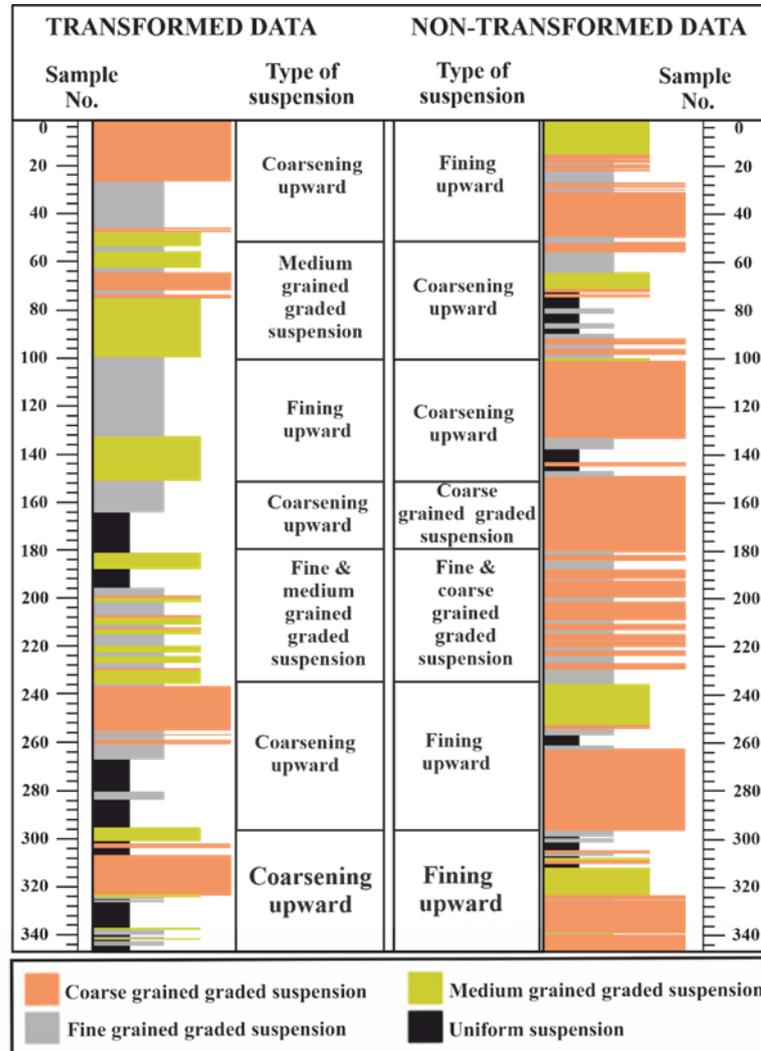
Figure 2.2 Vertical sequences of the samples belonging to the different clusters and the identified vertical units of the transport processes.

In the case of clr-transformed samples, the volume of the deposited RS increases with the increasing first background process. This process decreases the volume of the deposited QR (Table 2.1, "Transformed," PC1). This process is probably (current free) sedimentation in an oxbow lake environment. The second principal component represents such a process.

Increasing the M and C increases the coarse silt fraction (Table 2.1 "Transformed," PC1). In the CM diagram, the joint increase of M and C describes the QR of the traction carpet (Figure 2.1). In the clr-transformed dataset, this process affects only the coarse silt fraction (Table 2.1 "Transformed," PC1). In an oxbow lake, two independent

depositional processes can be outlined. The principal one was the quiet water sedimentation of the RS. This sedimentation was temporarily interrupted by the loads with traction carpet origin. In these periods, coarse silt grains were deposited. This deposition model adds essential detail to the conceptual model: the oxbow lake had weak periodic connections with the channel.

For non-transformed data, the first PC describes that the decrease of M and C increases the RS deposits and de- creases the coarser grain sizes (Table 2.1 "Non-transformed," PC1). PC1 describes the sedimentation from a traction carpet under decreasing mean flow velocity as the principal process of the deposition. There is only one grain size class in the second PC with large positive loading, i.e., the medium silt fraction is (Table 2.1 "Non-transformed," PC2). The critical consequence of this depositional model is that traction flows controlled the deposition with decreasing energy. This model supplement assumes that oxbow lake had a permanent but weak connection with the main channel. During the flooding periods, this weak connection became stronger.

The applied CA gave a possibility to define units of the depositional history. Seven depositional events could be described (Figure 2.2). The results showed that in the case of the clr-transformed dataset, the seven units corresponded to seven mainly CU cycles. This situation may suggest that the units can be connected to the periodic connection between the oxbow lake and the main channel. The non-transformed dataset has seven units characterized by FU cycles (Figure 2.2). In this case, the cycles can be drawn back to the periodically decreasing energy of the almost permanent weak traction f lows in the oxbow lake system. The situation assumes a permanent but weak connection with the adjacent channel. Figure 2.3 connects these two results with the compositional diagram of the cores.

2.4   Discussion

This paper explains the utility of centered log-ratio transformation, euclidean data analysis, and multivariate methods in characterizing the compositional grain size data. The advantage of using clr-transformed and non-transformed datasets coupled with multivariate statistics (PCA, CA, and LDA) is that the entire GSD is considered in the analysis, unlike applying only the individual size fractions.

The first sediments to be deposited after cutting off from the main channel are a plug of channel sands at each end of the oxbow lake. The exchange of water and sediments be-

tween the oxbow lake, and the active channel is maintained through a narrow channel (Rowland and Dietrich, 2006). These narrow channels (known as the Tie channel) develop during lake formation (Blake and Ollier, 1971). Oxbow Lake can receive relatively fine sediment transported as suspension onto the floodplain (Allen, 1970). Therefore, the oxbow lake sediments consist of silt, clay, and coarser sediments. The water enters the lake when the river is active, and the lake level is below the river during a flood. Characterizing this variability and processes through data decomposition into dimensionally reduced components enables maximal variance to be retained using PCA. PCA applied to the clr-transformed data discovered that the first and second PCs characterize 60% and 23% of the variance, respectively (Table 2.1). Correspondingly, for non-transformed data, 71% and 15% of the variance are explained by the first and second PC variances, respectively (Table 2.1).
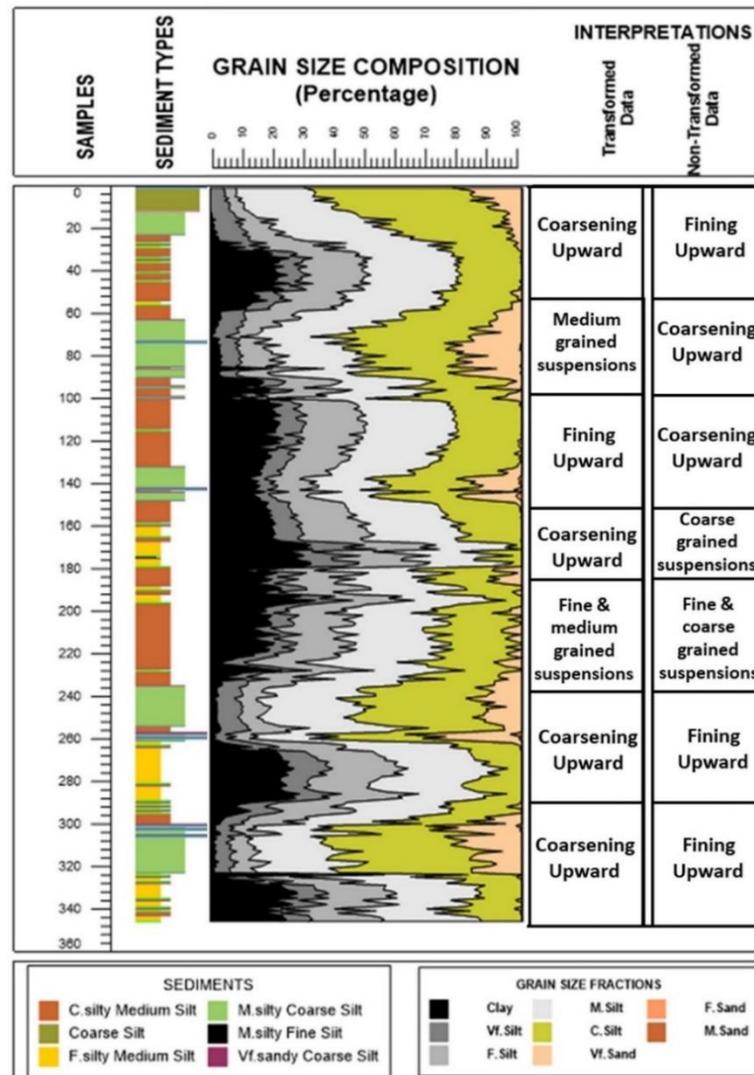
Figure 2.3 A general summary of the compositional charts and the depositional histories derived from the cluster analysis of the clr-transformed and non-transformed datasets.

The following substantial question is how this conceptual model interprets the results from analyzing the clr-transformed and the non-transformed datasets. In interpreting PCA, PCs are regarded as independent background processes. A particular background process significantly influences the variables with high component loadings (high positive or small negative numbers). The sign of the component loadings describes whether the increasing background process increases (positive sign) or decreases (negative sign) the affected variables. In the case of the clr-transformed data, the very fine suspension sediments (clay, very fine silt, and fine silt) have high positive loadings in PC1. All the coarser grains sizes and the M and C parameters have high negative loadings in the first PC. PC2, uncorrelated with PC1, had only one important variable, the medium silt fraction (Table 2.1), suggesting that the oxbow lake had periodic (during floods) connections with the main channel. The non-transformed data showed that the grain sizes from clay to medium silt had high positive loadings, while the coarser grain sizes showed a significantly negative PC loading. The M and C parameters and coarse silt fraction show positive loadings in PC2 (Table 2.1), indicating that the oxbow lake had a permanent but weak connection with the main channel, probably, through a Tie channel. During the flooding periods, this weak connection became strong.

The CA applied for the principal component scores of both datasets revealed four groups. The validation of the clustering results poses some difficulties as the clustering itself (Pfitzer et al., 2009). There are 'internal' and 'external' evaluation methods, but they are uncertain (Feldman and Sanger, 2007). However, the classification is statistically valid as the LDA indicates a high total percent corrects for clr-transformed and non-transformed datasets, 95% and 94%, respectively (Table 2.3). The sedimentological criteria used to judge the efficacy of these models is the CM diagram (Figure 2.1). Accordingly, the clustering of the clr-transformed data was inefficient in generating a distinct boundary between deposition by QR and the deposition by the RS (Figure 2.1a), as cluster 4 is presented as deposition of RS and QR. Therefore, the two approaches can be applied in similar situations where the sedimentological and geological criteria can assess the validity and test their efficacy.

2.5   Conclusion

This paper represents a methodology framework for characterizing the GSDs through the centered log-ratio transformation, Euclidean data analysis approaches, and

multivariate statistics. The presented methodologies eliminate challenges caused by the closed dataset with PCA extracting the maximum variance present. This variance was studied through CA and LDA. The study revealed that the deposition occurred as RS and QR loads in multiple stages interrupted with bottom current loads.

- The CA applied for the PC scores of both datasets revealed four groups of sequences. The vertical sequence can be subdivided into seven genetic units with similar boundaries. However, in the case of the clr-transformed dataset, they corresponded to mainly CU, suggesting the periodic connection between the oxbow lake and the main channel. Contrary to the clr-transformed dataset, the seven units of non-transformed data are characterized by FU cycles, indicating the permanent weak traction f lows with periods of decreasing energy in the oxbow lake system. The situation assumes a permanent but weak connection with the main channel.

- This finding demonstrates a great potential for applying clr-transformation, and Euclidean data analysis approaches coupled with PCA, CA, and LDA to characterize the GSD and interpret oxbow lakes' deposition and sedimentation processes. The results are statistically significant and sedimentologically interpretable for both datasets. However, the details by which they supplemented the conceptual model are significantly different, resulting in a different temporal model of the depositional history. Therefore, the reliability of the models derived from such methods must be cross-checked with sedimentological and geological criteria, as they cannot guarantee a meaningful result.

3 Application of Parameterized Grain-Size Endmember Modeling in the Study of Quaternary Oxbow Lake Sedimentation: A Case Study of Tövises Bed Sediments in the Eastern Great Hungarian Plain

Abdelrhim Eltijani[*1] iD ; Dávid Molnár[1]; László Makó[1]; János Geiger[1,2]; Pál Sümegi[1]

Abstract

Abandoned channels are essential in the Quaternary floodplains, and their infill contains different paleoenvironment recorders. Grain-size distribution (GSD) is one proxy that helps characterize the alluviation and associated sedimentological processes of the abandoned channels. The classic statistical methods of grain-size analysis provide insufficient information on the whole distribution; this necessitates a more comprehensive approach. Grain-size endmember modeling (EMM) is one approach beyond the traditional procedures that helps unmix the GSDs. This study describes the changes in the depositional process by unmixing the GSDs of a Holocene abandoned channel through parameterized EMM integrated with lithofacies, age–depth model, loss-on-ignition (LOI), and magnetic susceptibility (MS). This approach effectively enabled the quantification and characterization of up to four endmembers (EM1-4); the characteristics of grain-size endmembers imply changes in sedimentary environments since 8000 BP. EM1 is mainly clay and very fine silt, representing the fine component of the distribution corresponding to the background of quiet water sedimentation of the lacustrine phase. EM2 and EM3 are the intermediate components representing the distal overbank deposits of the flood. EM4 is dominated by coarse silt and very fine sand, representing deposition of overbank flow during the flood periods. This paper demonstrates that the parametrized grain-size EMM is reasonable in characterizing abandoned channel infill sedimentary depositional and sedimentation history.

3.1   Introduction

Grain-size distribution (GSD) is a fundamental aspect of sediments and is widely utilized to characterize different sediment transport processes. Fluvial transport is an essential transporting mechanism; the deposits of a fluvial channel are either accumulated through the channel or by infilling an abandoned channel. Abandoned channel fills are common in Quaternary fluvial systems; they result from complex meander migration over the floodplain (Shen, et al 2021). Three distinct stages can be described: cutoff initiation, plug bar formation, and channel disconnection (Toonen, et al 2012). The cutoff occurs when the discharge flows through the new channel and does not reach the meander loop (Hooke 1995; Toonen, et al 2012; Collinson and Lewin 1983); as a result, plug bar sediments accumulate, further triggering blockage of the discharge (Constantine et al. 2010), and the meander loop becomes a floodplain lake (Toonen, et al., 2012). Channel-fill sediments contain proxies that can be used for paleoenvironmental reconstructions (Willem H.J. Toonen, Kleinhans, and Cohen 2012). GSD of these infills is one of the paleoenvironmental proxies that help characterize the alluviation history of the abandoned channels. Fluvial sediments exhibit polymodal GSD (Sun et al. 2002), comprising many overlapping subpopulations (Collinson and Lewin 1983; Hooke 1995) corresponding to different transport mechanisms and depositional processes (Sun et al. 2002). Deciphering the depositional conditions from such distribution requires a thorough understanding of the individual components, which cannot be achieved by the classic statistical analysis (e.g., graphical descriptive method and statistical methods of quartile measures and method of moments). In recent years, grain-size EMM has become common in grain-size analysis (Shang et al. 2016; Liang et al. 2019; Kong et al. 2021) for unmixing the GSDs, unveiling the obscured information on the sedimentation process, and helping reconstruct paleoenvironments and the infill history of oxbow lakes. The grain-size endmembers (EMs) are obtained by statistically unmixing robust sub-populations, with each endmember representing a proportion of particles deposited by the same sedimentary processes (Meyer et al. 2020). EMM was applied to reconstruct sediment sources and sedimentation processes in various sedimentary environments; marine (Kong et al. 2021; Ha, Chang, and Ha 2021; de Mahiques et al. 2021; Huang et al. 2020), aeolian (Kong et al. 2021; Vandenberghe et al. 2018; Xiaonan Zhang et al. 2021), lacustrine (Meyer et al. 2020; IJmker et al. 2012; Dietze et al. 2013; E. Dietze et al. 2014; J. Wang et al., Li 2021), and fluvial (Toonen et al. 2015;) employing different unmixing algorithms and producing

promising results. Later on, in Quaternary sediments, the application of EMM has been promising, albeit it has not been extensively employed to study Quaternary abandoned channel sediments and is more often applied in the marine realm. After the first unmixing algorithm was published by (Weltje 1997), many investigators have developed new unmixing algorithms, e.g., EMMAgeo (Dietze et al. 2012b), AnalySize (Paterson and Heslop 2015), BEMMA (Yu et al., 2016), and BasEMMA (Zhang et al., 2020). Since oxbow lakes function as a sediment sink during floods with high preservation potential, and their sediments offer an excellent proxy concerning polymodality of the GSDs, thus providing essential information regarding depositional processes and dynamics, EMM can help reconstruct paleoenvironments and the infill history of oxbow lakes.

Extensive studies of Quaternary paleoenvironmental reconstructions have already been conducted in the Great Hungarian Plain (GHP); for instance, (Sümegi, et al., 2014; Vincze et al. 2020). Still, information on the Quaternary channel fill sediments is limited; many more records are required to fully document the spatiotemporal patterns of past environmental and depositional processes across the GHP. This paper intends to show that the abandoned channel-fill sequences are useful recorders of sedimentological processes and paleofloods by improving the understanding of the internal build-up of the oxbow lake sedimentary sequence through grain-size EMM. This study investigates the sediments of the Tövises channel, a Quaternary oxbow lake near Pocsaj in the eastern GHP (Figure 1). The topmost Holocene sequence of the study area and its vicinity is made up of loess sequences overlain by alluvial fan sediments (Szöőr, et al., 1991; Sümegi and Vissi e 1991)**.**

3.2   Results

3.2.1   Lithology and Stratigraphic Divisions

The ^{14}C ages of the Tövises bed core range from 7748 BP at 304 cm to 532 BP at 43 cm (Table 3.1) (Figure 3.1a). Details of the radiocarbon ages used to construct the age-depth models are shown in Table 3.1. The Bacon-built age-depth model (Figure 3.1a) is concurrent (R2 = 0.99) for the 7748 years represented by the Tövises bed sediments. Between circa 8000 and 6000 BP, the pre-dominant sediments accumulated are fine and medium silts intercalated coarse silt, with a minor fine sand fraction. Following a transition between approximately circa 6000 and 2500 BP, the sediment changed to the dominant sediments containing medium and coarse silts with fine silt, which was less

dominant. Between circa 2500 and 500 BP, the dominant sediment composition remained relatively unchanged (i.e., medium and coarse silt): fine and medium silt intercalated coarse silt with a minor, very fine sand fraction. Accumulation of the coarse silt continued until the channel became abundant.

Table 3.1 AMS radiocarbon dates for samples from Tövises bed core.

Lab ID	Depth (cm)	Sample Type	Conventional ^{14}C Age	Calibrated ^{14}C Age (2σ BP)	Weighted Mean ^{14}C Age (BP)
DeA-29986	43	Peat "bulk"	515 ± 18	514–545	532 ± 15.5
DeA-29550	116	P.corneus shell	1109 ± 22	958–1058	1007 ± 50
DeA-29551	121	P.corneus shell	1126 ± 22	959–1065	1016 ± 53
DeA-29988	224	Peat "bulk"	4546 ± 25	5052–5188	5161 ± 68
DeA-29990	252	Peat "bulk"	5465 ± 27	6266–6305	6273 ± 19.5
DeA-29992	305	Peat "bulk"	6926 ± 31	7678–7799	7748 ± 60.5
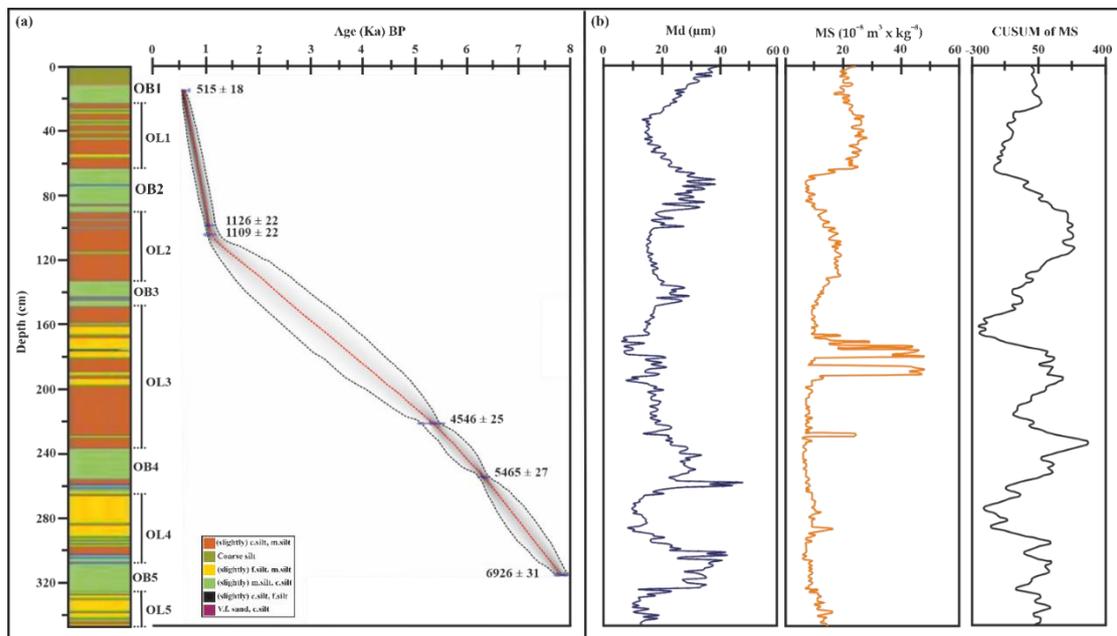
Figure 3.1 Stratigraphic division of the Tövises bed sequence. (a) Age–depth model, the calibrated 14C dates (blue, with 2σ errors), and darker grey shading represent the probable calendar ages; grey stippled lines indicate 95% confidence intervals; the red line is the single "best" model based on the weighted mean ages. (b) MS, the median grain size (Md), and the CUSUM curve of the MS. The notation OL represents the oxbow lake facies-dominated phase and OB for the overbank lithofacies-dominated phase. The depth scale is shared between the (a,b) panels.

Based on the ^{14}C data age–depth model and lithofacies data (Table 3.2) combined with the variation MS curve, cumulative sum (CUSUM) curve of MS, and median particle size (Md), the vertical subdivision of the Tövises bed sequence was conducted (Figure

3.1). The identified units represent identical floodplain environments with associated processes: abandoned channel, oxbow lake phase frequently intercalated with flood deposits. The bottom base sequence of circa 7500–6700 BP represents the oxbow infill sequence characterized by massive silt clay facies interbedded and overlaid with lacustrine gyttja of the coarse silt to clay organic-rich facies, which represent significant flood events. Intervals circa 2700–2000 BP and 500–0 BP are massive, rooted clay silt facies with downward decreasing root intensity representing distal overbank soil. The presence of plant wood debris, oxidized brown spots, and abundant shells characterize the layers at circa 2000–1000 BP and circa 850–500, with decreasing shell content towards the bottom; this layer represents lacustrine gyttja resulting from the infill of the abandoned channel. At the depth from circa 6700–5800 BP, the layer combines lacustrine gyttja and wetland histosol (peats) and is part of an interrupted oxbow lake infill phase. The massive organic-rich deposits (gyttjas) often contain terrestrial gastropods and brachiopod shells with abundant wood fragments. A wetland histosol sequence is characterized by peat facies (O) of varied thicknesses.

Table 3.2 The lithofacies association encountered in the Tövises bed core described following Hicks and Evans (Hicks and Evans 2022).

Lithofacies	Texture and Structures	Interpretations
Silt (SSm)	Massive silt, coarse-grained	Distal overbank
Mud (Ml)	Laminated silt and clay	Oxbow Lake flood layer
Mud (Mmo)	Massive silt, clay, OM	Oxbow lake infill
Mud (Mmr)	Massive silt and clay (rooted soil)	Distal overbank (soil)
Peat (O)	Massive peat, fibrous	Wetland histosol
Sand (Smf)	Massive, very fine-grained	Proximal overbank
Sand (Smo)	Massive (wood remains), very fine-grained	Proximal overbank

3.2.2   Grain-Size Distribution (GSD) and Endmember Modeling (EMM)

3.2.2.1   Grain-Size Distribution (GSD)

The GSDs of the sediments in the Tövises bed core consist of clay (<8 ϕ), silt (8 to 4 ϕ) with dominant values from 4.3 ϕ to 5.2 ϕ, and very fine sand (>4 to 3 ϕ) (Figure 2a). The silt content represents the dominant fraction with a content of 80.7%. The samples contained less clay (13%). The sand fraction represents the least dominant fraction and has a content of 6% (Figure 3.2b).

3.2.2.2 Endmember Modeling (EMM)

All the sediment samples collected from the core were subjected to EMM. The GSD analysis revealed several polymodal distributions ranging from 2.5 to 11 ϕ (Figure 3.3b). The samples were predominantly silt, with mean grain sizes fluctuating with depth. The optimum number of EMs was determined based on the inflection point of the determination coefficients (R^2) plot (Paterson and Heslop 2015) (Figure 3.3a). The four EMs model is the turning point in linear correlations (Figure 3.3a). The good of fit statistics demonstrates that these four endmembers offer great optimization of the grain-size subpopulations and R^2 (94% of the total variance); thus, the four EMs model meets the need for the least EM number with a high R^2. Figure 3.3b indicates the grain-size distribution of the selected EMs. Table 3.3 shows the endmember statistics; EM1 is characterized by moderately to well-sorted, fine skewed distribution between clay and very fine silt with a mean grain size of 7.45 ϕ. EM2 and EM3 are well to very well sorted, finely skewed coarse to fine silt with mean grain sizes 5.5 and 6.1 ϕ, respectively. EM4 is moderately sorted, finely skewed, and very fine sand to medium silt with a mean grain size of 4.65 ϕ. The proportional contributions of each endmember are 26.43% for EM1, 12.21% for EM2, 18.90% for EM3, and 42.46% for EM4.
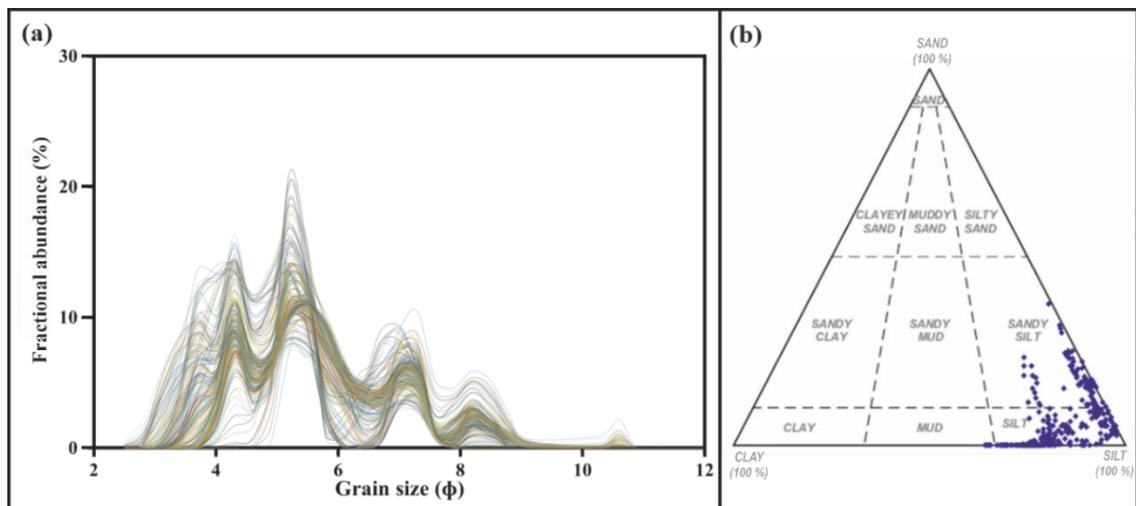
Figure 3.2 GSD characteristics of the Tövises bed core. (a) Overlay plot of the GSDs of all samples (n = 345), illustrating the polymodality of Tövises bed core sediments; (b) classification of grain-size composition in Tövises bed core, fine ternary plot (Folk 1954).

The mean grain size (Mz) relates to the overall grain size. The average sediment size is Mz, representing the index of depositional energy conditions. It exhibits local variations within the sediment core, ranging from silt to very fine sands; this variation in the Mz indicates fluctuations in the depositional energy. The correlation analysis shows a

positive correlation between Mz and EM4; therefore, EM4 can indicate the energy conditions of the deposition. The average EM4 of the Tövises bed sediments in the area points to the predominance of coarse silt, indicating a moderately low energy deposition condition. The endmember distributions in each identified layer of the Tövises bed core are calculated, and the results are presented in (Figure 3.4). The variation trend of average EM1 content is equal in all oxbow lake infill units (OL1–OL5) and higher than in all overbank units (OB1–OB5). EM2 and EM3's average content fluctuates in all units and appears equal for OL2 and OL5; meanwhile, EM2 is present in OB4 and OB5. The EM4 is present in all units, and its average contents in all overbank units (OB1–OB5) are higher than in the oxbow lake infill units.
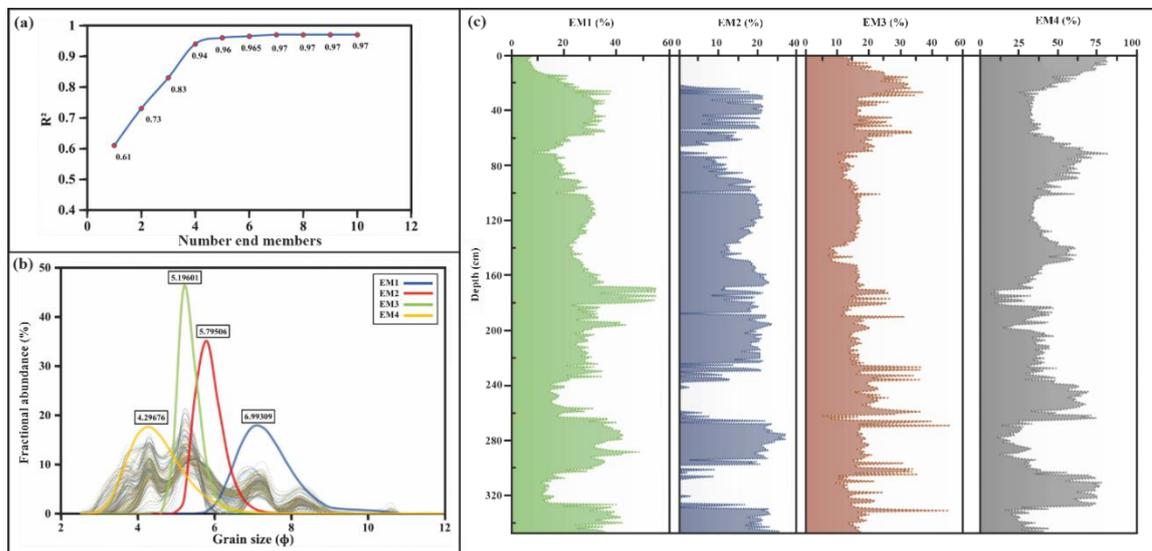
Figure 3.3 EMM results of sediments from the Tövises bed core. (a) The multiple coefficients of determination (R^2) function of the number of Ems. (b) Representative four EMs resulting from modeling. (c) The vertical distribution of the four EMs abundance of Tövises bed core.

Table 3.3 The statistical parameters of EMs of the Tövises bed core sediments.

Endmember	Mean Grain Size (Mz)	Sorting (σ)	Skewness (Sk)	Kurtosis (Kg)	Clay (%)	Silt (%)	Sand (%)
EM1	7.45	0.70	0.13	1.00	21.50	78.50	0.00
EM2	6.00	0.37	0.16	1.00	0.03	99.97	0.00
EM3	5.50	0.29	0.18	1.00	0.00	100.00	0.00
EM4	4.70	0.72	0.17	1.00	0.00	83.00	17.00

3.2.3 Loss-On-Ignition (LOI) and Magnetic Susceptibility (MS) Characteristics

The LOI record obtained from the core sediments shows distinct shifts in organic content that mark the different oxbow filling phases (Figure 3.1a), with fluctuations

within the oxbow lake infill facies-dominated phases associated with a gradual transition in sediment grain sizes.
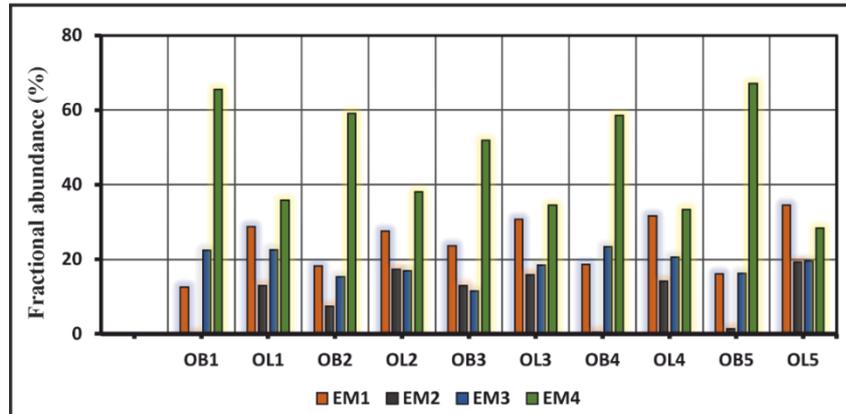
Figure 3.4 Average of endmembers abundance (EM1-4) of each core unit. OB = overbank unit, OL = oxbow lake unit.

In the Tövises bed sediments, the lower MS values range from $6 \times 10{-8}$ to $10 \times 10{-8}$ $m^3 \times kg^{-1}$, with an average value of $8 \times 10{-8}$ $m^3 \times kg^{-1}$, while the high MS ranges from 27 to $48 \times 10{-8} m^3 \times kg^{-1}$. The higher range of MS shows three distinct and prominent peaks; on the other hand, the low range exhibits fluctuation of increasing and decreasing trends (Figure 3.1b). The MS values change points at different intervals, and they are conformable with Md trends (Figure 3.1a,b); the highest value $(48 \times 10{-8})$ is between circa 3000 and 400 BP and between circa 3000 and 1700 BP, the value ranges $(5 to 19 \times 10{-8}^{m3} \times kg^{-1})$ and $(4$ to $14 \times 10{-8}$ $m^3 \times kg^{-1})$ at circa 1700–1000 BP, and the values at circa 400–1000 BP ranges $(19$ to $23 \times 10{-8}$ $m^3 \times kg^{-1})$.

3.3 Discussion

After cutting off from the main channel, the oxbow lake receives water and sediment delivered by a tie channel and overbank flow. The abandonment process starts with cutoff initiation, where the channel is separated from the meander loop and flows a new straight course; as the channel diverts, the plug bar forms, causing further blockage of the upper and lower junction of the oxbow lake (Collinson and Lewin 1983; Hooke 1995; Constantine et al. 2010). The regular channel discharge does not reach the developing abandoned channel (Toonen, et al., 2012). Suspended sediment accounts for most sediment loads in most river systems (Syvitski et al. 2005; Tooth 2013a). The subsequent high discharge removes the sediments deposited during the previous low discharge (Juez, et al., 2018).

The analyzed data show that the Tövises bed paleochannel evolved during the mid to late Holocene around 7748 ± 60–532 ± 15 BP, through an early lacustrine phase (oxbow lake) gradually infilled by lacustrine gyttjas, interbedded with wetland histosol, and periodically accumulating flood deposits. The lithofacies succession (Table 3.2) indicates different depositional settings and transport mechanisms. The overbank sequence of alternating dark layers of coarse silt and very fine sand represents flood sediments followed by finer sediment settling out of suspension after floods, respectively (Pannatier 1997). The organic-rich beds are likely to indicate the exposure of infill sediments to wetting and drought cycles which were probably part of the changing Holocene climate across Europe, where the different magnitude of changes was scattered in the content. For instance, during the mid-Holocene, Europe experienced the wettest period, with precipitation about 20 mm/year higher than the current. Around circa 8000 BP, warm and dry conditions prevailed in central Europe, with small-scale changes in temperature (Davis et al. 2003), while in the east cold and wet conditions were present; during the late Holocene, the climatic conditions were cold and wet in central Europe (Perşoiu et al. 2017) and northern Europe (Davis et al. 2003) and warm and dry in eastern Europe (Perşoiu et al. 2017)

The presence of peat indicates the accumulation of terrestrial organics (Hicks and Evans 2022). The peat-forming conditions are favored by preventing flood water accessibility to the accumulation site and mixing with clastic sediments (Moore 1989); warm and moist conditions are closely correlated with hydrologic factors (Moore 1989). However, the peat contains clastic sediments and is often intercalated with thin overbank facies indicating the lake transition into wetlands, probably caused by changes in lake and main channel connectivity (Marsh 1990). The peat layers at depths of 43, 224, 251, and 305 cm (Figure 3.1a) indicate consistent development of organic-rich sediments over a prolonged period. Moreover, peat layers between oxbow infill imply recurring drying up and reactivation of the oxbow lake. The fluctuations in organic content within the oxbow lake infill's facies-dominated phases are interpreted as sudden changes in the active channel vicinity due to abrupt migration through revived cutoffs in the main channel. The transition in sediment grain sizes within phases probably represents a gradual change in connectivity and lateral river migration. The high superimposed small-scale values probably reflect sedimentation during minor to moderate and more

significant floods, peat formation, inter-flood organic production, and other soil-forming processes.

The EMM of the GSDs of the last 8000 BP of the Tövises channel reveals valuable and independent information on the evolution and its alluviation process in the temperate mid-latitude GHP. The GSDs were separated into statistically robust EMs, each representing a distinct particle population assumed to be transported and deposited by a particular sedimentation process. Most of the suspension load in rivers originates outside the river channel and is delivered as overland flow; it can also originate from riverbank erosion (Tooth 2013b). EM1 is predominantly clay to very fine silt and relates to background fine sediments within the oxbow lake that occur irrespective of the river phase. Silt-sized particles are transported primarily as a suspended load, while sandy fractions are transported as a bedload, i.e., by traction or saltation (Gorsline 1984); the higher abundance of EM1 and EM2 is associated with the oxbow lake infill phase (OL1-5). In addition, the high MS values associated with EM1-2 and OL1-5 (Figures 3.4 and 3.5) indicate wet and reducing periods followed by drier periods resembling lacustrine conditions that are periodically exposed to dry seasons. MS reflects the mineral contents of the sediments; ferrimagnetic minerals contribute the most to the MS in the sediments, however, besides iron-rich clay minerals (Thompson and Oldfield 1986). MS is higher in the sediments and soils than in the parent rocks and topsoil than in the sub-soils (Thompson and Oldfield 1986). It results from the in situ alterations of the iron oxides from an antiferromagnetic form to the ferromagnetic form; this alteration requires wet conditions (Mullins 1977). During wet conditions, the anaerobic breakdown of soil organic matter leads to the reduction of iron; during the subsequent dry period, the iron is re-oxidized to form maghemite (Tite 1975). The intermediate components, EM2 and EM3, are mainly silt and are probably deposited during low to moderate flooding events.

The coarsest component, EM4, is coarse silt to very fine sand that is interpreted as coarse sediments entrained in suspension and transported to the overbank during moderate to significant flooding. However, fine sand fractions in most rivers are transported as a suspension load during the flood (Thompson and Oldfield 1986; Brownlie and Taylor 1981). The rivers transport sediments into oxbow lakes during short flooding periods; oxbow lakes develop distinctly between floods under more

stable lacustrine environmental conditions. The flood events are documented in preserved sediments (Gasiorowski and Hercman 2005).
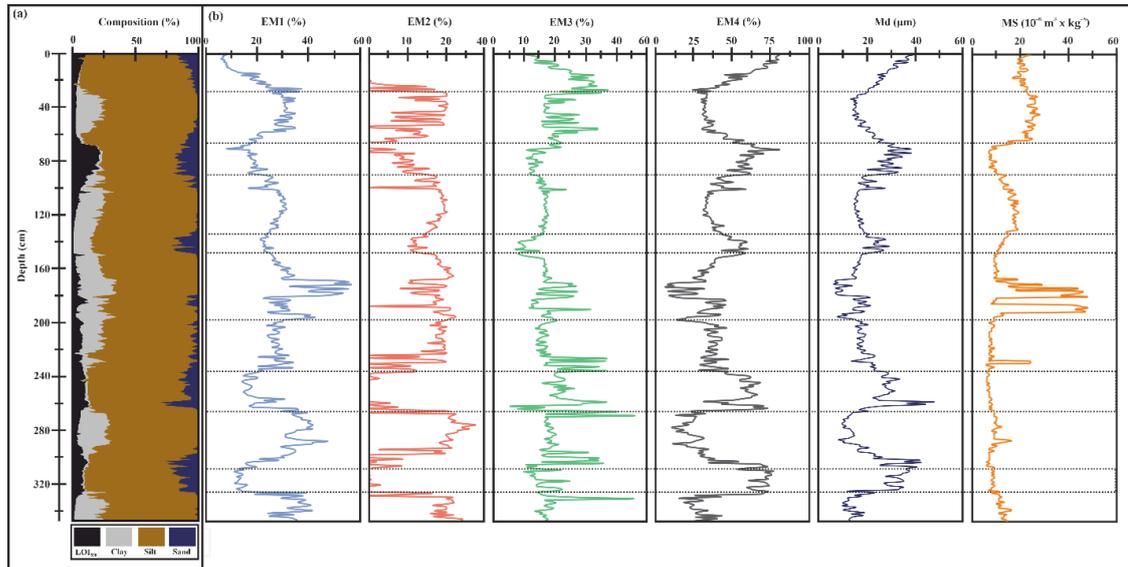

Figure 3.5 Sedimentary characteristics of the Tövises bed core. (a) Grain-size composition and LOI data. (b) Grain-size endmembers (EM1–4), magnetic susceptibility, and median particle size (Md) curves. The depth scale is shared between the (a,b) panels.

This is supported by the identified lithofacies (Table 3.1), where the overbank facies-dominated phase (OB1–OB5) is associated with a higher abundance of the EM4. The EM4 was dominated by the intensity and transported by river flood currents, with the increase in the proportion of EM4 with lower MS values (Figure 3.5) reflecting increasing flood strength and vice versa. Mechanisms by which sediments can be deposited during overbank flow are a function of grain size and sediment amount (Pizzuto 1987; Knighton 1998; Pierce and King 2008); the coarse sediments are deposited in the channel proximity and transported by traction currents, in this case as graded suspensions (Eltijani et al. 2022c), while the finer sediments are deposited at a distance as quiet water sedimentation as uniform suspensions (Pierce and King 2008; Eltijani et al. 2022c). The silt fractions of the EM4 are therefore considered to be deposited as distal overbank where the positive correlation EM3 showed a marginal correlation with other endmembers and the coarser percentile (D95) (Figure 3.6), indicating that EM3 is not sensitive to intensity changes in deposition processes during the lake filling. However, it could represent the distal over bank deposits. The indication of EM4 as a flood indicator is supported by its positive correlation with OM; the high carbon content can result from the deposition of carbon-rich sediments during floods

(Pinay et al. 1992). however, fine sediments with high OM can be attributed to the low and continuous sedimentation at a distance from rivers (Cierjacks et al. 2011).
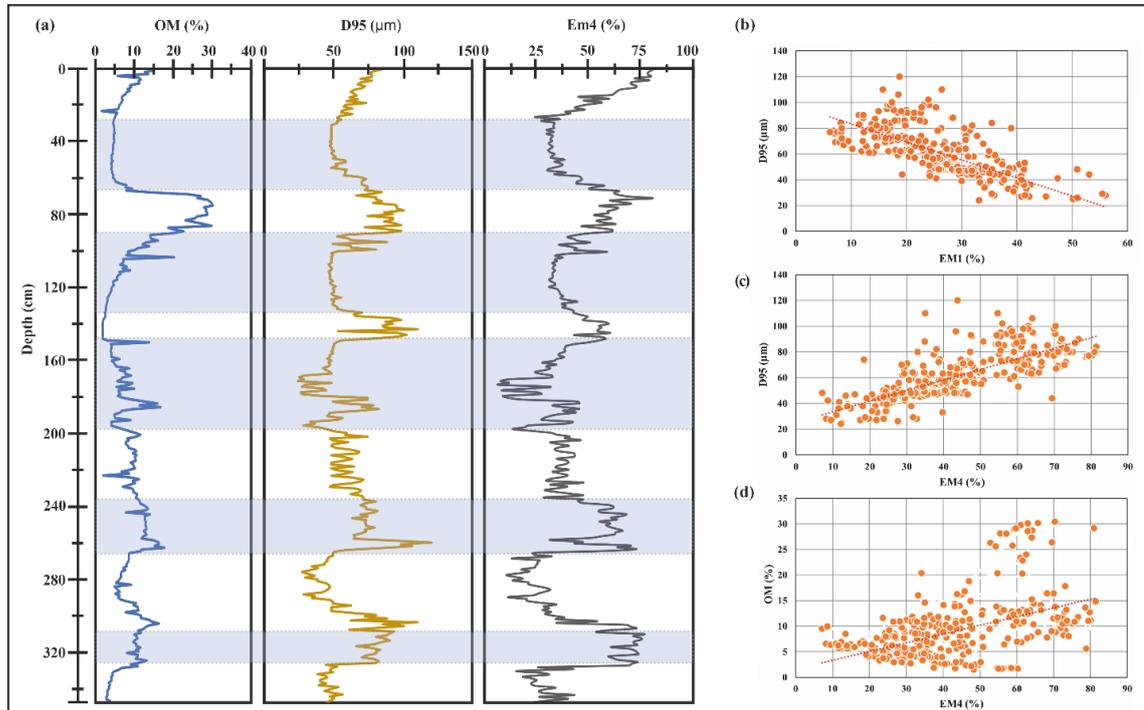
Figure 3.6 Sedimentary characteristics of the Tövises bed core. (a) OM, D95, and EM4 (b,c) linear correlations of D95 with EM1 and EM4, respectively. (d) Linear correlations of D95 with OM and EM4.

Comparing the coarse-grained components (EM4) with the NGRIP δ18O, annual precipitations, and mean annual precipitation records since 8000 BP, it is found that the EM4 records several cold and warm events during the relatively warmer early Holocene and the cooler late Holocene (Figure 3.7). Warm events with high precipitation are about circa 7500, 6000, 4700, 4000, 3600, 3000, 2500, 1400, and 700 BP. The warm event of the late Holocene was in 2000 BP, a colder event with high precipitation was in 1700, and a colder event with low precipitation was around 300 BP. The fluctuations are recorded at circa 7300–6700, 6000–4800, and 2700–1500 BP.
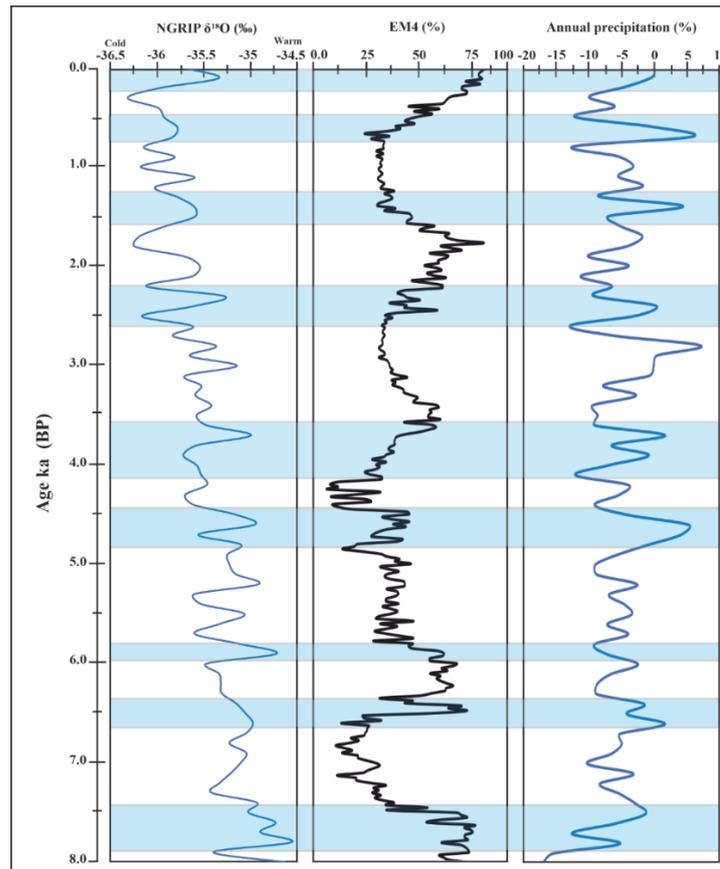
Figure 3.7 Comparison of the coarse grain-size endmembers (EM4) of Tövises bed core with annual mean precipitation as simulated by CCSM3 in TRaCE21ka experiment, averaged for "Europe" domain in Paleoview in Europe (Fordham et al. 2017) and the NGTRIP ∂ 18O (North Greenland Ice Core Project members 2004) of the last 8000 BP.

3.4   Conclusion

The endmember modeling reveals invaluable insight into the alluviation history of the abandoned channel. Based on paleoenvironmental proxy indicators such as the change of grain-size endmembers, LOI, and MS, this study showed:

The age–depth model indicates that the core sequence represents sediment accumulation since 8000 BP. The grain-size composition, lithofacies, LOI, and MS, reveal varying climatic conditions of wet and dry periods. The Tövises bed abandoned channel evolved through an early lacustrine phase (oxbow lake), gradually infilled by lake gyttjas, interbedded with wetland histosol, and periodically accumulated flood deposits.

The parameterized endmember modeling resulted in four endmembers (EM1-4), indicating different deposition conditions: EM1 is a clay to very fine silt component of the oxbow lake filling, representing the lake phase. EM2 and EM3 represent the

intermediate component representing mainly silt transported resulting from a moderate flood; EM4 is a fraction of material transported during a significant flood when the river discharge was relatively high; it represents the overbank deposition phase.

This study demonstrates that partitioning the grain-size distribution of abandoned channel sediments into statistically robust grain-size endmembers provides various quantitative proxies that help to model the complex paleoenvironmental changes and climate history in the region during the Holocene.

4 Characterizing Sedimentary Processes in Abandoned Channel using Compositional Data Analysis and Wavelet Transform

Abdelrhim Eltijani[*1] ; Dávid Molnár[1]; János Geiger[1]

Abstract

Grain size distribution (GSD) is essential for characterizing the deposition process. However, it is necessary to consider its compositional constraint to comprehend the statistical distribution of size fractions within the sediments. Compositional data analysis (CoDA) and wavelet transform (WT) represent alternative methods beyond traditional approaches, e.g., probability density function (PDF). This paper introduces a quantitative approach for characterizing Quaternary depositional and environmental changes using abandoned channel infill sediments. The proposed approach integrates CoDA and WT to thoroughly comprehend the depositional patterns observed in abandoned channels and the underlying environmental variability. The depositional model constructed based on CoDA showed coarsening-upward sequences, suggesting a periodic connection between the main channel and the oxbow lake. Three scales of cycles consistent with the depositional model constructed using CoDA were identified based on WT: small, medium, and large-scale cycles of processes. The large-scale cycles indicate the main depositional events, while the medium and small scale reflect the variation within and during deposition. CoDA and WT demonstrate excellent potential in characterizing the GSD and interpreting oxbow lakes' deposition and sedimentation processes.

Keywords: Sedimentary cycles, Compositional data, Wavelet transform, abandoned channel

4.1 Introduction

Meandering channels are omnipresent features of the quaternary fluvial system. Abandoned channels are distinct meandering sub-environments that result from complex channel migration over the floodplain (Shen et al., 2021); they develop through consecutive processes of cutoff initiation, plug bar formation, and channel disconnection (Costigan and Gerken, 2016; Sedláček et al., 2019; Ielpi et al., 2021). Therefore, understanding abandoned channels' sedimentary processes can contribute significantly to understanding Quaternary fluvial systems. The infill history of the abandoned channel can be observed in vertical sedimentary succession (Hicks and Evans, 2022). The sediments of the abandoned channel contain paleoenvironmental proxies that help interpret the depositional processes (Toonen et al., 2012). The infill record exhibits vertical variations corresponding to interconnected sedimentation mechanisms of distinct ranks. Hence, deciphering sedimentary cycles is necessary to understand the hierarchy and extent of sedimentation controls. Grain size distribution (GSD) is a proxy that provides insights into the transport processes and depositional history of abandoned channels. GSDs of fluvial sediments are generally polymodal, with many overlapping individual grain size populations (Collinson and Lewin 1983; Hooke 1995; Eltijani et al., 2022a) deposited by different processes within the same environment (Sun et al., 2002).

GSDs are quantified commonly using log-normal distribution, probability density function (PDF), and multivariate statistics, e.g., principal component analysis (PCA) and cluster analysis (CA). While log-normal results in identical log-normal distribution coefficients for different polymodal GSDs, as GSDs are regularly not log-normal (Roberson and Weltje 2014), PDF may obscure significant variability in GSD. Although multivariate statistics can solve the variability problems, the compositional constraint of GSD requires a prior mathematical treatment because the sum of percentages of the size fractions of GSD sum 100% making a closed system thus, carrying relative information on the whole distribution (Flood et al., 2015).

To address the constraint of GSD, (Eltijani et al., 2022b)employed compositional data analysis (CoDA) and Euclidean data analysis (EuDA) coupled with PCA and CA to examine GSD of a Quaternary abandoned channel located in the southeast Great Hungarian Plain. In the CoDA approach, the raw GSDs were transformed using the centered log-ratio (clr) transformation. In the EuDA approach, raw GSDs were regarded

as Euclidean data, needing no prior transformation. The result was distinct models (i.e., EuDA and CoDA models) that are statistically and sedimentologically plausible; their contributions to the conceptual model were substantially distinct. Our interest was to construct depositional models with statistical significance and geological interpretability. It is worth noting that the models constructed using CoDA and EuDA approaches were deemed plausible based on the statistical assessment and sedimentologic criteria. Nonetheless, the unique contribution of each approach to the conceptual models led to notable differences in the temporal reconstruction of the depositional models.

Although the CoDA helped elucidate the constraint imposed by GSD, and PCA, CA enabled the consideration of the whole GSD and revealing genetic sequences; a comprehensive insight into the evolution requires a quantitative analysis of the vertical sedimentation pattern (i.e., orders and magnitudes) in spatial and temporal scales. This involves characterizing cyclicity manifested on various orders of magnitude in response to changes in processes that are evident in GSDs, ranging from laminations to beds of centimeter scale.

Cyclic sedimentation is a pattern where the sedimentary sequences develop regularly and repeatedly (Schwarzacher 2000; Michael Allaby 2020). It can be generated by external factors and changes in provenance, e.g., climate and tectonic activity (allocyclic), or by internal factors triggering redistribution of energy and sediment transport within the system (autocyclic) (Michael Allaby, 2020). The latter can be represented by meandering channel processes that eventually form an abandoned channel and further influence its sedimentary infills.

Time series analysis is commonly used to discover a steady pattern in regularly measured datasets for identifying and quantifying processes (Gossel and Laehne, 2013). The traditional time series models are dependable when applied to long series with hundreds of unreadily available observations that require significant effort to reduce modeling uncertainty. A conventional approach to extract cyclicity uses Fourier transform. The drawback of the FT is that it captures the frequencies over the whole signal; the analysis window cannot capture features in the signals that are either longer or shorter than the window (Prokoph and Patterson, 2004). The wavelet transforms (WT) can address the limitations of the FT by enabling the characterization of shorter

time series, multiscale, non-steady processes. WT can simultaneously extract local and temporal information, including periods of cyclic events in spatial and time domains (Saadatinejad et al., 2011). Researchers have widely explored the use of WT to interpret geological data, for instance (Perez-Muñoz et al., 2013; Kadkhodaie and Rezaee, 2017; Duesing et al., 2021; Li et al., 2022). The commonly employed wavelet techniques are continuous wavelets transformation (CWT) and discrete wavelets transformation (DWT), and the Morlet family is the most used CWT family in earth science. WT can detect gradual and abrupt changes in GSDs and interpret the associated transport and deposition processes.

This paper presents a statistical methodology for analyzing sedimentary time series using GSD data from a Quaternary abandoned channel in the southeast Great Hungarian Plain. Our approach utilizes CoDA coupled with principal component analysis (PCA) and cluster analysis (CA) to construct a model with statistical significance and geological interpretability. The study showed that the CoDA provides a statistically valid, sedimentologically interpretable depositional model. It also demonstrated the application of wavelet transforms (WT) in quantifying the related patterns of transport and deposition processes. The study aimed to construct a stochastic model that provides a suitable mathematical representation of the alluviation history of the abandoned channel throughout the mid-late Holocene in the southeast Great Hungarian Plain (GHP).

4.2   Results

4.2.1   Results of the CoDA modeling

The application of PCA on clr-transformed data yielded two principal components that accounted for 83.22% of the total variance (Figure 4.1a). These two principal components represent two distinct depositional processes: the dominant process is indicated by PC1, which explains approximately 60% of the variance, and the less dominant process is indicated by PC2, accounting for around 22% of the variance. As illustrated in Figure 3b, PC1 exhibits a positive correlation with fine fractions (i.e., clay to medium silts) and a negative correlation with coarser fractions (i.e., medium silt and very fine sand). On the other hand, PC2 shows a positive correlation with the coarse fraction (i.e., medium silt to very fine sand) and a negative correlation with the finer fractions (i.e., clay to fine silt). Thus, the primary process influences the fine fractions,

while the second process primarily affects the coarser fractions, indicating the presence of two distinct processes in the oxbow lake.

The PC scores were subjected to cluster analysis, partitioning the sequence into four groups corresponding to specific grain size fractions. The successes of clustering were statistically significant, as indicated by discriminant analysis (DA), with ~ 95% percent correct (Table 4.1). These groups corresponded to specific depositional processes, as inferred from CM diagram (Figure 4.2a). The clusters identified in the analysis were associated with different types of suspension transport: uniform suspensions (SR) and graded suspensions (QR) (Figure 4.2b). Three types of QR were described: FG-QR, MG-QR, and CG-QR. These suspensions primarily showed cycles of the coarse CG-QR and MG-QR (Figure 4.2a), indicating a periodic connection between the oxbow lake and the main channel (Eltijani et al., 2022b). These suspensions primarily showed cycles of the coarse CG-QR and MG-QR (Figure 4.2b), indicating a periodic connection between the oxbow lake and the main channel.
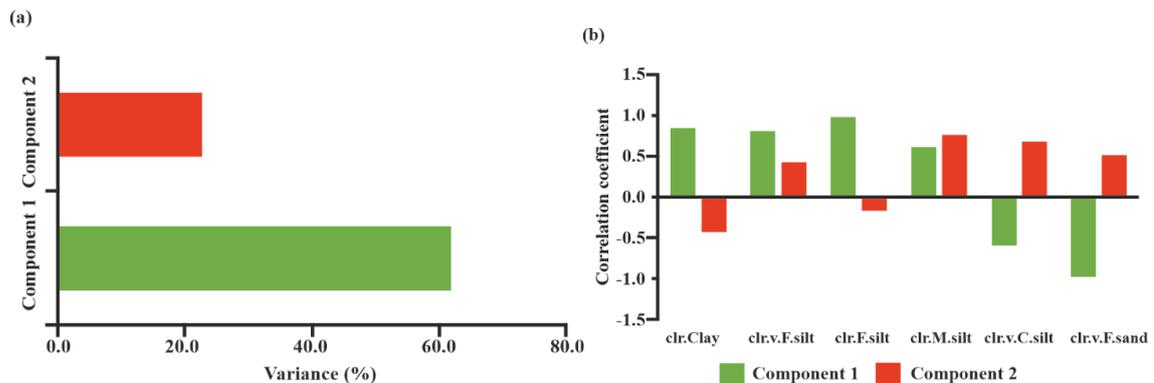
Figure 4.1 a. Scree plot of the variance % explained by the first and second principal components (PC1 and PC2). b. Correlation coefficients (loadings) of the PC1 and PC2.

4.2.2   Wavelet transform (WT)

The results of CWT based on the Morlet family applied to sand content of 340 cm intervals are presented in Figure 5. Due to the finite length of the time series, (edge effects) occur at the two ends of the wavelet spectrum, resulting from stretching the wavelets beyond the sand content series intervals. The information represented by the scalogram is reliable within the cone of influence (Figure 4.3). The color patterns of the scalogram indicate the energy magnitude of each wavelet coefficient; the blue patterns

correspond to high energy, while the red patterns indicate low energy. The small scales correspond to energy in the input signal at higher frequencies.

The medium and large scales correspond to energy in the input signal at medium and lower frequencies. The CWT for the sand fraction (Figure 4.3) shows quasi-periodic small, medium, and large-scale cyclicities. The top of the figure represents detailed, small-scale periods, the middle indicates medium-scale periods, and the bottom represents a smoothed overview of more extended-scale periods.

The edge effect is much higher in the bottom than in the middle and upper parts. Thus, this error significantly influences the large-scale periods. Consequently, the wavelets of the entire time series cannot be considered in the case of prolonged periods, only the following intervals: from ~ 100 – 240, 80 – 260, 40 – 290, and 25 – 320 cm. The sand content series (Figure 4.4) shows two large, medium, and three small-scale cycles. Figure 4.5 clearly shows those depth intervals where these cycles are significant.

The data in Figure 4.4 demonstrates that significant cycles of 64 cm occur at the top of the sequence, with less significant cycles at the bottom. The middle of the sequence contains significant cycles of 38 cm, specifically within the range of 140-100 cm and 340-260 cm. Medium scale cycles of 22.5 cm are significant at 180-120 cm and 320-240 cm, while 13 cm cycles are significant at 300-240 cm. The small-scale cycles show the significance for 8 cm cycles at 300-240 cm, while 4.8 cm and 2.5 cm cycles recur significantly throughout the entire sequence. These recurring cycles suggest a predictable pattern in the sequence, which is probably greatly influenced by the autogenic processes of the river system.
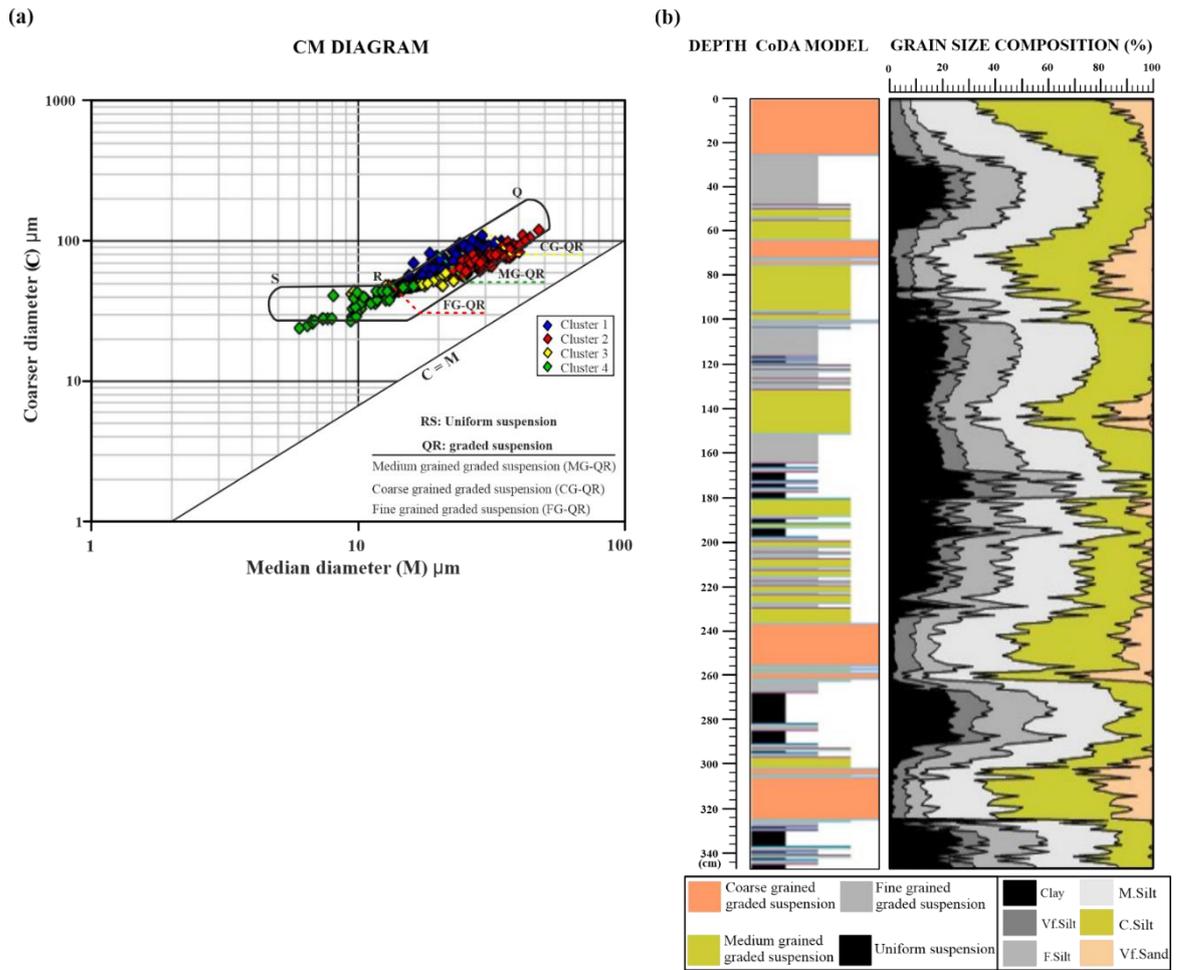
Figure 4.2 a. The relations between the CoDA clusters and their corresponding depositional processes in the CM diagram and b. The vertical depositional model was constructed based on CoDA; modified after Eltijani et al. 2022b).
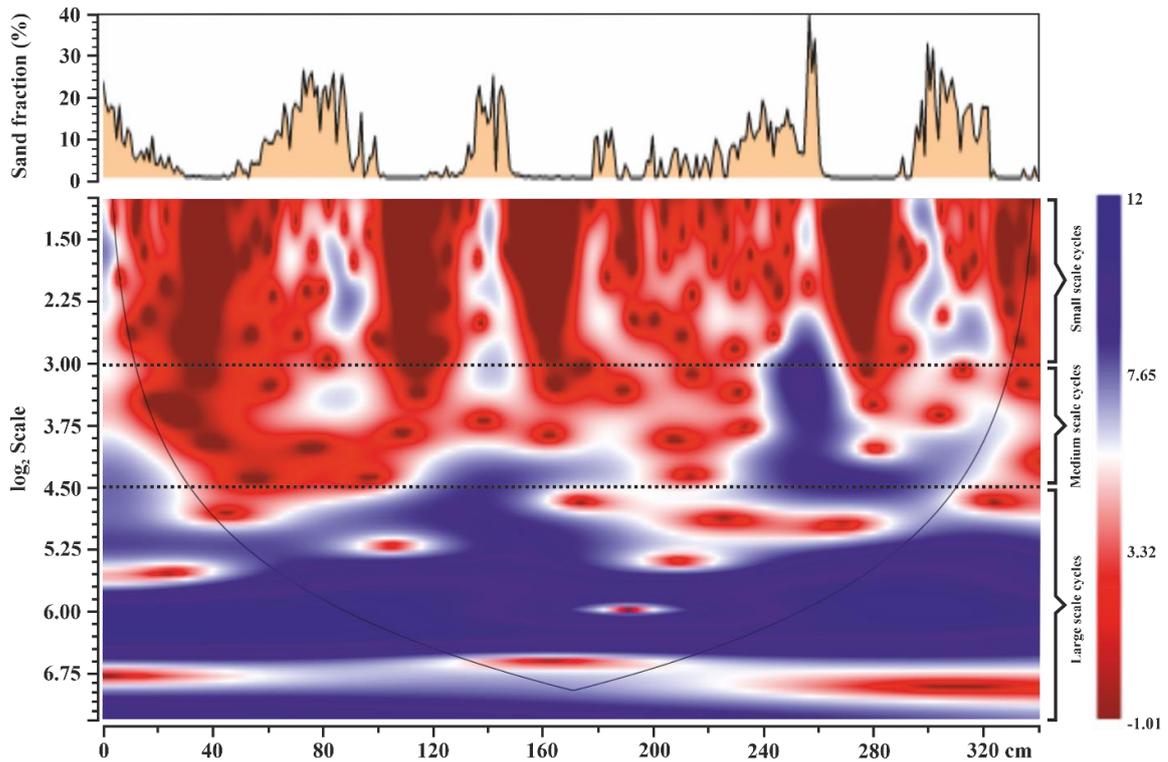
Figure 4.3 Represents the analyzed sand content series and its Morlet scalograms representation; the solid line envelope represents the cone of influence. The horizontal dashed lines separate the different scales of details provided by the CWT.

4.3   Discussion

Employing a CoDA approach in combination with PCA and CA offers the advantage of analyzing the entire GSD, which is not achievable using PDF methods. The fundamental concept is that PDF methodologies entail inherent incongruities besides the lack of sedimentological rationale to support the use of model coefficient parameters for assessing efficacy (Flood et al. 2015). The CoDA model uses an external sedimentological criterion CM diagram, which validates the approach and enables the characterization of two primary suspension processes: RS and QR, with QR further categorized into CG-QR, MG-QR, and FG-QR. This partitioning results from the interaction between overbank depositional processes and abandoned channels' infill sediment (Citterio and Piegay 2009). Thus, it is likely that CG-QR denotes the stratification resulting from floodplain mechanisms and subsequently integrated into suspended sediments or bedload sediments conveyed by the flood channel. These two alternatives imply the occurrence of high-magnitude flow events.
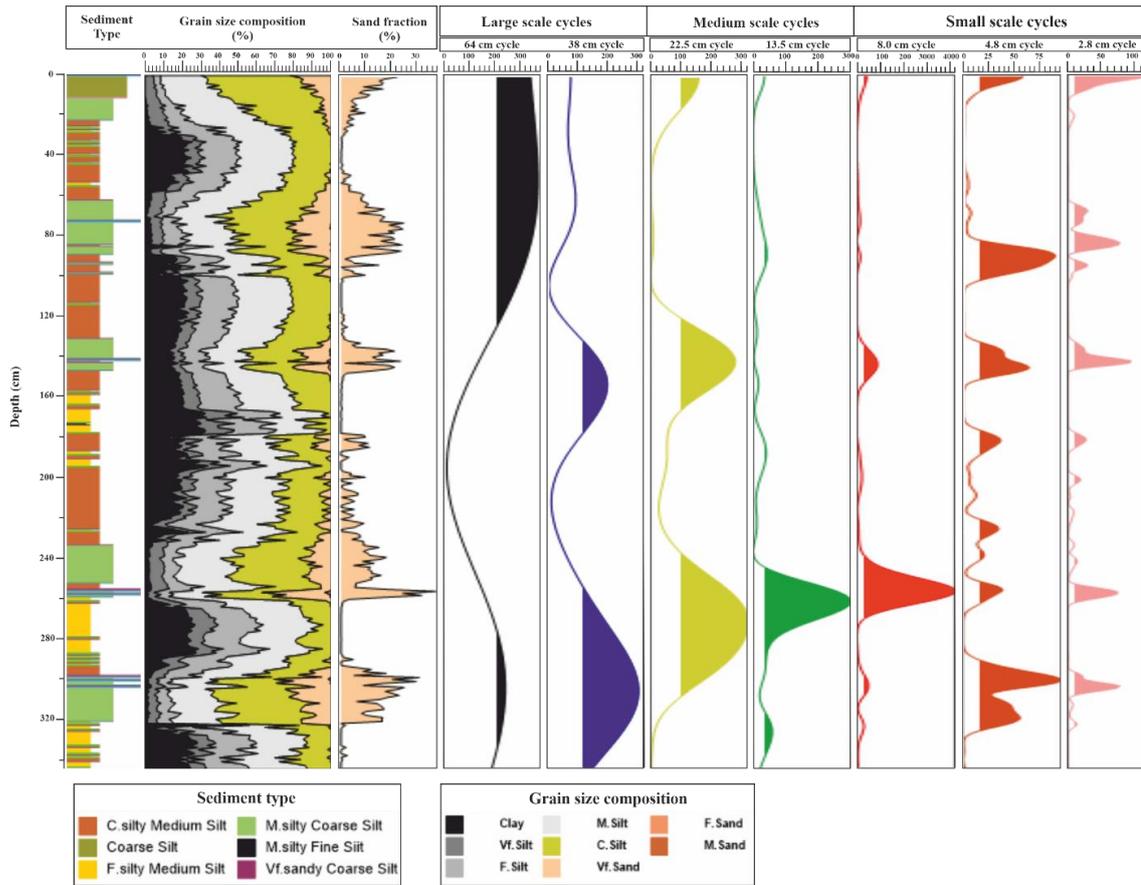
Figure 4.4 Represents the analyzed sand content series and representation of the revealed cycles; at small, medium, and large scales.

Analyzing the sand fraction through WT can identify the intensity and the recurrence degree of these transport depositional events and processes. The large-scale cycles depict the general trend of the sedimentation processes and environmental conditions. In contrast, the medium-scale (i.e., intense ~ 22.4, 10, and 8 cm cycles between circa 7 and 6 ka yrs BP and weaker ~ 10 and 8 cm cycles between circa 4 – 2 ka yrs BP) (Figure 4.6a) and small-scale cycles (i.e., ~ 4.5 and 2.8 cm at the bottom of the sequence during the circa 8 -7.5 Ka yrs BP and ~ 4.8 cm at the top of the sequence between circa 2.5 – 2 ka yrs BP) (Figure 4.6b) may reflect the local variability of the transport agent.
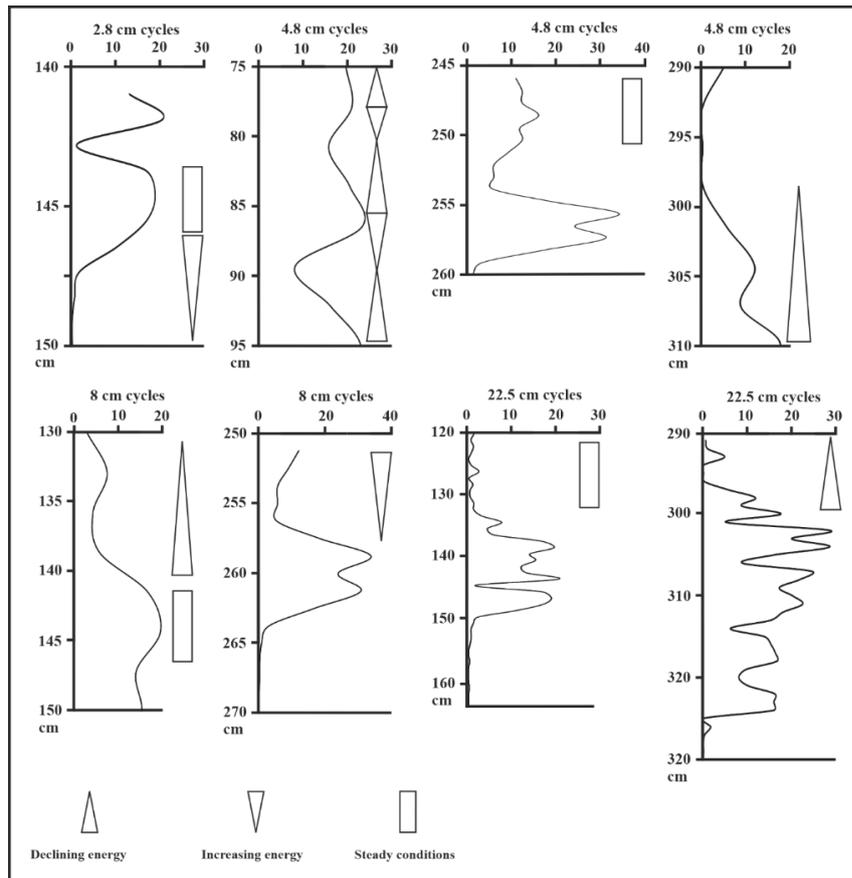
Figure 4.5   Represents the medium and small-scale cycles of sand content based on average power spectrum and significance intervals.

The background-graded sedimentary sequences are commonly deposited during short-lived events (Hiscott, 2003), probably during flooding. These cycles are characterized by upward, downward decreasing, and fluctuation in the sand content (Figure 4.5). The upward decrease in the sand content indicates the upward declining energy of the deposition. During this period, turbulence dominates the bottom with a progressive decrease in the energy for all sediment fractions (Hiscott, 1994). The upward increase of sand content indicates the upward increase in the available energy during the depositional event. This trend may result from the grain-to-grain collisions caused by the shearing of the overriding bottom current or by the downward percolation of the finer sediments between the coarser grains (Bagnold, 1956). The cycles where variations in sand content repetition show stratified or laminated depositional sequences indicating a recurring short-year flood event.

Comparing these results with the CoDA model (Eltinaji, 2022b, 2022a), the small-scale cycles (Figure. 4.6b) correspond to CG-QR and MG-QR in the CoDA model indicating

cyclic deposition as QR through the active channel or during the overbank flow. The possibility of deposition during a flood is supported by the correlation with overbank (OB) units in the Endmember (EMM) model (Figure 4.6a and b) (Eltijani et al., 2022a). As the analyzed time series is a CG-QR and MG-QR, the deciphered cycles represent peak depositional energy. The hydraulic settling of the size fractions within QR (in this case CG-QR, MG-QR, and FG-QR) is controlled by a similar process that is partially independent (Passega 1964); two types of currents, competent and incompetent, support the coarser suspension particles. As the competent current velocity is reduced gradually, the coarse to medium silt is deposited with a small amount of clay.
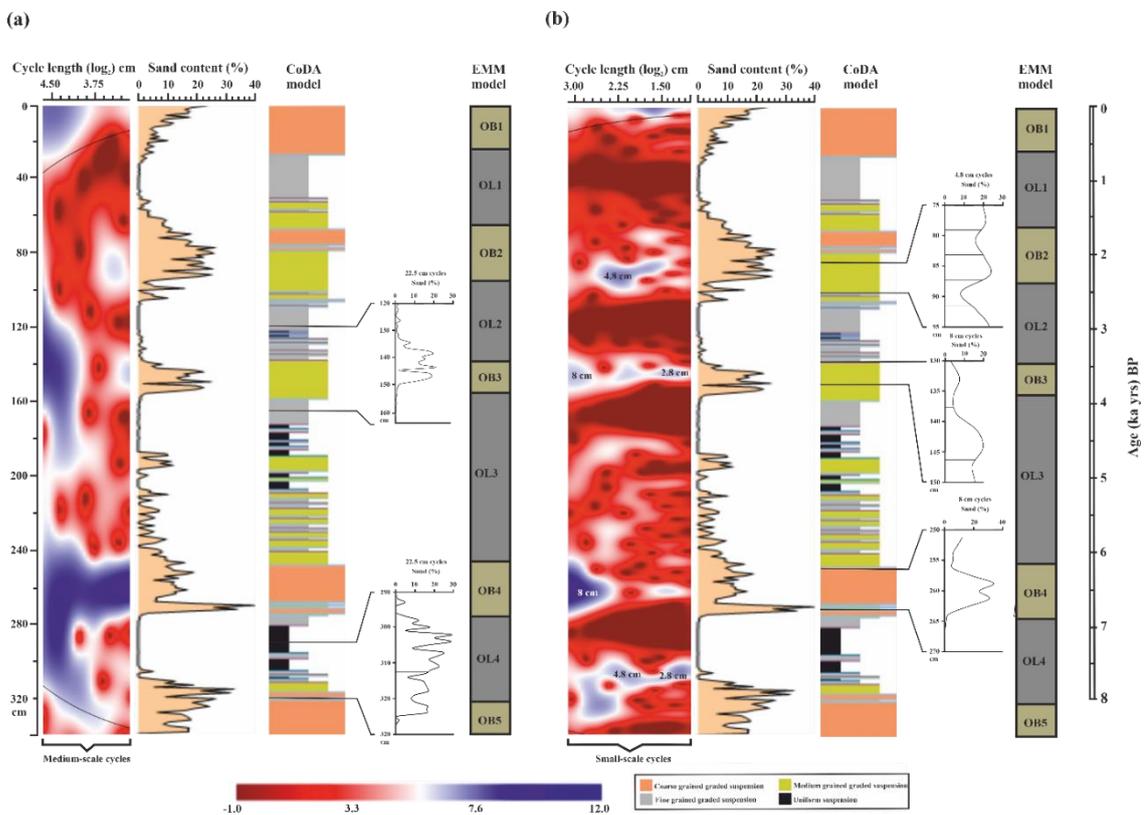
Figure 4.6 Represents the medium and small-scale cycles of sand content with their Morlet representation and its Morlet scalograms at different scales, average power spectrum, and significance intervals. (a) represents part A of the small-scale, and detailed cyclicity (b) represents part B of the small-scale cycles.

4.4 Conclusion

This paper aims to provide a quantitative description of the depositional processes based on CoDA and the wavelet transform of sand content from the Tövises bed core. The CoDA eliminated the constraint of the GSDs before the applications of PCA and CA;

PCA and CA characterized the depositional processes based on the whole distribution. Wavelet analysis provided valuable information for the cyclic alluviation processes of the abandoned channel analyzed from grain size distribution. The study shows that the adequate explanation of the wavelet decomposition and Morlet scalogram helps identify the thickness, depth, and level of significant high-energy depositional events.

The study revealed that the deposition occurred as uniform and graded suspension loads in multiple stages interrupted with bottom current loads. In the case of CoDA, they corresponded to coarsening-upward sequences, suggesting the periodic sediment flux to the oxbow lake. In contrast, EDA corresponds to fining upward cycles, indicating the permanent weak traction flows with periods of decreasing energy.

This finding shows excellent potential for using CoDA, coupled with PCA and CA, to interpret the sedimentation processes of abandoned channel sediments with satisfactory statistical significance and geologic interpretations. However, the reliability of this approach should be cross-checked with sedimentological and geological criteria as it cannot consistently deliver a significant result.

Comparing these cycles with the CoDA model and EMM, the results indicate that cycles can relate to higher energy conditions of different magnitudes. Significant floods occurred at the bottom of the sequence between circa 7 – 6 ka yrs BP, while moderate flood events can be seen at the top of the sequence during circa 3.5 ka yrs BP; how the middle of the sequence between circa 6 – 4 ka yrs BP witnessed week flood events.

5 Thesis Summary

Grain Size Distribution (GSD) serves as a proxy for understanding past environmental conditions and deposition processes. Previous research has employed statistical parameters like median, mean, sorting, and skewness to interpret depositional conditions from GSDs. This research applied the concept of compositional data analysis (CoDA), endmember modeling (EMM), and wavelet transformation (WT), which provided a comprehensive understanding of the abandoned channel alluviation process by allowing the entire GSD to be considered and analyzed in spatial and temporal domains.

This thesis delves into utilizing grain size distribution (GSD) analysis in conjunction with statistical techniques to comprehend depositional environments and processes within oxbow lake sediments. Examining sediment samples from a Great Hungarian Plain oxbow lake, the aim is to comprehend sedimentological processes and alluviation history by characterizing the grain-size distribution (GSD) through compositional data analysis and endmember modeling, and wavelet transformation. In general, The Tövises channel sequence spans 8000 years, displaying shifts in organic content, indicating wet and dry phases. Sediments include gyttjas, histosols, and flood deposits, representing channel evolution through various phases.

The study involves analyzing sediment samples from an oxbow lake in the Great Hungarian Plain. The workflow included calculating various percentiles from cumulative grain size distributions and transforming these data using clr. Two datasets were generated: one with frequency percentages and another with clr-transformed percentages of grain size fractions. PCA is applied to both datasets to identify principal components that capture the most variance. CA is then employed to cluster samples based on their principal component scores. LDA assesses the validity of these clusters. The results are interpreted in terms of depositional processes and environmental changes.

The research emphasized the employability of the centered log-ratio (clr) transformation, Euclidean data analysis, and multivariate methods "Principal Component Analysis (PCA), Cluster Analysis (CA), and Linear Discriminant Analysis (LDA)" to scrutinize GSDs. PCA identifies key components capturing variance in both datasets, corresponding to two distinct sedimentation mechanisms, impacting distinct grain sizes. CA validated these mechanisms, by differentiating suspension types linked

to channel- main system connections. LDA validated these clusters, yielding insights into depositional processes and environmental shifts.

Results suggested that clr-transformed and non-transformed datasets yield insights into the lake's sedimentation history, identifying depositional processes linked to grain size variations. A temporal evolution model is provided, highlighting periods of weak and strong oxbow-main channel connections. Both datasets are valid and interpretable, though they offer distinct historical insights. However, the results must be validated with sedimentological and geological criteria to ensure their reliability and accuracy.

To showcase how endmember modeling of GSD, combined with other proxies like MS and LOI, can provide a comprehensive understanding of the sedimentological processes, paleoenvironmental changes, and climatic variations in the region over the Holocene period. Endmember modeling is used to extract meaningful information from the grain-size distribution. Four distinct endmembers (EMs) are identified: EM1 represents clay to very fine silt, associated with the oxbow lake phase; EM2 and EM3 represent intermediate components of mainly silt, likely deposited during moderate flooding; EM4 indicates coarse silt to very fine sand, transported during significant flood events. These endmembers help characterize the deposition conditions and provide insights into the changing river dynamics and sediment sources. Magnetic susceptibility (MS) and loss-on-ignition (LOI) analyses further support the findings. High MS values are linked to wet and reducing conditions, indicating lacustrine phases, while low MS values correspond to drier periods. The LOI data show fluctuations in organic content, suggesting variations in wet and dry cycles and climate changes.

The study also employed wavelet transform (WT) to analyze cyclicity in the depositional processes and events. WT, unlike the Fourier transform, captures both short and long-term cycles. The results reveal quasi-periodic cyclicities in the sand content data, providing insights into sedimentation processes over time. Large, medium, and small-scale cycles are identified, reflecting both external factors like floods and internal processes like channel dynamics.

In conclusion, the combination of CoDA, PCA, CA, and WT to analyze GSD data from an abandoned channel offered a statistically significant and geologically interpretable model of sedimentation processes. The findings highlight the presence of distinct depositional mechanisms and provide insights into cyclic patterns driven by both external and internal factors. The study contributes to understanding the history of sedimentation in abandoned channels and Quaternary fluvial systems.

6   Acknowledgment

7   References

Aitchison, J. 1986. "The Statistical Analysis of Compositional Data." *The Statistical Analysis of Compositional Data* 44 (2): 139–77. https://doi.org/10.1007/978-94-009-4109-0.

Aitchison, J., C. Barceló-Vidal, J. A. Martín-Fernández, and V. Pawlowsky-Glahn. 2000. "Logratio Analysis and Compositional Distance." *Mathematical Geology* 32 (3): 271–75. https://doi.org/10.1023/A:1007529726302.

Álvarez, Giselle, Bruno Sansó, Reinaldo J. Michelena, and Juan Ramón Jiménez. 2003. "Lithologic Characterization of a Reservoir Using Continuous-Wavelet Transforms." *IEEE Transactions on Geoscience and Remote Sensing* 41 (1): 59–65. https://doi.org/10.1109/TGRS.2002.808065.

Beierle, Brandon D, Scott F Lamoureux, Jaclyn M H Cockburn, and Ian Spooner. 2002. "A New Method for Visualizing Sediment Particle Size Distributions." *Journal of Paleolimnology* 27: 279–83.

Brownlie, Williamn R., and Brent D. Taylor. 1981. "Coastal Sediment Delivery by Major Rivers in Southern California. EQL REPORT NO. L7-C." *EQL REPORT NO. L7-C*, no. February.

Buccianti, Antonella, G Mateu-Figueras, and Vera Pawlowsky-Glahn. 2006. "Compositional Data Analysis in the Geosciences: From Theory to Practice." *Geological Society, London, Special Publications* 264.

Cierjacks, A, B Kleinschmit, I Kowarik, M Graf, and F Lang. 2011. "Organic Matter Distribution in Floodplains Can Be Predicted Using Spatial and Vegetation Structure Data." *River Research and Applications* 27 (8): 1048–57. https://doi.org/https://doi.org/10.1002/rra.1409.

Citterio, Anne, and Herve PiegayY. 2009. "Overbank Sedimentation Rates in Former Channel Lakes: Characterization and Control Factors." *Sedimentology* 56 (2): 461–82. https://doi.org/https://doi.org/10.1111/j.1365-3091.2008.00979.x.

Coconi-Morales, E., G. Ronquillo-Jarillo, and J. O. Campos-Enríquez. 2010. "Multi-Scale Analysis of Well-Logging Data in Petrophysical and Stratigraphic Correlation." *Geofisica Internacional* 49 (2): 55–67. https://doi.org/10.22201/igeof.00167169p.2010.49.2.113.

Collinson, J. D., and J. Lewin. 1983. "Alluvial Cutoofs in Wales and the Borderlands." *Modern and Ancient Fluvial Systems.* https://doi.org/10.1016/0037-0738(84)90031-9.

Constantine, José Antonio, Thomas Dunne, Hervé Piégay, and G. Mathias Kondolf. 2010. "Controls on the Alluviation of Oxbow Lakes by Bed-Material Load along the Sacramento River, California." *Sedimentology* 57 (2): 389–407. https://doi.org/10.1111/j.1365-3091.2009.01084.x.

Costigan, Katie H, and Joseph E Gerken. 2016. "Channel Morphology and Flow Structure of an Abandoned Channel under Varying Stages." *Water Resources Research* 52 (7): 5458–72. https://doi.org/https://doi.org/10.1002/2015WR017601.

David Knighton. 1998. "Fluvial Forms & Processes. A New Perspective. Don Mills, Ontario, Oxford University Press, 383 p. (ISBN 0-340-66313-8)." *Cahiers de*

Géographie Du Québec 43 (118): 148. https://doi.org/10.7202/022796ar.

Davis, B. A.S., S. Brewer, A. C. Stevenson, J. Guiot, J. Allen, H. Almqvist-Jacobson, B. Ammann, et al. 2003. "The Temperature of Europe during the Holocene Reconstructed from Pollen Data." *Quaternary Science Reviews* 22 (15–17): 1701–16. https://doi.org/10.1016/S0277-3791(03)00173-2.

Dietze, E., F. Maussion, M. Ahlborn, B. Diekmann, K. Hartmann, K. Henkel, T. Kasper, G. Lockot, S. Opitz, and T. Haberzettl. 2014. "Sediment Transport Processes across the Tibetan Plateau Inferred from Robust Grain-Size End Members in Lake Sediments." *Climate of the Past* 10 (1): 91–106. https://doi.org/10.5194/cp-10-91-2014.

Dietze, Elisabeth, Bernhard Diekmann, Stephan Opitz, and Kai Hartmann. 2013. "Quantifying Depositional Processes in Sediment Archives Using End-Member Modelling of Grain Size Data" 15: 2013.

Dietze, Elisabeth, Kai Hartmann, Bernhard Diekmann, Janneke IJmker, Frank Lehmkuhl, Stephan Opitz, Georg Stauch, Bernd Wünnemann, and Andreas Borchers. 2012. "An End-Member Algorithm for Deciphering Modern Detrital Processes from Lake Sediments of Lake Donggi Cona, NE Tibetan Plateau, China." *Sedimentary Geology* 243–244: 169–80. https://doi.org/10.1016/j.sedgeo.2011.09.014.

Duesing, Walter, Nadine Berner, Alan L. Deino, Verena Foerster, K. Hauke Kraemer, Norbert Marwan, and Martin H. Trauth. 2021. "Multiband Wavelet Age Modeling for a ~293 m (~600 Kyr) Sediment Core From Chew Bahir Basin, Southern Ethiopian Rift." *Frontiers in Earth Science* 9 (March): 1–15. https://doi.org/10.3389/feart.2021.594047.

Eltijani, Abdelrhim, David Molnar, and Janos Geiger. 2023. "Characterizing Sedimentary Processes in Abandoned Channel Using Compositional Data Analysis and Wavelet Transform." *GEM - International Journal on Geomathematics* 123: 1–16. https://doi.org/10.1007/s13137-023-00223-y.

Eltijani, Abdelrhim, Dávid Molnár, László Makó, János Geiger, and Pál Sümegi. 2022a. "Applying Grain-Size and Compositional Data Analysis for Interpretation of the Quaternaoxbow Lake Sedimentation Processes: Eastern Great Hungarian Plain." *Studia Quaternaria* 39 (2): 83–93. https://doi.org/10.24425/sq.2022.140885.

Eltijani, Abdelrhim, Dávid Molnár, László Makó, János Geiger, and Pál Sümegi. 2022b. "Application of Parameterized Grain-Size Endmember Modeling in the Study of Quaternary Oxbow Lake Sedimentation: A Case Study of Tövises Bed Sediments in the Eastern Great Hungarian Plain." *Quaternary* 5 (4). https://doi.org/10.3390/quat5040044.

Evans, David J A, and Douglas I Benn. 2014. *A Practical Guide to the Study of Glacial Sediments*. Routledge.

Feldman, Ronen, and James Sanger. 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge university press.

Fieller, N R J, E C Flenley, and W Olbricht. 1992. "Statistics of Particle Size Data." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 41 (1): 127–46.

Filzmoser, Peter, Karel Hron, and Clemens Reimann. 2009. "Principal Component Analysis for Compositional Data with Outliers." *Environmetrics* 20 (6): 621–32. https://doi.org/https://doi.org/10.1002/env.966.

Flood, R. P., J. D. Orford, J. M. McKinley, and S. Roberson. 2015. "Effective Grain Size Distribution Analysis for Interpretation of Tidal-Deltaic Facies: West Bengal Sundarbans." *Sedimentary Geology* 318: 58–74. https://doi.org/10.1016/j.sedgeo.2014.12.007.

Folk, R.L., Ward, W.C., 1957. 1957. "Brazos River Bar: A Study in the Significance of Grain Size Parameters. Journal of Sedimentary Petrology 27, 3–26.Pdf." *Journal of Sedimentary Petrology* 27 (1): 3–26. https://doi.org/10.1306/74D70646-2B21-11D7-8648000102C1865D.

Folk, Robert L. 1954. "The Distinction between Grain Size and Mineral Composition in Sedimentary-Rock Nomenclature." *The Journal of Geology* 62 (4): 344–59. https://doi.org/10.1086/626171.

Fordham, Damien A., Frédérik Saltré, Sean Haythorne, Tom M.L. Wigley, Bette L. Otto-Bliesner, Ka Ching Chan, and Barry W. Brook. 2017. "PaleoView: A Tool for Generating Continuous Climate Projections Spanning the Last 21 000 Years at Regional and Global Scales." *Ecography* 40 (11): 1348–58. https://doi.org/10.1111/ecog.03031.

Fredlund, Murray D, D G Fredlund, and G Ward Wilson. 2000. "An Equation to Represent Grain-Size Distribution." *Canadian Geotechnical Journal* 37 (4): 817–27.

Gary Nichols. 2009. *Sedimentology and Stratigraphy*. *European Journal of Soil Science*. 2nd ed. Vol. 61. Wiley-Blackwell. https://doi.org/10.1111/j.1365-2389.2009.01225.x.

Gasiorowski, Michał, and Helena Hercman. 2005. "Recent Changes of Sedimentation Rate in Three Vistula Oxbow Lakes Determined by 210 Pb Dating." *Geochronometria* 24 (January): 33–39.

Gorsline, Donn S. 1984. "Introduction to a Symposium on Fine-Grained Sedimentology." *Geo-Marine Letters* 4 (3–4): 133–38. https://doi.org/10.1007/BF02281693.

Gossel, W, and R Laehne. 2013. "Time Series Analysis with Sample Applications in Geosciences Applications of Time Series Analysis in Geosciences: An Overview of Methods and Sample Applications Time Series Analysis with Sample Applications in Geosciences." *Hydrol. Earth Syst. Sci. Discuss* 10: 12793–827. https://doi.org/10.5194/hessd-10-12793-2013.

Greenacre, Michael. 2019. "Variable Selection in Compositional Data Analysis Using Pairwise Logratios." *Mathematical Geosciences* 51 (5): 649–82. https://doi.org/10.1007/s11004-018-9754-x.

Greenacre, Michael. 2020. "Amalgamations Are Valid in Compositional Data Analysis, Can Be Used in Agglomerative Clustering, and Their Logratios Have an Inverse Transformation." *Applied Computing and Geosciences* 5 (September 2019): 100017. https://doi.org/10.1016/j.acags.2019.100017.

Gulagiz, Fidan Kaya, and Sahin Suhap. 2017. "Comparison of Hierarchical and Non-Hierarchical Clustering Algorithms." *International Journal of Computer Engineering and Information Technology* 9 (1): 6–14. www.ijceit.org.

Ha, Hun Jun, Tae Soo Chang, and Ho Kyung Ha. 2021. "Using End-Member Analysis to Determine Sediment Dispersal and Depositional Processes on the Heuksan Mud Belt, Southwest Korean Shelf." *Geo-Marine Letters* 41 (1): 1–13. https://doi.org/10.1007/s00367-020-00672-6.

Halkidi, M., Y. Batistakis, and M. Vazirgiannis. 2001. "Clustering Algorithms and Validity Measures." *Proceedings of the International Conference on Scientific and Statistical Database Management, SSDBM*, 3–22. https://doi.org/10.1109/ssdm.2001.938534.

Hicks, Jocelyn L., and James E. Evans. 2022. "Oxbow Lakes as Geological Archives of Historical Changes in Channel Substrate, Swan Creek, Toledo, Ohio (USA)." *Open Journal of Modern Hydrology* 12 (02): 32–54. https://doi.org/10.4236/ojmh.2022.122003.

Hiscott, R. N. 1994. "Traction-Carpet Stratification in Turbidites - Fact or Fiction?" *Journal of Sedimentary Research A: Sedimentary Petrology & Processes* 64 A (2): 204–8. https://doi.org/10.1306/d42681ad-2b26-11d7-8648000102c1865d.

Hiscott, Richard N. 2003. "Grading, Graded Bedding BT - Encyclopedia of Sediments and Sedimentary Rocks." In , edited by Gerard V Middleton, Michael J Church, Mario Coniglio, Lawrence A Hardie, and Frederick J Longstaffe, 333–35. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-3609-5_101.

Hooke, J. M. 1995. "River Channel Adjustment to Meander Cutoffs on the River Bollin and River Dane, Northwest England." *Geomorphology* 14 (3): 235–53. https://doi.org/10.1016/0169-555X(95)00110-Q.

Huang, Yun, Jule Xiao, Rong Xiang, Shengfa Liu, Somkiat Khokiattiwong, Narumol Kornkanitnan, Jiawei Fan, Ruilin Wen, Shengrui Zhang, and Jianguo Liu. 2020. "Holocene Indian Summer Monsoon Variations Inferred from End-Member Modeling of Sediment Grain Size in the Andaman Sea." *Quaternary International* 558 (December 2019): 28–38. https://doi.org/10.1016/j.quaint.2020.08.032.

Ielpi, Alessandro, Mathieu G A Lapôtre, Alvise Finotello, and Massimiliano Ghinassi. 2021. "Planform-Asymmetry and Backwater Effects on River-Cutoff Kinematics and Clustering." *Earth Surface Processes and Landforms* 46 (2): 357–70. https://doi.org/https://doi.org/10.1002/esp.5029.

IJmker, Janneke, Georg Stauch, Elisabeth Dietze, Kai Hartmann, Bernhard Diekmann, Gregori Lockot, Stephan Opitz, Bernd Wünnemann, and Frank Lehmkuhl. 2012. "Characterisation of Transport Processes and Sedimentary Deposits by Statistical End-Member Mixing Analysis of Terrestrial Sediments in the Donggi Cona Lake Catchment, NE Tibetan Plateau." *Sedimentary Geology* 281 (December): 166–79. https://doi.org/10.1016/j.sedgeo.2012.09.006.

John C. Davis. 2002. *Statistics and Data Analysis in Geology*. 3rd ed. John Wiley & Sons New York.

Juez, Carmelo, Marwan A. Hassan, and Mário J. Franca. 2018. "The Origin of Fine Sediment Determines the Observations of Suspended Sediment Fluxes Under

Unsteady Flow Conditions." *Water Resources Research* 54 (8): 5654–69. https://doi.org/10.1029/2018WR022982.

Kadkhodaie, Ali, and Reza Rezaee. 2017. "Intelligent Sequence Stratigraphy through a Wavelet-Based Decomposition of Well Log Data." *Journal of Natural Gas Science and Engineering* 40: 38–50. https://doi.org/10.1016/j.jngse.2017.02.010.

Kong, Fanbiao, Shujian Xu, Mei Han, Haitao Chen, Xiaodong Miao, Xianglun Kong, and Guangju Jia. 2021. "Application of Grain Size Endmember Analysis in the Study of Dust Accumulation Processes: A Case Study of Loess in Shandong Province, East China." *Sedimentary Geology* 416: 105868. https://doi.org/10.1016/j.sedgeo.2021.105868.

Kuo, R J, L M Ho, and C M Hu. 2002. "Cluster Analysis in Industrial Market Segmentation through Artificial Neural Network." *Computers & Industrial Engineering* 42 (2): 391–99. https://doi.org/https://doi.org/10.1016/S0360-8352(02)00048-7.

Lau, K. M., and Hengyi Weng. 1995. "Climate Signal Detection Using Wavelet Transform: How to Make a Time Series Sing." *Bulletin - American Meteorological Society* 76 (12): 2391–2402. https://doi.org/10.1175/1520-0477(1995)076<2391:CSDUWT>2.0.CO;2.

Li, Helong, Xiaoyan Deng, and Hongliang Dai. 2007. "Structural Damage Detection Using the Combination Method of EMD and Wavelet Analysis." *Mechanical Systems and Signal Processing* 21 (1): 298–306.

Li, Zongyu, Zhilin Sun, Jing Liu, Haiyang Dong, Wenhua Xiong, Lixia Sun, and Hanyu Zhou. 2022. "Prediction of River Sediment Transport Based on Wavelet Transform and Neural Network Model." *Applied Sciences (Switzerland)* 12 (2). https://doi.org/10.3390/app12020647.

Liang, Xiaolei, Qinghe Niu, Jianjun Qu, Bing Liu, Benli Liu, Xiaohui Zhai, and Baicheng Niu. 2019. "Applying End-Member Modeling to Extricate the Sedimentary Environment of Yardang Strata in the Dunhuang Yardang National Geopark, Northwestern China." *Catena* 180 (September 2018): 238–51. https://doi.org/10.1016/j.catena.2019.04.029.

Liu, Yuming, Xingxing Liu, and Youbin Sun. 2021. "QGrain: An Open-Source and Easy-to-Use Software for the Comprehensive Analysis of Grain Size Distributions." *Sedimentary Geology* 423: 105980. https://doi.org/10.1016/j.sedgeo.2021.105980.

Mahiques, Michel Michaelovitch de, Samara Cazzoli y Goya, Maria Carolina da Silva Nogueira de Matos, Rodrigo Augusto Udenal de Oliveira, Bianca Sung Mi Kim, Paulo Alves de Lima Ferreira, Rubens Cesar Lopes Figueira, and Marcia Caruso Bícego. 2021. "Grain-Size End-Members and Environmentally Sensitive Grain-Size Components: A Comparative Study in the Mud Shelf Depocenters off Southern Brazil." *International Journal of Sediment Research* 36 (2): 317–27. https://doi.org/10.1016/j.ijsrc.2020.07.004.

Marsh, P. 1990. "Permafrost and Lakes in the Mackenzie Delta." *Collection Nordicana* 54 (In Proceedings of the Fifth Canadian Permafrost Conference): 131–36.

Meyer, Inka, Maarten Van Daele, Niels Tanghe, Marc De Batist, and Dirk Verschuren.

2020. "Reconstructing East African Monsoon Variability from Grain-Size Distributions: End-Member Modeling and Source Attribution of Diatom-Rich Sediments from Lake Chala." *Quaternary Science Reviews* 247: 106574. https://doi.org/10.1016/j.quascirev.2020.106574.

Michael Allaby. 2020. *A Dictionary of Geology and Earth Sciences*. 5th ed. Oxford University Press. https://doi.org/10.1093/acref/9780198839033.001.0001.

Moore, Peter D. 1989. "The Ecology of Peat-Forming Processes: A Review" 12: 89–103.

Mullins, C E. 1977. "MAGNETIC SUSCEPTIBILITY OF THE SOIL AND ITS SIGNIFICANCE IN SOIL SCIENCE – A REVIEW." *Journal of Soil Science* 28 (2): 223–46. https://doi.org/https://doi.org/10.1111/j.1365-2389.1977.tb02232.x.

North Greenland Ice Core Project members. 2004. "High-Resolution Record of Northern Hemisphere Climate Extending into the Last Interglacial Period." *Nature* 431 (7005): 147–51.

Pál Sümegi, Mihály Molnár, Gusztáv Jakab, Gergı Persaits, Péter Majkut, Dávid G Páll, Sándor Gulyás, A J Timothy Jull, Tünde Törcsik. 2014. "Radiocarbon-Dated Paleoenvironmental Changes on a Lake and Peat Sediment Sequence from the Central Great Hungarian Plain (Central Europe) During The," no. January 2011. https://doi.org/10.1017/S0033822200034378.

Pannatier, E G. 1997. *Sediment Accumulation and Historical Deposition of Trace Metals and Trace Organic Compounds in the Mackenzie Delta (Northwest Territories, Canada)*.

Passega, R. 1964. "Grain Size Representation by CM Patterns as a Geologic Tool." *Journal of Sedimentary Research* 34 (4): 830–47. https://doi.org/10.1306/74d711a4-2b21-11d7-8648000102c1865d.

Passega, R. 1957. "Texture as Charcteristic of Clastic Deposition" 41 (9): 1952–84.

Paterson, Greig A, and David Heslop. 2015. "Geochemistry, Geophysics, Geosystems," 4494–4506. https://doi.org/10.1002/2015GC006070.Received.

Pawłowski, Dominik, Grzegorz Kowalewski, Krystyna Milecka, Mateusz Płóciennik, Michał Woszczyk, Tomasz Zieliński, Daniel Okupny, Wojciech Włodarski, and Jacek Forysiak. 2015. "A Reconstruction of the Palaeohydrological Conditions of a Flood-Plain: A Multi-Proxy Study from the Grabia River Valley Mire, Central Poland." *Boreas* 44 (3): 543–62. https://doi.org/10.1111/bor.12115.

Pawlowsky-Glahn, Vera, and Antonella Buccianti. 2011. *Preface. Compositional Data Analysis: Theory and Applications*. https://doi.org/10.1002/9781119976462.

Perez-Muñoz, Teresa, Jorge Velasco-Hernandez, and Eliseo Hernandez-Martinez. 2013. "Wavelet Transform Analysis for Lithological Characteristics Identification in Siliciclastic Oil Fields." *Journal of Applied Geophysics* 98 (October 2017): 298–308. https://doi.org/10.1016/j.jappgeo.2013.09.010.

Perşoiu, Aurel, Bogdan P. Onac, Jonathan G. Wynn, Maarten Blaauw, Monica Ionita, and Margareta Hansson. 2017. "Holocene Winter Climate Variability in Central and Eastern Europe." *Scientific Reports* 7 (1): 1–8. https://doi.org/10.1038/s41598-017-01397-w.

Pfitzner, Darius, Richard Leibbrandt, and David Powers. 2009. "Characterization and Evaluation of Similarity Measures for Pairs of Clusterings." *Knowledge and Information Systems* 19 (3): 361–94. https://doi.org/10.1007/s10115-008-0150-6.

Pierce, Aaron R., and Sammy L. King. 2008. "Spatial Dynamics of Overbank Sedimentation in Floodplain Systems." *Geomorphology* 100 (3–4): 256–68. https://doi.org/10.1016/j.geomorph.2007.12.008.

Pinay, G, A Fabre, Ph Vervier, and F Gazelle. 1992. "Control of C, N, P Distribution in Soils of Riparian Forests." *Landscape Ecology* 6 (3): 121–32.

Pizzuto, James E. 1987. "Sediment Diffusion during Overbank Flows." *Sedimentology* 34 (2): 301–17. https://doi.org/https://doi.org/10.1111/j.1365-3091.1987.tb00779.x.

Prokoph, Andreas, and R. Timothy Patterson. 2004. "Application of Wavelet and Regression Analysis in Assessing Temporal and Geographic Climate Variability: Eastern Ontario, Canada as a Case Study." *Atmosphere - Ocean* 42 (3): 201–12. https://doi.org/10.3137/ao.420304.

R. Thompson and F. Oldfield. 1986. *Environmental Magnetism. Syria Studies*. Vol. 7. https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civil wars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625.

Rivera, Nestor A., S. Ray, Jerry L. Jensen, Andrew K. Chan, and Walter B. Ayers. 2004. "Detection of Cyclic Patterns Using Wavelets: An Example Study in the Ormskirk Sandstone, Irish Sea." *Mathematical Geology* 36 (5): 529–43. https://doi.org/10.1023/B:MATG.0000037735.34280.42.

Roberson, Sam, and Gert Jan Weltje. 2014. "Inter-Instrument Comparison of Particle-Size Analysers." *Sedimentology* 61 (4): 1157–74. https://doi.org/10.1111/sed.12093.

Saadatinejad, M R, and H Hassani. 2013. "Application of Wavelet Transform for Evaluation of Hydrocarbon Reservoirs: Example from Iranian Oil Fields in the North of the Persian Gulf." *Nonlinear Processes in Geophysics* 20 (2): 231–38.

Schwarzacher, W. 2000. "Repetitions and Cycles in Stratigraphy," 51–75.

Sedláček, Jan, Veronika Kapustová, Daniel Šimíček, Ondřej Bábek, and Milan Sekanina. 2019. "Initial Stages and Evolution of Recently Abandoned Meanders Revealed by Multi-Proxy Methods in the Odra River (Czech Republic)." *Geomorphology* 333: 16–29. https://doi.org/https://doi.org/10.1016/j.geomorph.2019.02.027.

Shang, Yuan, Christiaan J. Beets, Hui Tang, Maarten A. Prins, Yann Lahaye, Roel van Elsas, Leena Sukselainen, and Anu Kaakinen. 2016. "Variations in the Provenance of the Late Neogene Red Clay Deposits in Northern China." *Earth and Planetary Science Letters* 439: 88–100. https://doi.org/10.1016/j.epsl.2016.01.031.

Shen, Zhixiong, Molly Aeschliman, and Nicholas Conway. 2021. "Paleodischarge Reconstruction Using Oxbow Lake Sediments Complicated by Shifting Hydrological Connectivity." *Quaternary International*.

https://doi.org/10.1016/j.quaint.2021.07.004.

Sinha, Satish, Partha Routh, and Phil Anno. 2009. "Instantaneous Spectral Attributes Using Scales in Continuous-Wavelet Transform." *Geophysics* 74 (2): WA137–42.

SÜMEGI P.-VISSI E. 1991. "A Pocsaji Láp Kialakulása És Fejlődéstörténete." *Calandrella* 5: 15–27.

Sun, Donghuai, J. Bloemendal, D. K. Rea, J. Vandenberghe, Fuchu Jiang, Zhisheng An, and Ruixia Su. 2002. "Grain-Size Distribution Function of Polymodal Sediments in Hydraulic and Aeolian Environments, and Numerical Partitioning of the Sedimentary Components." *Sedimentary Geology* 152 (3–4): 263–77. https://doi.org/10.1016/S0037-0738(02)00082-9.

Syvitski, James P M, Charles J Vo, Albert J Kettner, and Pamela Green. 2005. "Impact Ofhumans on the Flux of Terrestrial Sediment to the Global Coastalocean." *Science* 308 (April): 376–80.

Szilágyi, Szilvia Sebok, and János Geiger. 2012. "Sedimentological Study of the Szőreg-1 Reservoir (Algyő Field, Hungary): A Combination of Traditional and 3D Sedimentological Approaches." *Geologia Croatica* 65 (1): 77–90. https://doi.org/10.4154/GC.2012.06.

Szöőr, Gy, P Sümegi, and É Balázs. 1991. "Sedimentological and Geochemical Facies Analysis Upper Pleistocene Fossil Soil Zones Discovered in the Hajdúság Region, NE Hungary." *Quaternary Environment in Hungary. Studies in Geography in Hungary* 26: 47–59.

Tarar, Zeeshan Riaz, Sajid Rashid Ahmad, Iftikhar Ahmad, and Zahra Majid. 2018. "Detection of Sediment Trends Using Wavelet Transforms in the Upper Indus River." *Water (Switzerland)* 10 (7): 1–17. https://doi.org/10.3390/w10070918.

Tite, M.~S. 1975. "Effect of Climate on the Magnetic Susceptibility of Soils." \nat 256 (5518): 565–66. https://doi.org/10.1038/256565a0.

Toonen, W H J, T G Winkels, K M Cohen, M A Prins, and H Middelkoop. 2015. "Lower Rhine Historical Flood Magnitudes of the Last 450 Years Reproduced from Grain-Size Measurements of Flood Deposits Using End Member Modelling." *Catena* 130: 69–81.

Toonen, Willem H.J., Maarten G. Kleinhans, and Kim M. Cohen. 2012. "Sedimentary Architecture of Abandoned Channel Fills." *Earth Surface Processes and Landforms* 37 (4): 459–72. https://doi.org/10.1002/esp.3189.

Tooth, S. 2013. "Dryland Fluvial Environments: Assessing Distinctiveness and Diversity from a Global Perspective." In, edited by John F B T - Treatise on Geomorphology Shroder, 612–44. San Diego: Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-374739-6.00257-8.

Vandenberghe, Jef, Youbin Sun, Xianyan Wang, H A Abels, and X Liu. 2018. "Grain-Size Characterization of Reworked Fine-Grained Aeolian Deposits." *Earth-Science Reviews* 177: 43–52.

Vincze, Ildikó, Walter Finsinger, Gusztáv Jakab, Mihály Braun, Ildikó Vincze, Walter Finsinger, Gusztáv Jakab, Mihály Braun, and Katalin Hubay. 2020. "Paleoclimate Reconstruction and Mire Development in the Eastern Great Hungarian Plain for

the Last 20 , 000 Years To Cite This Version : HAL Id : Hal-02279803." *Review of Palaeobotany and Palynology*, 104112. https://doi.org/10.1016/j.revpalbo.2019.104112.

W.M. Ma, Eden, and Tommy W S Chow. 2004. "A New Shifting Grid Clustering Algorithm." *Pattern Recognition* 37 (3): 503–14. https://doi.org/https://doi.org/10.1016/j.patcog.2003.08.014.

Wang, Junyi, Zhongtai He, and Linlin Li. 2021. "Palaeoseismic Records in Lacustrine Sediments—A Case Study of the Daqingshan Piedmont Fault and Hasuhai Lake in Inner Mongolia, China." *Basin Research* 33 (1): 681–704. https://doi.org/10.1111/bre.12490.

Wang, Shuo, Ninglian Wang, Yuzhu Zhang, Chang Huang, Yan Zhu, Qili Xiao, Dou Chen, et al. 2023. "Geochemistry of a Paleo-Oxbow Lake Sediments and Its Implications for the Late Holocene Extreme Overbank Fl Ooding History of the Yellow River within the Zoige Basin , NE Tibetan Plateau," no. February: 1–14. https://doi.org/10.3389/feart.2023.1144283.

Ward, J.H., 1963. 1963. "Hierarchical Grouping to Optimize an Objective Function" 58 (301): 236–44.

Weltje, Gert Jan. 1997. "End-Member Modeling of Compositional Data: Numerical-Statistical Algorithms for Solving the Explicit Mixing Problem." *Mathematical Geology* 29 (4): 503–49. https://doi.org/10.1007/bf02775085.

Woolway, R. Iestyn, and Christopher J. Merchant. 2019. "Worldwide Alteration of Lake Mixing Regimes in Response to Climate Change." *Nature Geoscience* 12 (4): 271–76. https://doi.org/10.1038/s41561-019-0322-x.

Wren, Daniel G., Jason M. Taylor, J. R. Rigby, Martin A. Locke, and Lindsey M.W. Yasarer. 2019. "Short Term Sediment Accumulation Rates Reveal Seasonal Time Lags between Sediment Delivery and Deposition in an Oxbow Lake." *Agriculture, Ecosystems and Environment* 281 (May): 92–99. https://doi.org/10.1016/j.agee.2019.05.007.

Wu, Li, Wout Krijgsman, Jian Liu, Chaozhu Li, Rujian Wang, and Wenshen Xiao. 2020. "CFLab: A MATLAB GUI Program for Decomposing Sediment Grain Size Distribution Using Weibull Functions." *Sedimentary Geology* 398. https://doi.org/10.1016/j.sedgeo.2020.105590.

Yu, Shi Yong, Steven M. Colman, and Linxiong Li. 2016. "BEMMA: A Hierarchical Bayesian End-Member Modeling Analysis of Sediment Grain-Size Distributions." *Mathematical Geosciences* 48 (6): 723–41. https://doi.org/10.1007/s11004-015-9611-0.

Zhang, Guoqing, Tandong Yao, Hongjie Xie, Kun Yang, Liping Zhu, C K Shum, Tobias Bolch, et al. 2020. "Response of Tibetan Plateau Lakes to Climate Change: Trends, Patterns, and Mechanisms." *Earth-Science Reviews* 208: 103269. https://doi.org/https://doi.org/10.1016/j.earscirev.2020.103269.

Zhang, Xiaodong, Hongmin Wang, Shumei Xu, and Zuosheng Yang. 2020. "A Basic End-Member Model Algorithm for Grain-Size Data of Marine Sediments." *Estuarine, Coastal, and Shelf Science* 236 (February): 106656. https://doi.org/10.1016/j.ecss.2020.106656.

Zhang, Xiaonan, Hucai Zhang, Fengqin Chang, Umar Ashraf, Han Wu, Wei Peng, Qi
    Liu, Fengwen Liu, Yun Zhang, and Lizeng Duan. 2021. "Sedimentary Grain-Size
    Record of Holocene Runoff Fluctuations in the Lake Lugu Watershed, SE Tibetan
    Plateau." *Holocene* 31 (3): 346–55. https://doi.org/10.1177/09596836209727777.