

# Modelling the joint variability of grain size and chemical composition in sediments

M.R. Bloemsma <sup>a,\*</sup>, M. Zabel <sup>c</sup>, J.B.W. Stuut <sup>b,c</sup>, R. Tjallingii <sup>b</sup>, J.A. Collins <sup>c</sup>, G.J. Weltje <sup>a</sup>

<sup>a</sup> Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Geotechnique, Stevinweg 1, NL-2628CN Delft, The Netherlands

<sup>b</sup> Royal Netherlands Institute of Sea Research (NIOZ), Department of Marine Geology, Landsdiep 4, NL-1797SZ 't Horntje (Texel), The Netherlands

<sup>c</sup> MARUM – Center for Marine Environmental Sciences, University of Bremen, Leobener Strasse, D-28359 Bremen, Germany

## ARTICLE INFO

### Article history:

Received 18 July 2011

Received in revised form 14 March 2012

Accepted 16 April 2012

Available online 26 April 2012

### Keywords:

Partial least squares

Multi-proxy analysis

Compositional data analysis

Geochemical proxies

Singular value decomposition

Provenance

## ABSTRACT

The geochemical composition of siliciclastic sediments correlates strongly with grain size. Hence, geochemical composition may serve as a grain-size proxy. In the absence of grain-size variations, geochemical data of siliciclastic sediments may be used to characterise size-independent processes, i.e., sediment provenance, weathering, mixing, shape/density sorting and diagenesis. In general, however, geochemical data sets contain both types of information. In order to formalise interpretation of geochemical data, we propose a mathematical method to decompose the total geochemical variability of a series of genetically related specimens into a grain-size dependent (the shared signal) and a grain-size independent part (the residual signal). The former may serve as a proxy for grain size whereas the latter represents geochemical variability that would have been observed if all sediments would have had the same grain-size distribution. The two data sets are jointly decomposed by means of Partial Least Squares (PLS) and orthogonal projection. Subsequently, the presence of significant grain-size independent geochemical variability in the residual signal is determined in a statistically rigorous manner using a  $\chi^2$ -test. Using a synthetic example, we show that the residual record effectively reveals an imposed provenance signal which could not have been resolved from the geochemical or grain-size data sets individually.

We analysed the relation between grain size and geochemical composition in three Quaternary marine sediment cores located offshore West Africa and South America (GeoB7920-2, GeoB9508-5 and GeoB7139-2). Both sites are characterised by biogenic sediment input, in addition to fluvial and aeolian sediment input from the continent. It was found that all cores show a strong, but different correlation between the mean grain size and the bulk geochemical composition. These results demonstrate that geochemical grain-size proxies are empirical and site-specific. It was also found that the geochemical and grain-size data in cores GeoB7920-2 and GeoB7139-2 do not contain unique information, whereas in core GeoB9508-5 Ti varies independently from the grain size. This residual Ti-signal correlates with the transport mechanism, as demonstrated by statistically different values of aeolian and fluvial-dominated sediments. However, a unique interpretation of this residual signal in terms of the postulated grain-size independent mechanisms could not be provided without additional information.

We conclude that the proposed model facilitates identification and validation of different element ratios as grain-size proxies and, more importantly, as proxies for size-independent processes. For this reason, the model paves the way for rigorous analysis of multi-proxy data, which are widely used in palaeoceanographic and palaeoclimatic research.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Geochemical analysis is a powerful method of quantifying bulk sediment properties. It may be used to characterise the composition of the parent rock, or the climatic-physiographic conditions under which the sediment was formed (Pettijohn et al., 1987; Johnsson, 1993; Basu, 2003; Weltje and von Eynatten, 2004). Alternatively, it may be used to assess compositional modifications caused by

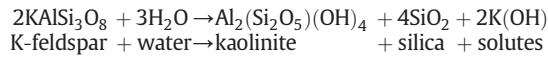
weathering, sorting, mixing and diagenesis (McLennan et al., 1993; Nesbitt and Young, 1996; Amorosi et al., 2002; Dinelli et al., 2007; Pe-Piper et al., 2008). Because geochemical variation typically reflects the superposition of these factors, independent information is often required to resolve them individually. This applies specifically to analysis of clastic sediments, where we can postulate multiple equally plausible hypotheses explaining the observed compositional data.

It is well known that grain size and geochemical composition of clastic sediments are highly correlated. The tight connection between grain size and bulk chemistry is a consequence of the processes which govern the generation of sediments from crystalline rocks. Chemical weathering leads to release of unstable elements as solutes, while

\* Corresponding author.

E-mail address: [m.r.bloemsma@tudelft.nl](mailto:m.r.bloemsma@tudelft.nl) (M.R. Bloemsma).

stable elements such as Al remain in the solid phase (e.g., Nesbitt and Young, 1984). Consider for instance chemical weathering of K-feldspar:



Since kaolinite particles are relatively fine grained, whereas the K-feldspar grains are of sand size, chemical alteration is accompanied by textural modification of the sediment. If the K-feldspar and kaolinite are transported away from their source area and deposited in a sedimentary basin, the resulting product is a sediment with a spatially variable grain-size distribution. If we sample this material at different localities, we expect Si/Al of the bulk sediment to correlate positively with mean grain size. Quite often, the objective of geochemical analysis is not to approximate grain size, but to provide a unique compositional fingerprint of the sediment for the purpose of paleoclimate or provenance reconstruction. In such cases, we would like to eliminate the size-dependency from our data.

A method to eliminate variation of bulk grain-size distribution from sediment geochemical analysis is to analyse a narrow size fraction (Weltje and Brommer, 2011). In the above example, the composition of a narrow size fraction would be constant across the basin. However, reality is generally more complicated. If we assume that the extent of weathering in the source area of the sediment has varied over time, we expect that the composition of a narrow size fraction extracted from samples taken at different stratigraphic levels records this variation. For example, the Si/Al ratio within the sand fraction will correlate positively with the extent of weathering, whereas the Si/Al ratio of the clay fraction will correlate negatively with the extent of weathering. Because separation of samples into narrow size fractions is time consuming and expensive, we should try to find more practical ways of eliminating the grain-size bias from geochemical data.

In marine geosciences it is commonplace to use a so-called multi-proxy approach to infer paleo-climatological signals. In this approach, high-resolution records of physical and chemical properties, for example grain-size distribution and bulk chemical composition, are measured on the same sediment core. Subsequently, these data are displayed side by side to highlight consistency between two data sets, or to illustrate patterns of variability unique to each data set. A widely accepted method to select the relevant signals (i.e., the relevant element ratios and/or grain-size parameters) from multivariate data sets is not yet in place, and the same applies to generic procedures to examine the internal consistency of selected proxies. In this contribution, we propose a fully quantitative and formal method for simultaneous analysis of multiple data sets. Our analysis of the relation between the geochemical composition and grain-size distribution of sediment will serve to illustrate the implications of the proposed statistical framework for multi-proxy analysis.

## 2. Material and methods

### 2.1. Materials

We use three marine soft-sediment cores from the margin off West African and Chile to investigate geochemical variation with grain size in lithogenic sediments. These cores are well suited for testing this variation since both West Africa and Chile receive a considerable amount of sediment from the continent, transported by wind and rivers (e.g., Sarnthein et al., 1981; Stuut et al., 2007). In addition, both areas are characterised by strong year-round upwelling, resulting in a considerable biogenic sediment input (Haslett and Smart, 2006; De Pol-Holz et al., 2007).

Cores GeoB9508-5 and GeoB7920-2 were retrieved from the continental margin off West Africa, and core GeoB7139-2 was retrieved from the margin off Chile (Fig. 1). Core GeoB9508-5 (15°30'N/17°57'W, 2384 m water depth) is approximately 9.5 m long and

was obtained during RV Meteor cruise M65/1 (Mulitza and cruise participants, 2006). Core GeoB7920-2 (20°45'N/18°35'W 2278 m water depth) is approximately 16 m long and was retrieved during RV Meteor cruise M53/1 (Meggers and cruise participants, 2002). The Chilean core GeoB7139-2 (30°12'S/71°59'W) was retrieved during the PUCK expedition on RV Sonne in 2001 (Hebbeln and cruise participants, 2001) and is approximately 8 m long. In all three cores, the sediments were dated using stable oxygen isotopes and radiocarbon dates, which show that they date back to 57 kyr BP (GeoB9508-5; Mulitza et al., 2008), 118 kyr BP (GeoB7920-2; Tjallingii et al., 2008) and 65 kyr BP (GeoB7139-2; De Pol-Holz et al., 2007).

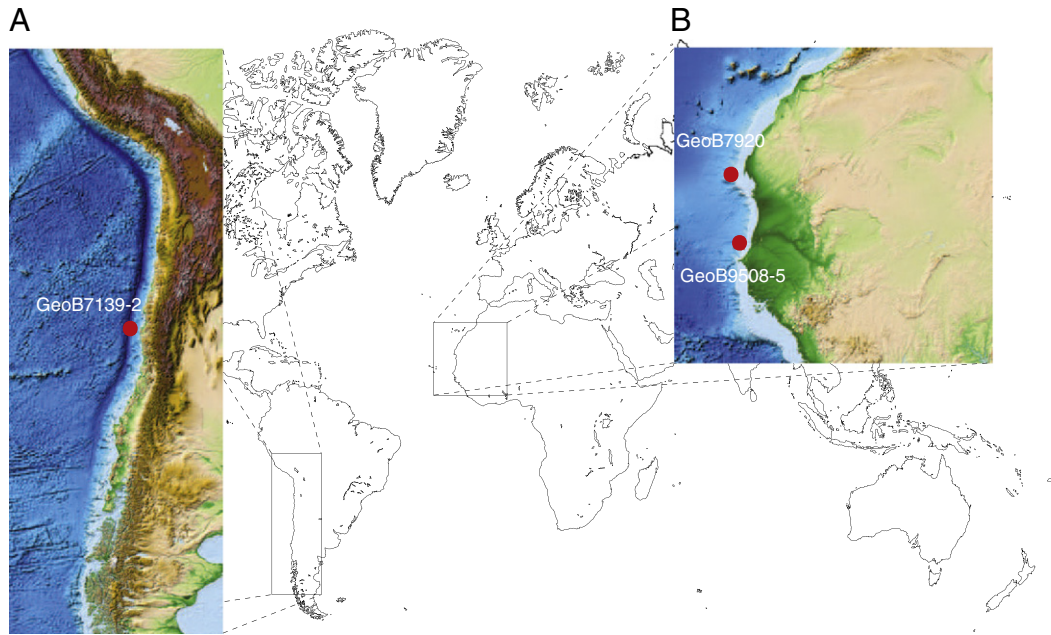
Marine sediment cores have served as a valuable source of information to infer Late Quaternary variations in paleoclimatic conditions in both South America (e.g., Lamy et al., 1998) and West Africa (e.g., Tjallingii et al., 2008). Climatic conditions have been inferred from grain-size records (e.g., Holz et al., 2004; Stuut and Lamy, 2004; Tjallingii et al., 2008) and from geochemical records (e.g., Haslett and Davies, 2006; Mulitza et al., 2008), in addition to various combinations of sediment properties (e.g., Kaiser et al., 2008; Romero et al., 2008).

In West Africa, present-day samples from fluvial and aeolian sources show that aeolian dust is relatively coarse grained, compared to fluvial sediment (Gac and Kane, 1986; Stuut et al., 2005). Grain-size variations in the lithogenic fraction of marine cores are therefore considered to reflect the transporting mechanism (Sarnthein, 1978; Koopmann, 1981; Holz et al., 2004; Stuut et al., 2007). Proximal to the source, sediment with a grain size below 6 µm is typically assumed to be of fluvial origin, whereas sediment coarser than 6 µm is assumed to be transported by wind (Koopmann, 1981; Lamy et al., 1998; Holz et al., 2004). Nevertheless, if transport distances are relatively long, the typically large grain size of the aeolian dust may be reduced by proximal to distal fining (e.g., Weltje and Prins, 2003; Stuut et al., 2005). The same information (i.e., transporting medium, its energy and transport distance) may be contained in geochemical signals; Boyle (1983) postulated Al/Ti as a generic grain-size proxy because it is controlled by the concentration of heavy minerals transported along with the coarse fraction. For this reason, downcore variations in Al/Ti were considered a proxy for aridity and/or wind strength (Yarincik et al., 2000).

Besides the shared information, geochemical records are considered to contain additional information regarding source area characteristics. Sediment that is derived from chemically weathered terrain typically has high Al and Fe concentrations (Moreno et al., 2006; Mulitza et al., 2008), which may therefore serve as a proxy for humidity (Sarnthein, 1978; deMenocal et al., 1993). Schneider et al. (1997) interpreted elevated Al/K towards the tropics as being induced by enrichment in kaolinite relative to K-feldspar as a result of increased chemical weathering intensity. Yarincik et al. (2000) used the same ratio as a proxy for the amount of illite relative to kaolinite which is also considered to be controlled by the intensity of chemical weathering. Compared to West Africa, the effect of chemical and physical weathering on the geochemical composition of sediments from the continent is low, due to short transport distances and arid climate conditions (Lamy et al., 2000). Because this implies that the sediments retain their initial composition, geochemical records from this area may serve as a tracer for parent rock lithology (Lamy et al., 1998; Klump et al., 2000; Lamy et al., 2000).

### 2.2. Analytical methods

Samples for grain-size analysis were acquired every 5 cm for cores GeoB9508-5 (n = 191) and GeoB7920-2 (n = 330), and every 2 cm in core GeoB7139-2 (n = 389). The siliciclastic sediment fraction was isolated by dissolving carbonate, organic matter and biogenic opal in HCl, H<sub>2</sub>O<sub>2</sub> and NaOH, respectively. The samples were heated with about 300 mg of Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> · 10H<sub>2</sub>O directly before measuring to avoid the formation of aggregates in the fine-grained fraction. The grain-size distribution was determined with a Coulter LS200 laser particle sizer,



**Fig. 1.** Location of the cores used in this study. Core GeoB7139-2 was retrieved off the Chilean coast (inset A), whereas cores GeoB9508-5 and GeoB7920-2 were retrieved off the West-African coast (inset B). The insets are derived from the ETOPO1 topographical map (Amante and Eakins, 2008).

which detects grains in the range from 0.4 to 2000  $\mu\text{m}$  (i.e., between 11 and -1 phi-units) (Fig. 2).

All cores have been scanned with an Avaatech core scanner (e.g., Richter et al., 2006) at a 2 cm resolution, a source current of 10 kV and a measurement time of 30 s. The measured XRF spectra were converted to a record of net element intensities using the WinAxil software package (Vekemans et al., 1994). In order to be able to convert the element intensities to concentrations, quantitative geochemical analysis of powdered sediment samples was carried out with a portable Spectro Xepos Analyzer (Wien et al., 2005). For core GeoB9508-05, 229 samples were geochemically analysed, whereas the geochemical reference data sets of GeoB7920-2 and GeoB7139-2 comprise 168 and 20 samples, respectively.

### 2.3. Data pre-processing

Quantitative calibration of the XRF core scanning records was performed using the log-ratio calibration equation (LRCE) (Weltje and Tjallingii, 2008). The predictive power of each alternative calibration model (i.e., linear models with a certain denominator element) is empirically quantified by the taking the median of the squared discrepancies between the predicted and the reference geochemical composition. We calculate these discrepancies using leave-one-out cross-validation (LOOCV) (Geisser, 1993). As a result, they account for all factors influencing the deviation between XRF core scanning records and destructive analysis (e.g., difference in analysed sample volume). A detailed description of these factors is provided by Tjallingii et al. (2007).

Based on the Aitchison distance between predicted and reference composition, we select the optimal denominator element for each core, which are Si (GeoB7139-2), Al (GeoB7920-2) and Fe (GeoB9508-5). For these calibration models, the LOOCV median squared discrepancies of the different log-ratios are shown in the upper diagonal of Table 1. The lower diagonal shows half of the width of 95% confidence limits, which are directly derived from the variances in the upper diagonal (see Weltje and Tjallingii, 2008).

To analyse the relation between grain size and geochemical composition, we use samples of which both grain size and geochemical composition were measured. Because the resolution of the grain size

analysis is different from that of the core scan, a subset of the grain-size and bulk geochemical data meets these requirements, namely 185 (GeoB7139-2), 163 (GeoB7920-2) and 92 (GeoB9508-5) samples.

In order to make application of log-ratio analysis to the grain size data possible, we removed the zero-valued channels. In this study, all channels of the grain-size distribution which contain a zero in any of the observations are amalgamated with an adjacent channel that contains only positive values. Generally this leaves more than 80% of the channels unchanged. Given the high level of redundancy in grain-size data (Weltje and Prins, 2003), we consider this step as having a negligible effect on the information content of the data.

## 3. Modelling approach

### 3.1. Conceptual model

From a theoretical perspective, the information provided by textural and geochemical data may be subdivided into (1) information shared by the two records, (2) information unique to the geochemical record, and (3) information unique to the grain-size record. The partitioning of the geochemical and grain-size variability into these three independent components is illustrated in Fig. 3A. The significance of each of these three parts will be discussed below.

If a series of sediment samples has been derived from a single source area in which the extent of chemical weathering did not vary significantly over time, most variation in bulk chemical composition may be attributed to fractionation during entrainment, transport, and deposition. Hence, if it were possible to apply a “grain-size correction” to the bulk chemical composition of this hypothetical data set, the residual geochemical composition would not show any significant downcore variation. Such a one-to-one correspondence between grain-size and geochemistry implies that the geochemical record carries no unique information and the bulk chemical composition can be accurately predicted from the grain-size distribution and vice versa.

In reality this one-to-one correspondence between composition and grain size will be an exception, and a single grain-size distribution may correspond to different geochemical signatures. Various

mechanisms for generating size-independent geochemical variability may be envisaged (see Fig. 3):

1. Chemical weathering: if the degree of chemical weathering varies over time at a single locality, the residual Si/Al ratio will correlate positively with the extent of weathering. Similar trends will be present in other element ratios (Calvert and Pederson, 2007; Xiong et al., 2010).
2. Hydraulic/aerodynamic sorting: sediment-forming minerals in the silt to sand fraction span a wide range of densities (from 2.6 to about 5.0 g/cm<sup>3</sup>) and shapes (spherical to platy). Mineral grains which are susceptible to size-independent fractionation (i.e., depletion and enrichment not mirrored in the bulk grain-size distribution) must have a combination of density and shape which deviates strongly from the bulk, and be present in small proportions. Chemical elements which exclusively reside in such minerals (e.g., Ti and Zr) are therefore excellent tracers of fractionation.
3. Mixing: if different source areas shed sediments with distinct geochemical signatures within a given grain-size range, the residual geochemical signal represents the variation of mixing coefficients of these compositionally distinct sediment types.

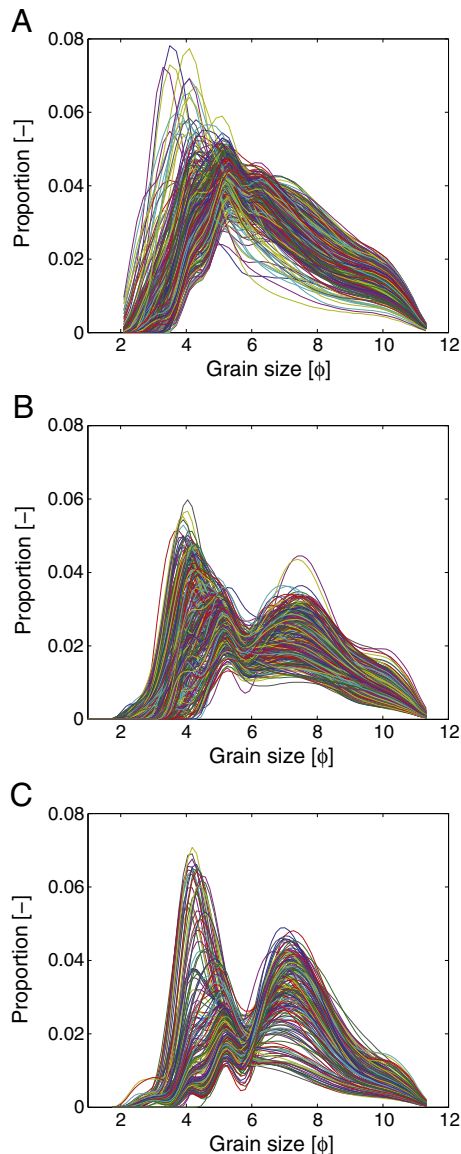


Fig. 2. Grain size distributions of samples from cores GeoB7139-2 (A), GeoB7920-2 (B) and GeoB9508-5 (C).

4. Diagenesis: if the degree of diagenetic modification varies across a series of samples with similar grain-size distributions, a size-independent signal will be generated. Examples include dissolution and precipitation at different stratigraphic levels within a core.

For the sake of completeness, one should also consider occurrence of grain-size variations which do not correspond to geochemical variability (Fig. 3), which might be the case if rocks have been mechanically ground by glaciers, or mechanically weathered by wind-action. However, even under glacial or extremely arid conditions, geochemical variability was found to be accompanied by grain-size variations (Nesbitt and Young, 1996; Solano-Acosta and Dutta, 2005). We therefore consider this a theoretical possibility rather than something which merits further investigation.

### 3.2. Statistical model

If we can remove the shared variability from grain size and geochemical data (the gray part in Fig. 3A), the residual record may reveal patterns which cannot be distinguished in the original record. We propose a generic approach to achieve this, based on the assumption that a monotonic relation exists between grain size and geochemical composition. Grain-size distributions and geochemical compositions are both compositional in nature (values are non-negative by definition, and the sum over all grain size classes, or all chemical constituents, should equal unity), which requires that we model them as log-ratios (Aitchison, 1986). More specifically, we use a centred log-ratio (clr) transformation so that classical multivariate methods can be used to explore their relation.

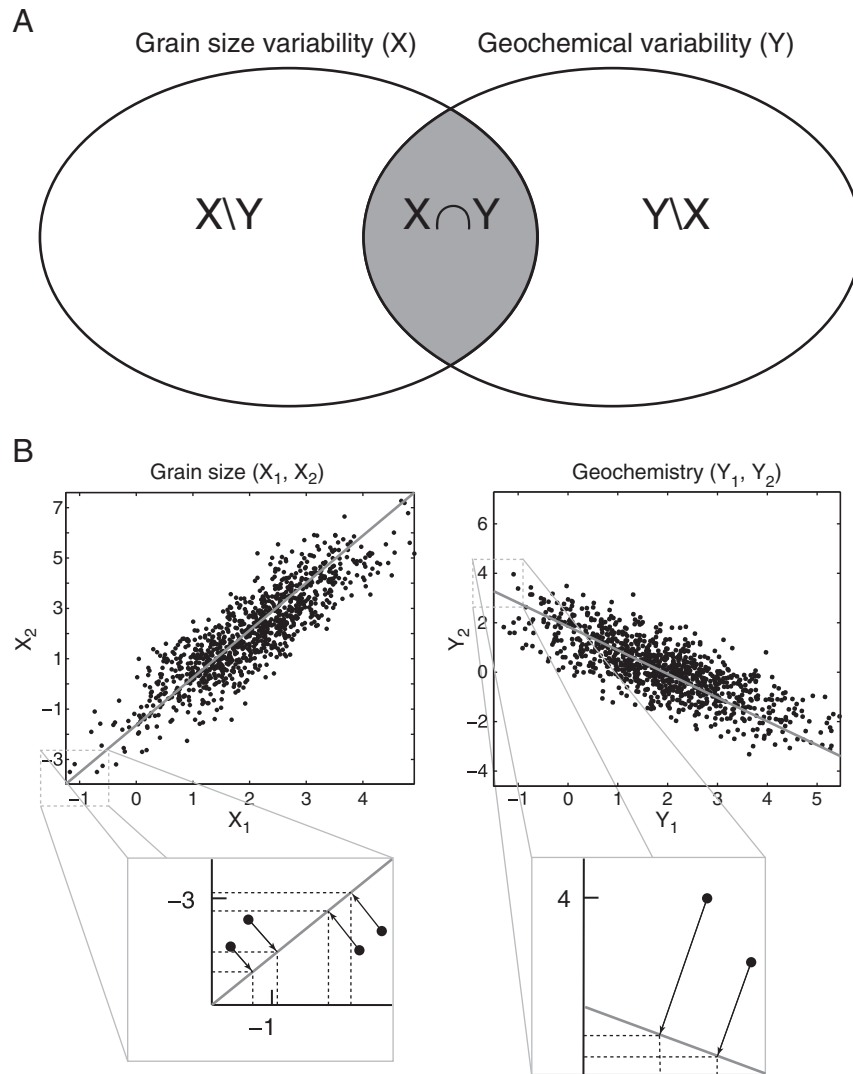
Linear models of log-ratio transformed variables (i.e., log-ratio linear models) are compatible with trends of compositional variability of sediments induced by variation in weathering conditions and grain size (e.g., von Eynatten et al., 2003; Tolosana-Delgado and von Eynatten, 2008). Exceptions to this compatibility were raised by Tolosana-Delgado and von Eynatten (2010) who found nonlinearity in the 'intrinsic' relation between chemistry and composition, derived from chemical analyses of narrow size fractions. However, since we focus only on the relation between bulk grain size and bulk chemistry, which is typically more smooth, we consider the log-ratio linear suitable to model their relation.

A graphical illustration of the model we introduce in this contribution is shown in Fig. 3B. The model establishes two bases (coordinate systems) which maximise the joint variability of grain size and geochemical composition. We derive these bases using Partial Least Squares (PLS) (e.g., see Martens and Naes, 1989). By projecting both datasets onto their basis vectors, we obtain the scores on each vector.

Table 1

The upper diagonal shows the median variances between the observed geochemistry and the geochemistry predicted on the basis of the XRF core scan. The lower diagonal shows the 95% confidence limits corresponding to these residual variances.

	Al	Fe	K	Si	Ti
GeoB7139-2					
Al	–	0.0020	0.0003	0.0014	0.0019
Fe	0.0784	–	0.0032	0.0008	0.0035
K	0.0322	0.0995	–	0.0016	1.7470
Si	0.0660	0.0504	0.0695	–	0.0012
Ti	0.0759	0.1039	2.3280	0.0621	–
GeoB7920-2					
Al	–	0.0024	0.0039	0.0072	0.0039
Fe	0.0792	–	0.0068	0.0157	0.0056
K	0.1030	0.1319	–	0.0049	0.0041
Si	0.1383	0.2034	0.1139	–	0.0154
Ti	0.1025	0.1211	0.1049	0.1990	–
GeoB9508-5					
Al	–	0.0046	0.0032	0.0015	0.0021
Fe	0.1119	–	0.0033	0.0056	0.0033
K	0.0930	0.0954	–	0.0019	0.0032
Si	0.0642	0.1239	0.0716	–	0.0055
Ti	0.0750	0.0949	0.0931	0.1223	–



**Fig. 3.** In A, we illustrate geochemical and grain-size variabilities in the form of a Venn-diagram. The variability shared by both data sets is indicated by  $X \cap Y$ , which is highlighted in gray. In contrast to this shared component, the variability unique to the geochemical data is the portion that potentially holds relevant signals (e.g., provenance). The unique geochemical variability is indicated by  $Y \setminus X$ , representing the residual geochemical variability. As shown in B, the shared signals are derived by an orthogonal projection of each data set on a set of basis vectors. The direction of these basis vectors is such that they jointly maximise the common covariance in the two data sets. The residual signal may then be obtained by subtracting the shared signals from their corresponding raw data set.

If the scores of the grain-size and geochemical data on the  $k$ -th basis vector correlate significantly, we project both data sets on their corresponding bases to obtain their *shared signals*. Subsequently, the variability unique to each data set (i.e., the residual signal) is derived by subtracting the shared signals from their corresponding raw data set.

The number of orthogonal components that are removed is a hyperparameter, and may take on any value between zero (the raw geochemical, and grain-size data) and  $N - 1$ , with  $N$  being the minimum number of variables (either grain-size classes or chemical elements). For every order  $k$ , a  $\chi^2$ -test is used to decide whether the residual variance of a log-ratio of elements  $X$  and  $Y$  (i.e.,  $\log(X/Y)$ ) deviates significantly from the error variance of the corresponding geochemical signal. More detailed information about the mathematical model may be found in the [Appendix](#).

#### 4. Synthetic examples

To illustrate the working principle of the model, we will discuss two synthetic examples. These examples reflect limiting cases; one dataset (dataset 1) reflects the case where compositional variability is only induced by grain size variations, whereas the other dataset

(dataset 2) reflects the case where compositional variability is induced by both grain-size variations and provenance.

##### 4.1. Synthetic example 1

The first step in acquiring a synthetic dataset is stochastic simulation of grain-size distributions. These distributions are constructed to be lognormal, with a random mean and a standard deviation that is proportional to the mean. We impose a standard deviation of fine-grained sediments which is larger than that of coarse-grained sediments. The fact that these two parameters are perfectly correlated implies that the rank of the grain-size data is one. In order to simulate bulk chemical composition, we postulate a log-ratio linear relation between composition and grain size (see Fig. 4A). This relation must be transformed back into proportions using the inverse clr-transformation in order to obtain the function which depicts element concentrations as a function of grain size (see Fig. 4B). The bulk chemical composition is then obtained by multiplying this function with the grain size distribution. This provides a unique bulk geochemical composition for any given grain size distribution. As a final step, these bulk chemical compositions are clr-transformed and noise is added.

Fig. 5A shows the simulated grain-size record, whereas Fig. 5B shows the final bulk geochemical record, expressed in proportions. The residual grain-size and geochemical signals are shown in Fig. 5 (C and D). Since the rank of the grain-size data is one, and the geochemistry is directly inferred from the grain size, the shared signal carried by the first vectors in both solution spaces explains all systematic variation. As a consequence, Fig. 5D reflects only the superimposed noise, centred around the mean bulk chemical composition. We derive the dominant signal in the residual and the shared geochemical signal using Principal Component Analysis (PCA). The Principal Components (PCs) are obtained by means of a Singular Value Decomposition (SVD) (Press et al., 1994). Figs. 5E and 5F show the scores on PC1 of the shared and the residual signal, respectively. Gray bars represent 95% confidence limits derived from the imposed noise. Note that in the PC1 scores of the residual geochemical signal, the proportion of observations plotting outside of these limits is around 5%, which leads to the conclusion that there is no significant geochemical variability. The scores of the shared geochemical signal, on the other hand, show significantly more exceedences, indicating a strong correlation between the shared geochemical signal and the grain size.

Although the individual records of grain size and bulk chemistry (Fig. 5, A and B) display variation which may be difficult to interpret, the model returns that they contain the same information, i.e., the common size-composition trend. Using the model, this common trend was successfully eliminated from the data without any prior knowledge.

#### 4.2. Synthetic example 2

In the second example, we increased the complexity of the simulated data by imposing random mixing of two distinct sediment sources. The size-composition function of the additional source is shown in Fig. 4 (C and D). The grain size distributions were simulated according to the procedure outlined in example 1, and therefore contain no information about provenance (Fig. 6B). A real-world analog of this example is a river with two tributaries, draining different parent-rock types. We assume that the proportional contribution of the tributaries have varied over time and we repeatedly sampled the river-mouth sediments. The grain-size distributions of these sediments reflect the hydrodynamic conditions under which the sediments have been transported, and thus contain no information regarding provenance.

As shown in Figure 6B, dataset 2 exhibits a geochemical record which does not appear to be very different from that of dataset 1. This is not surprising because the size-composition functions of the two sources are very similar (see Fig. 4). As a result, differences induced by variations in the source are overshadowed by variability induced by grain size variations. In contrast to example 1, the residual geochemical record suggests significant variability independent from the grain size (Fig. 6D). This is confirmed by the scores on PC1 of the residual geochemical record, which exceed the 95% confidence interval of the superimposed noise. Without knowing the underlying structure, we would therefore conclude that there is evidence for grain-size independent geochemical variability.

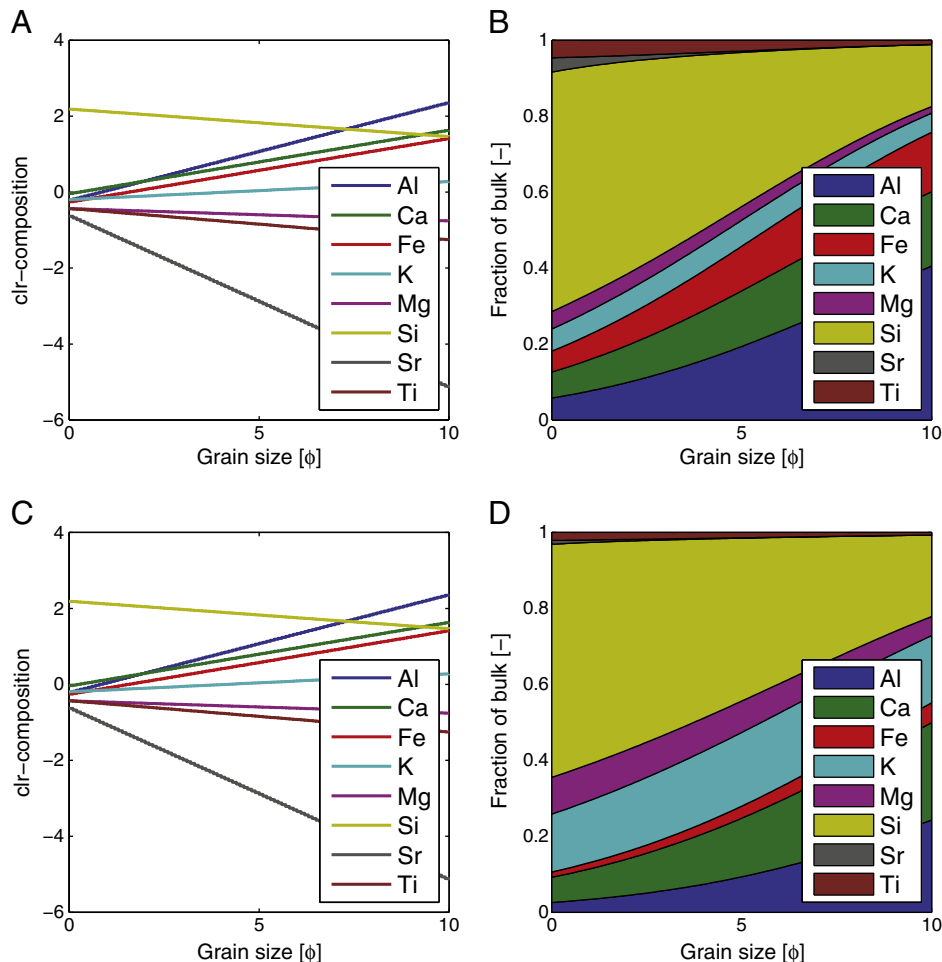
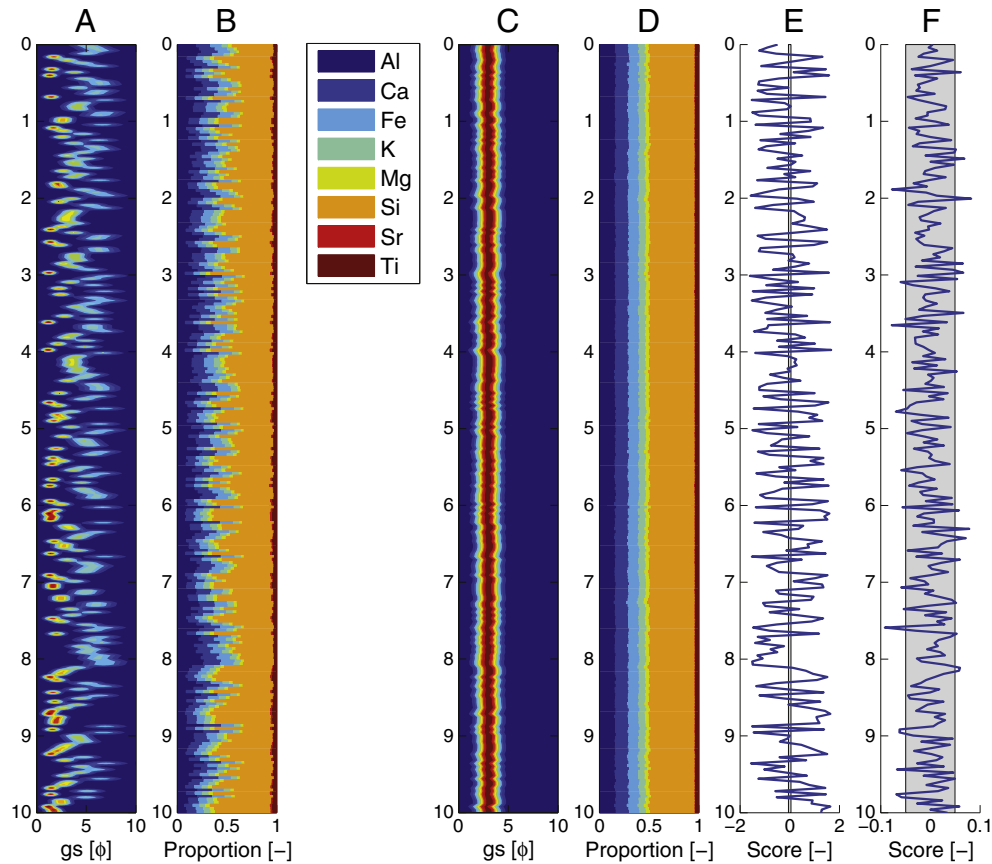


Fig. 4. In the synthetic example, the fingerprints of source 1 and source 2 are parameterised by a log-ratio linear function between grain size and composition (A and C). These log-ratio linear functions can be transformed back to proportions by means of the inverse cfr-transformation (B and D).



**Fig. 5.** This figure shows synthetic dataset 1 ( $n = 250$ ), reflecting one provenance. Plots A and B show the grain size distributions and geochemical record, respectively. The residual grain size and geochemical signals are shown in C and D, respectively. The first principal component scores of the shared geochemical signal is shown in E, whereas plot F shows the first principal component scores of the residual geochemical signal. In Fig. 6E and 6F, a gray bar represents a 95% confidence interval, derived from the imposed noise. Note that approximately 5% of the residual PC1 scores (F) are located outside the 95% confidence interval.

If our aim is to infer either grain size or provenance from the raw geochemical data, a straightforward approach is to analyse the scores on PC1 (i.e., the main direction of variance). Fig. 7A shows the relation between these scores and the mean grain size, whereas Fig. 7B shows the relation between these scores and the provenance (i.e., the contribution from source 1). In both cases there is a weak correlation.

Instead of using the raw data, we may also use the model outcome to derive proxies for both grain size and provenance; PC1 scores of the shared geochemical record represent the grain-size proxy, whereas the residual geochemical record may serve as a provenance proxy. The relation between the grain-size proxy, and the grain size and (known) source contributions, respectively, are shown in Fig. 7 (C and D). Fig. 7C demonstrates that the grain-size proxy correlates strongly with the mean grain size and does not contain any provenance information. The provenance proxy (Fig. 7, E and F), on the other hand, varies independently from the mean grain size, but records the known changes in provenance perfectly.

## 5. Results

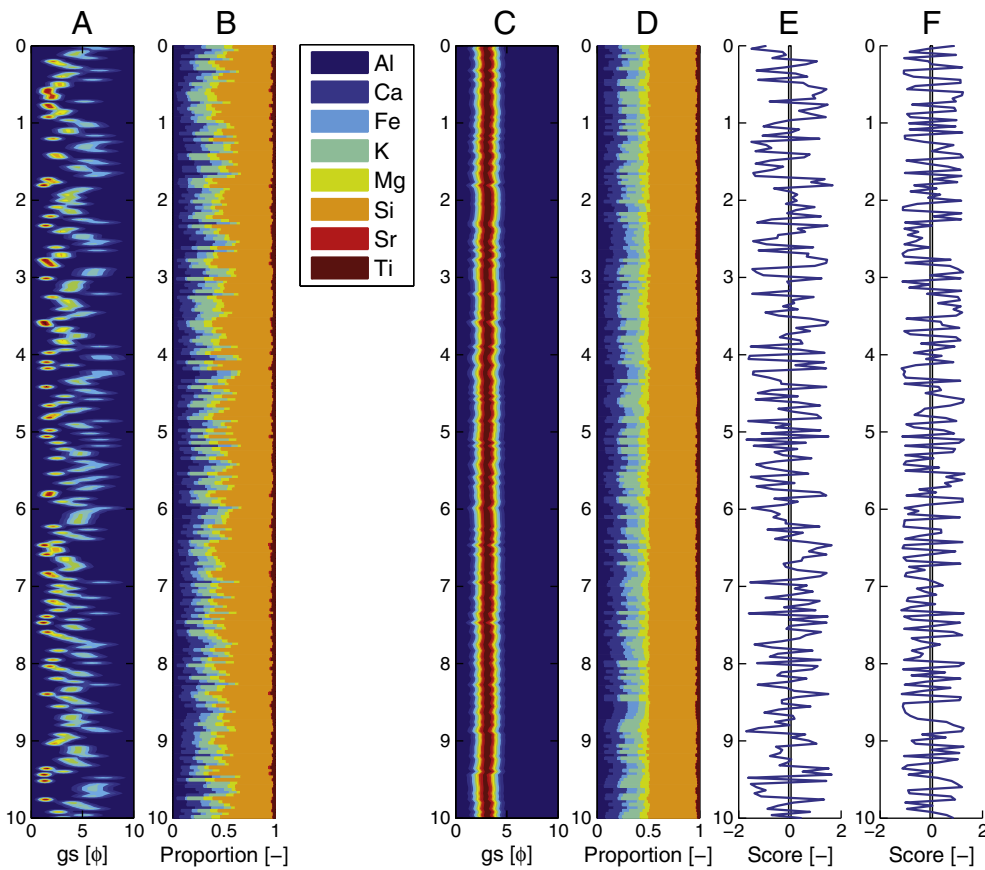
We applied the model to the data sets corresponding to the three Quaternary marine sediment cores. However, before studying the model outcome, we analyse the correlation structure between composition and the mean grain size in each of the three cores (see Fig. 8). The steepness of the fitted log-ratio linear models of cores GeoB7920-2 and GeoB9508-5 in particular, demonstrate that the geochemical composition strongly depends on the mean grain size. Based on these plots, we also conclude that our data shows no systematic

departures from the compositional linear trend, which justifies the use of a log-ratio linear approach.

Fig. 8D shows Pearson's correlation coefficients between the different log-ratio-transformed chemical elements and the mean grain size. In addition, the correlation coefficients between the mean grain size and the PC1 scores, derived from the raw geochemical data, are shown. In core GeoB9508-5, the high correlation coefficient ( $>0.9$ ) between these PC1 scores and the mean grain size demonstrates the large portion of redundancy between geochemistry and grain size. In core GeoB7139-2, on the other hand, the correlation coefficient of 0.15 shows that grain-size control on the geochemical variability is low; only Ti shows a fairly strong (negative) correlation with the mean grain size. It should be noted, though, that the total geochemical variance in this core is also very low.

After applying the proposed model to the three cores, significance tests on the residual records ( $\alpha = 5\%$ ) allows us to identify additive log-ratio pairs which show significant residual variability (see Table 2). The residual signals having the highest signal-to-noise ratio are shown in Figs. 9 and 10. In contrast to the synthetic example, the residuals are plotted without their mean added, so they simply reflect the difference between modeled and observed signals. An exception is made for  $k = 0$ ; since these patterns are by definition only a centred version of the input data, we show the raw data instead.

The downcore plots demonstrate that the variance of the residuals is largely controlled by the number of removed components (the order of the residuals). In addition, the cores show a marked difference, in the sense that the third-order residual geochemical signal strength ranges from weak (GeoB7920), to strong (GeoB9508-5). Note that the limited number of chemical elements ( $D = 5$ ) used in our study limits the maximum number of removed shared signals to four.



**Fig. 6.** This figure shows the data and model outcome of synthetic dataset 2 ( $n = 250$ ), which is constructed by random mixing of two different provenances. Figure 6A shows the grain-size distributions and Fig. 6B shows the geochemical record. The residual grain size and geochemical record are shown in C and D, respectively. The first principal component of the shared geochemical signal is shown in E, whereas plot F shows the first principal component scores of the residual geochemical signal. In both E and F, a gray bar represents a 95% confidence interval derived from the imposed noise. Note that the residual PC1 scores (F) exceed the 95% confidence interval at almost every stratigraphic level.

## 6. Discussion

### 6.1. Correlation between composition and grain size

The correlation coefficients in Fig. 8D show that in cores GeoB7920-2 and GeoB9508-5, Al and Fe correlate positively with the mean grain size whereas Si correlates negatively with the mean grain size. These trends are in line with typical chemical weathering trends, leading to relative enrichment of clay-minerals (reflected by Al and Fe) in the fine fraction, and enrichment of quartz (reflected by Si) in the coarse fraction. In core GeoB7139-2 the correlation structure is different, in the sense that Si shows a weak, but positive correlation with the mean grain size. Moreover, Ti is the only element closely linked to grain size, reflected by the highly negative correlation with the mean grain size in phi-units.

The weak grain-size control on the geochemical composition in this core may be due to continental climate conditions. Core GeoB7139-2 is situated close to the Atacama desert, which is characterised by extreme aridity (Stuut and Lamy, 2004; Kaiser et al., 2008). The absence of chemical weathering under arid conditions implies that the size-composition trend primarily reflects mechanical weathering, whose impact on the size-composition trend is apparently small. The weak, but positive correlation between Si and the mean grain size may be attributable to additional input of biogenic silica, which obviously disturbs the initial size-composition trend of the terrestrially derived sediments.

Based on these results, we may select a proxy for grain size by choosing a ratio of an element showing a high correlation coefficient, and an element showing a very low (i.e., negative) correlation coefficient with the mean grain size. This yields Al/Ti (GeoB7139-2), Ti/Si (GeoB7920-2) and Fe/Si (GeoB9508-5). Differences among these

proxies illustrate the empirical nature of grain-size proxies, and other geochemically-based proxies in general.

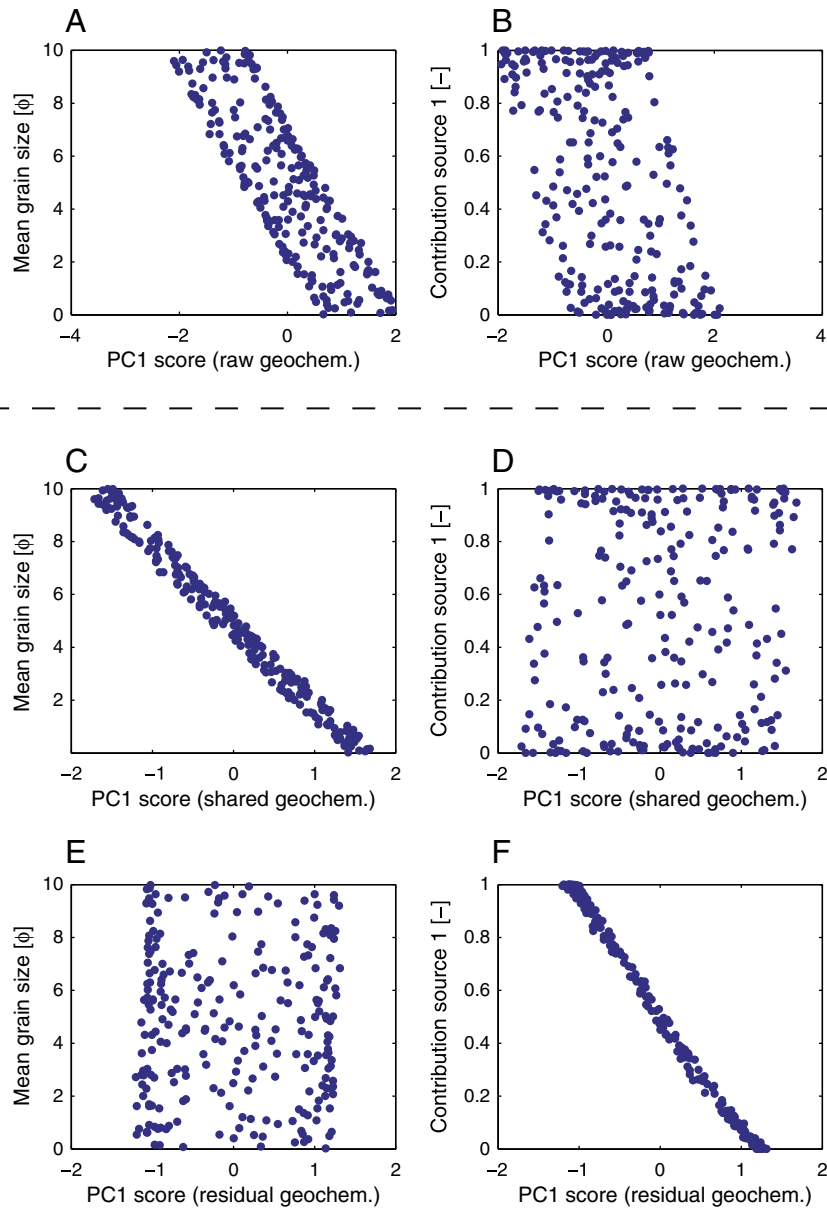
### 6.2. Residuals analysis

In order to evaluate the modeling results in a statistically correct manner, we analyse only the residual log-ratios for which the significance test yields a p-value smaller than  $\alpha (= 5\%)$ . (see Table 2). The test yields that, depending on the order, the cores reflect residual patterns of Ti (GeoB9508-5), K (GeoB7920) and Al/K (GeoB7139-2). Whereas in the synthetic example, the number of dimensions of the grain-size data was exactly known, in reality this is certainly not the case. It is obvious that the more shared signals we remove from the data, the smaller the variance of the residual records. The limitations of our data are reached at  $D - 1$ , which equals four in these data set (we measured only five elements). At this point, there is still some residual grain-size variability.

Following the principle of parsimony, we should stop subtracting signals when the shared grain size variability is within acceptable limits of the measured grain variability, or when the  $k$ -th order shared grain size signal does not correlate with the  $k$ -th order shared geochemical signal. This means that quantitative control on the model outcome does not only require uncertainties of the geochemical data, but also of the grain size data. Unfortunately, we can only infer the uncertainties associated with the geochemical data, meaning that we are not able to identify the 'true' order of the residuals. However, we can interpret the results for the different cores relative to each other.

The downcore records in Figs. 9 and 10 suggest that from these three cores, GeoB9508-5 contains the strongest residual signals. The same is reflected by the large number of log-ratios that are identified as significant (Table 2). We also observe that within one model order, Al/Ti,





**Fig. 7.** Crossplots in plots A and B show the PC1 scores of the raw geochemical data against the mean grain size, and the contribution from source 1, respectively. Note that the raw PC1 scores correlate poorly with both provenance and grain size. Plot C, on the other hand, shows that the correlation between the shared geochemical signal scores and the mean grain size is strong. Similarly, plot F shows that residual scores correlate strongly with the contribution from source 1.

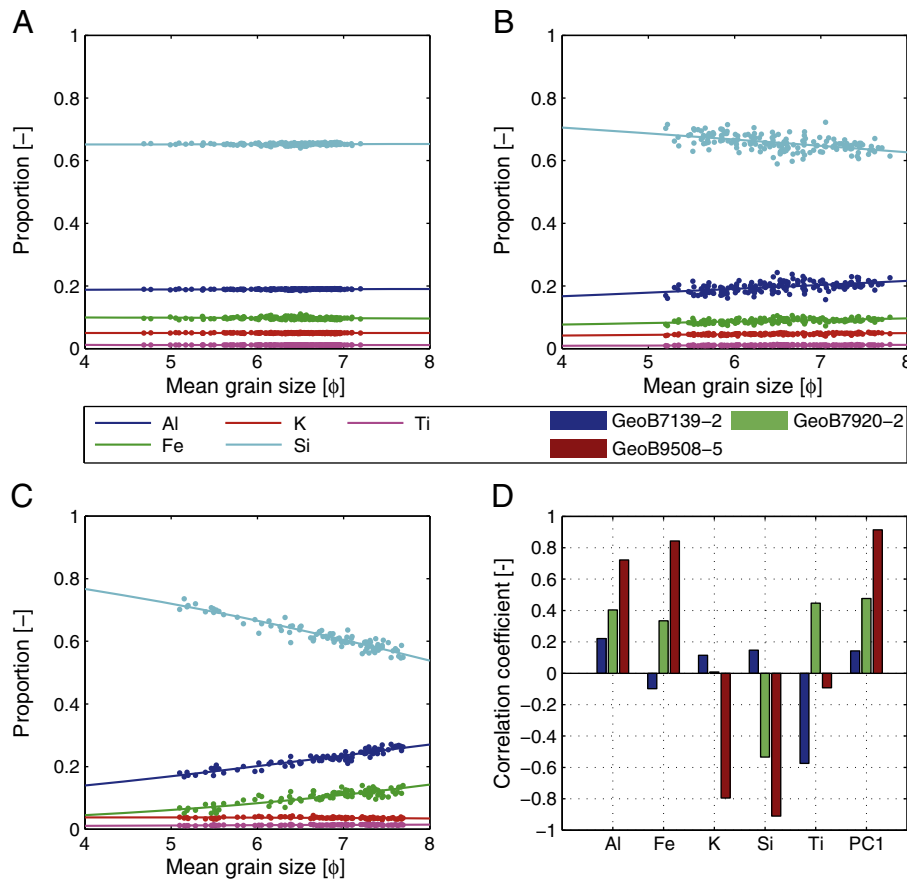
Fe/Ti and Si/Ti are highly correlated. This suggests that the residual variability is attributable to changes in Ti, rather than the other elements. Apparently, sediments with a constant grain-size distribution have different Ti-concentrations. This result is supported by analyses of present-day dust samples collected off northwest Africa; dust samples with the same mean grain size showed different Al/Ti values ( $\text{Log}(\text{Al}/\text{Ti}) \sim [2.65, 2.9]$ ; *Stuut et al., 2005*). Moreover, the variability of these Al/Ti values is in the same range as the residuals observed in this study.

Finally we observe that for every alternative model (i.e., model outcome of a certain order), the residual Al/Ti, Fe/Ti and Si/Ti show several profound dips, suggesting a relative Ti-enrichment. Based on the consistency of this pattern among the different model orders, the plausibility of Ti reflecting residual variability and the fact that the fit to the grain size exhibits no significant improvements when subtracting four instead of three components, we assume that the ‘true order’ of the residuals in core GeoB9508-5 is three. For this order we obtain significant Al/Ti, Fe/Ti, Si/Ti, as well as K/Si residuals.

The residual grain size signals in cores GeoB7139-2 and GeoB7920-2 (*Fig. 9*) suggest that at least three common patterns should be removed to obtain a grain-size invariant geochemical record. At this point, there are no significant residual geochemical signals present in the data (see *Table 1*). As a result, we conclude that these cores contain no significant residual geochemical variability.

### 6.3. Paleo-climatological interpretation

In order to relate the residuals to the proposed size-independent mechanisms, mineralogical information is indispensable. Although mineralogical data is not available, chemical elements can be associated with mineral phases, though with a varying level of confidence. In core GeoB9508-5, K most likely, but not exclusively, resides in K-feldspar whereas Si resides in detrital quartz and possibly in opal. Because (1) the signal-to-noise ratio of the associated residual signal (i.e., K/Si) is 60–80% smaller than that of the Ti residuals, and (2) we cannot link



**Fig. 8.** The chemical composition as function of the mean grain size of cores GeoB7139-2 (A), GeoB7920-2 (B) and GeoB9508-5 (C). The line passing through the observations reflects the least-squares solution to a log-ratio linear model. The correlation between the different clr-transformed chemical components and the mean grain size is shown in D. In addition, the correlation coefficients between the mean grain size and the first principal component of the raw clr-transformed geochemical data are shown.

these elements to minerals with an acceptable level of confidence, we do not attempt to interpret this residual signal.

Ti can be more confidently associated with mineral phases; it is considered to be of exclusively terrigenous origin (Murray and Leinen, 1996) and mostly resides in heavy minerals such as sphene, rutile and anatase (Spears and Kanaris-Sotiriou, 1976). These minerals have a high density relative to minerals such as quartz, feldspars and clays, which make up the bulk of siliciclastic sediments. Although Ti-rich heavy minerals are considered essentially inert components (Young and Nesbitt, 1998), authigenic formation of these minerals have been described under highly specific conditions (Goldberg and Arrhenius, 1958; Pe-Piper et al., 2011). However, given the fact that the observed

**Table 2**

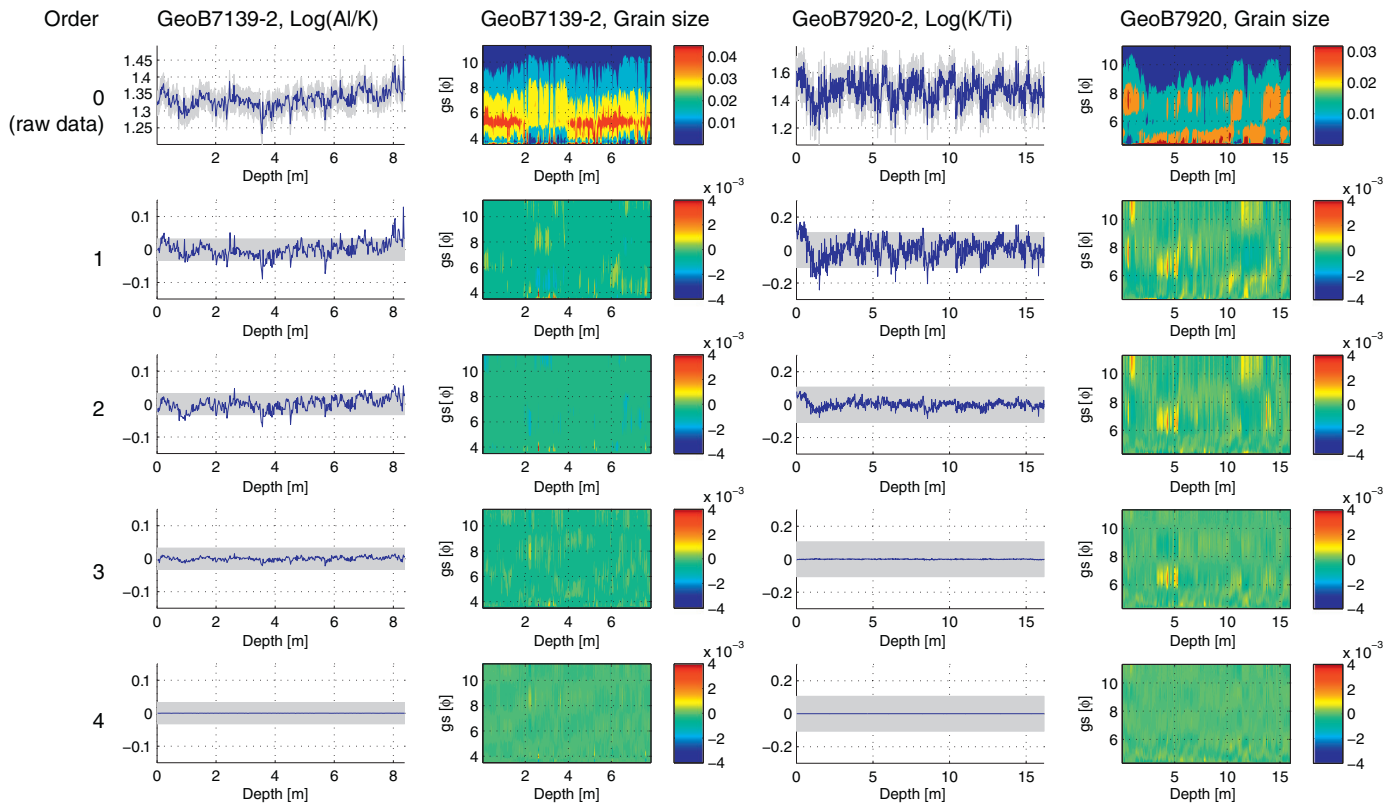
The significant residuals of the different cores ( $\alpha=5\%$ ). A **\*\*** indicates that all elements (except the numerator) exhibit significant residuals.

Core	Order	Significant residuals
GeoB7139-2	1	Al/K
GeoB7139-2	2	Al/K
GeoB7139-2	3	None
GeoB7139-2	4	None
GeoB7920-2	1	K/Ti, K/Al
GeoB7920-2	2	None
GeoB7920-2	3	None
GeoB7920-2	4	None
GeoB9508-5	1	Al/*, Ti/*, K/Si
GeoB9508-5	2	Al/Si, Ti/*, K/Si
GeoB9508-5	3	Al/Ti, Fe/Ti, Si/Ti, K/Si
GeoB9508-5	4	None

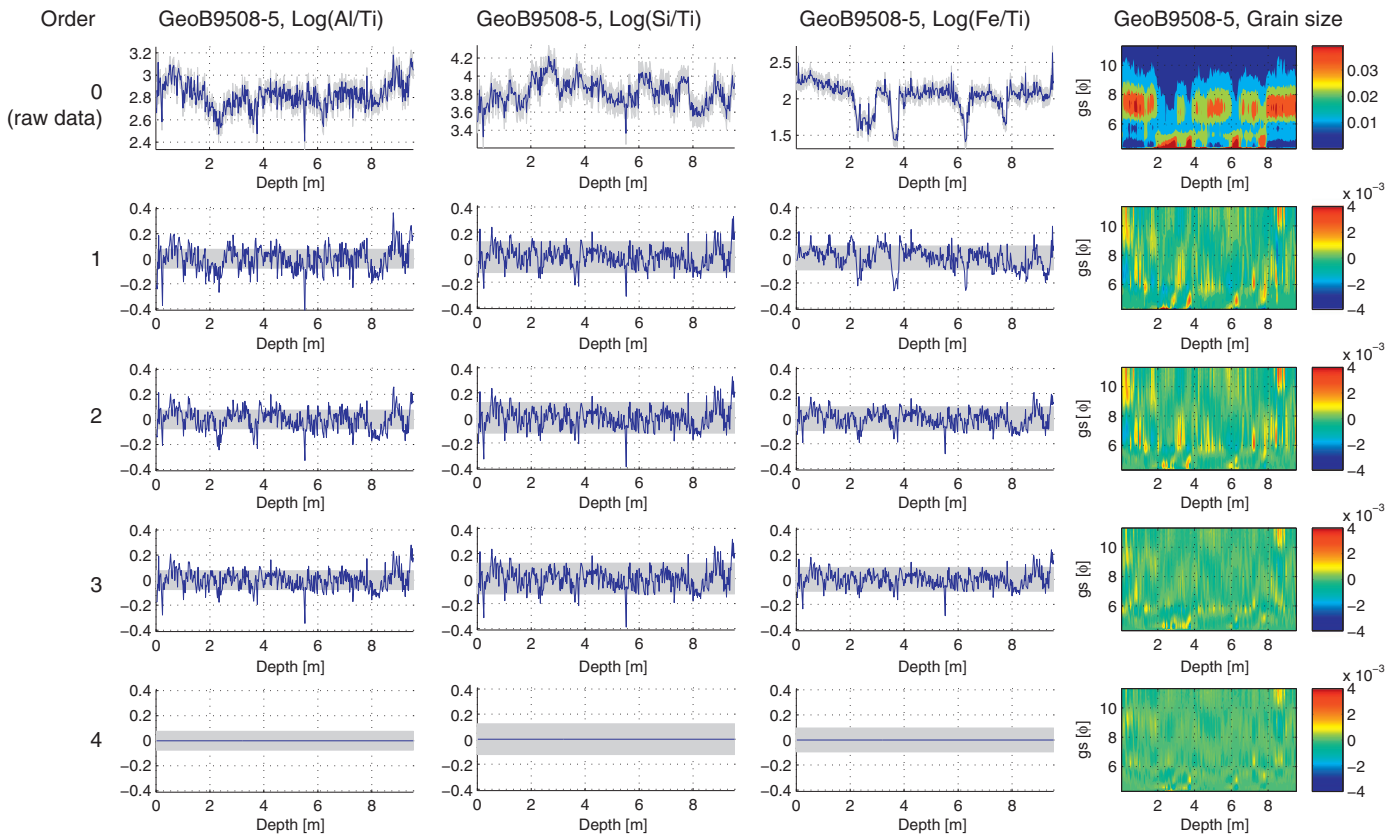
Al/Ti are approximately equal to Al/Ti values found in other marine sediment cores located in the Atlantic (Zabel et al., 1999), we consider diagenetic modification a highly unlikely cause for the observed Ti residuals. This implies that they may be explained by any of the postulated grain-size independent processes that take place before deposition such as weathering, density/shape sorting and mixing.

In the area of core GeoB9508-5, contrasting patterns of deposition have been linked to climate changes (Mulitza et al., 2008). Dry and cold periods were characterised by strong winds and low fluvial sediment input, whereas the opposite holds for relatively warm and humid intervals (e.g., Sarnthein, 1978; deMenocal et al., 1993). Hence, we are dealing with a situation in which the grain-size characteristics of the sediments, and therefore also their transport mechanisms, are likely to correlate with provenance. We make use of this knowledge by dividing the residual geochemical data set into aeolian-dominated, fluvial-dominated, and mixed sediments samples, based on the median grain size (P50 of the cumulative grain size distributions) with cutoffs set at 5.5 and 6.5 phi-units (Fig. 11A). The relation between transport mechanism and residual Ti/Al is shown in Fig. 11B. With the null hypothesis being that the Log(Ti/Al) in the fluvial and aeolian source are samples from identical distributions having the same median, we estimate the probability under the null hypothesis using a Wilcoxon rank sum test (Hollander and Wolfe, 1973). Based on this p-value and a significance level of 5%, we conclude that the third-order Ti/Al residuals are related to the transport mechanism (and therefore also the Al/Ti residuals, because log-ratios are symmetrical Weltje and Tjallingii, 2008).

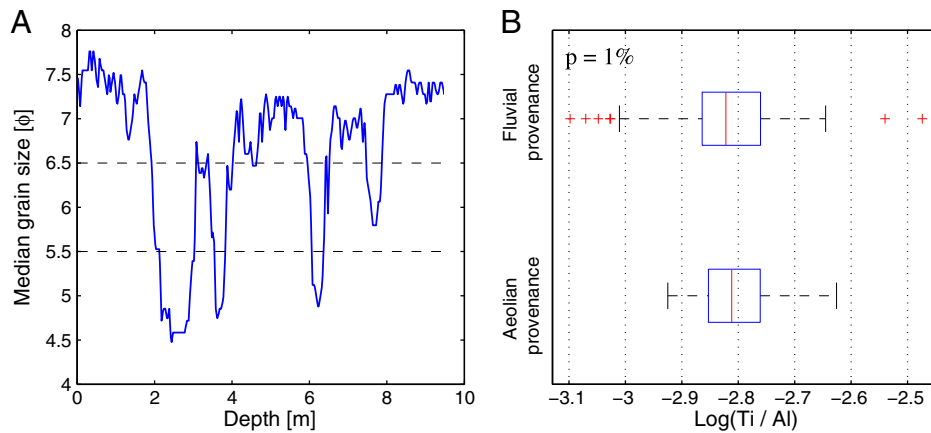
We can postulate several scenarios that may have produced the residual patterns, which differ with respect to the underlying



**Fig. 9.** The plots show the strongest (in terms of signal-to-noise ratio) downcore residual signals in cores GeoB7139-2 and GeoB7920-2, with the model order ranging between zero (the raw data) and four. The residual geochemical signals are presented as variations around the mean, with a gray bar indicating a 95% confidence limit. The residual grain size distributions are presented as the difference between the raw and modeled clr-transformed grain size distributions.



**Fig. 10.** The plots show the three strongest downcore residual signals (the signal-to-noise ratio decreases from left to right) in cores GeoB9508-5, with the model order ranging between zero (the raw data) and four. The residual geochemical signals are presented as variations around the mean, with a gray bar indicating a 95% confidence limit. The residual grain size distributions are presented as the difference between the raw and modeled clr-transformed grain size distributions.



**Fig. 11.** As indicated in A, we partitioned the data into fluvial, aeolian and mixed provenance, based on the median grain size. For both the fluvial and aeolian sediment, we show the residual Ti/Al in plot B. Based on the p-value derived from the Wilcoxon rank sum test, and a significance level of 5%, we conclude that the third-order residuals are related to the transporting medium.

assumptions. If, for instance, the sediment deposited in the sink has been exclusively derived from one source, and the conditions under which the sediment was transported were constant in time, the residual signal reflects varying climatic conditions in the source area(s). An alternative scenario is that the residuals reflect the different areas that have served as a source, rather than varying conditions within the source area(s). In both cases, the underlying assumption is that the sediment is transport-invariant, i.e., that it has been transported under the same hydro- and aerodynamic conditions (Weltje, 2004). Another scenario is that the residual variability is induced by mixing of fluvial and aeolian sediment and/or sorting processes during either river- or wind-transport. This scenario is characterised by the assumption of "provenance-invariance".

It is evident that, without additional constraints, there is no unique solution to this problem. Even when adopting the (strong) assumption that the source area and the climatic conditions were constant in time, we cannot ascribe the residuals to either selective transport or mixing. This is due to the fact that preferential entrainment of Ti-bearing minerals in the windblown sediment due to increased wind strength (Sarnthein, 1978; Rea, 1994) will have the same effect on Ti residuals as mixing of Ti-rich windblown sediment and Ti-poor riverine sediment having a constant geochemical composition. Moreover, Stuut et al. (2005) showed that even on a short time-scale (days), both source area and pathway length (which controls selective transport) substantially vary, which means that assuming provenance-invariance, as well as transport-invariance, are both not defensible. We therefore conclude that, on the basis of these data, it is impossible to unambiguously interpret the residuals in terms of the proposed grain-size independent mechanisms.

#### 6.4. General discussion

In the analysed cores, the elements showing the strongest correlation with the mean grain size were Ti (GeoB7139-2) and Si (GeoB7920-2 and GeoB9508-5). In addition, Ti was designated as the main residual component in GeoB9508-5. This leads to the conclusion that geochemistry-based proxies for mean grain size depend on the specific setting, which automatically implies that the same is true for geochemistry-based proxies considered to be associated with other processes. We showed that, using the model and a set of overlapping grain-size and geochemical observations, these proxies can be identified in a robust way. However, whether or not signals associated with the size-independent processes can be inferred from the data depends on the magnitude of this signal, relative to the uncertainty in both data sets. Likewise, more significant signals may

appear if the uncertainty of the input data will be reduced. Furthermore, applying the model to a more comprehensive set of geochemical variables is more likely to reveal informative residual patterns.

As for the interpretation of bulk geochemical records in general, interpretation of the residuals is difficult because they may reflect numerous different processes. In addition, elements cannot always be uniquely associated with a single mineral. We can thus not provide a 'cookbook' for the interpretation of the residuals. However, by correcting the data for grain size, at least one hypothesis can be eliminated, i.e., that the geochemistry only reflects grain-size variations.

#### 7. Conclusions

The joint analysis of bulk geochemical records and grain-size distributions allows us to decompose the geochemical record into a grain-size independent and a grain-size related part. This method therefore facilitates identification of grain-size proxies, and interpretation of residual geochemical signals in terms of weathering, shape/density sorting, mixing and diagenesis. Differences between the size-compositional trends highlight the need for empirical models; in core GeoB7139-2 (offshore Chile) Ti and Si were, respectively, strong and weak proxies for the mean grain size whereas in core GeoB9508-5 (offshore Senegal) it was the other way around. Moreover, the model showed that the residual Ti in the case of core GeoB9508-5 varies independently of the grain size, and discriminates between the two known sources (fluvial and aeolian). Several hypotheses were postulated on the basis of this model outcome although they could not be confirmed nor rejected without additional information. From the synthetic example and the case study we conclude that multi-proxy analysis can be formalised using multivariate methods (e.g., Partial Least Squares). However, to apply the methodology in future studies, comprehensive quantification of the uncertainties in grain-size and geochemical data is necessary.

#### Acknowledgments

Inka Meyer is thanked for laboratory assistance during size separation of the sediment samples. M.R. Bloemsma gratefully acknowledges funding by Wintershall. G.J. Weltje acknowledges financial support of the Hanse Institute for Advanced Study, Delmenhorst, Germany, for initiating this research during a stay as Research Fellow in Ocean Sciences. We thank editor Hilmar von Eynatten, Georgia Pe-Piper, David Piper, Raimon Tolosana-Delgado and an anonymous reviewer for their constructive comments.

## Appendix A. Joint modelling of grain size and geochemical composition

Bulk chemistry and grain-size distribution are both compositional data. The compositional nature of a data set spanning  $D$  components implies that (Aitchison, 1986):

$$\mathbf{x} = (x_1, x_2, \dots, x_D)', x_i > 0, \sum_{i=1}^D x_i = \kappa \quad (\text{A.1})$$

where  $\kappa$  may be 1, 100 or 1.000.000, depending on the units of measurement. The constant-sum and positivity constraint make that a  $D$ -part compositional dataset  $\mathbf{x} = \{x_1, x_2, x_3, \dots, x_D\}$  does not span a real space  $\mathbb{R}^D$ , but the so called  $D$ -part simplex, or  $S^D$ , which has rank  $D-1$ .

From a modelling perspective, the most important implication of the compositional nature is that classical statistical methods cannot be applied. Instead, processing of compositional data should be done using log-ratio transformations, or log-ratios. For multivariate statistical analysis, the data should be centred log-ratio (clr) transformed (Aitchison, 1986):

$$\mathbf{y} = \text{clr}(\mathbf{x}) = \left( \log \frac{x_1}{g(\mathbf{x})}, \log \frac{x_2}{g(\mathbf{x})}, \dots, \log \frac{x_D}{g(\mathbf{x})} \right) \quad (\text{A.2})$$

where  $g(\mathbf{x})$  denotes the geometric mean over all components.

The objective of this study is to decompose the geochemical record into a part which is correlated with grain size, and a part which varies independently from the grain size. We achieve this by finding a basis  $Q$  in  $\mathbb{R}^D$  (the clr-transformed geochemical solution space) which maximizes the geochemical variance explained by the grain size. A multivariate method to maximize the common covariance, and obtain this basis  $Q$ , is Partial Least Squares, or PLS (Wold et al., 1982), which is the method being used in this study.

If we have measured both the grain-size distribution  $X$  ( $L$  grain-size classes) and the geochemical composition  $Y$  ( $D$  variables) of a set of  $m$  specimens, we fit the model on data matrices  $X$ , and  $Y$ , which contain the grain size and bulk chemical composition of this set of samples, respectively. If the mean is subtracted from these data matrices, we obtain  $X_{c*}$  and  $Y_{c*}$ . In this study, we then use the SIMPLS algorithm (de Jong, 1993) to calculate the PLS matrix decomposition:

$$\begin{aligned} X_{c*} &= T_* P^T \\ Y_{c*} &= U_* Q^T \end{aligned} \quad (\text{A.3})$$

where  $T$  and  $U$  are the  $m \times D$  and  $m \times L$  scores, and  $P$  and  $Q$  are the loadings of size  $D \times D$  and  $L \times L$ , respectively. To ensure the decomposition to be orthogonal, the bases  $P$  and  $Q$  are orthogonalised using Singular Value Decomposition (Press et al., 1994) after which the score matrices  $T_*$  and  $U_*$  are recalculated (for the methodology, see Eq. A.5). Then, we determine the maximum number of components that may be removed by performing a test on the correlation between the scores  $T_*$  and  $U_*$ . Significance of correlation between the geochemical and grain-size scores on the  $k$ -th basis vector (i.e., the  $k$ -th column of  $T_*$  and  $U_*$ ) is tested using the following criterion (Kendall and Stuart, 1973):

$$|r| \sqrt{\frac{m-2}{1-r^2}} > t^{-1}(m-2, p). \quad (\text{A.4})$$

We derive  $r$  for any order  $k$ , where  $r$  is Pearson's correlation coefficient between the  $k$ -th column of  $T_*$  and  $U_*$  (i.e. the projection of the data on the  $k$ -th basis vector). The right-hand side is the inverse  $t$ -

distribution with probability  $p$  and  $m-2$  degrees of freedom. Given a level of confidence  $\alpha$  and  $p = 1 - \alpha$ , we remove the first  $k$  shared signals if for  $k$ , the criterion in Eq. A.4 is met.

Now the model order and the model parameters (i.e., the linear bases) are calculated, we can apply the model to all grain-size and geochemical observations in the data set, i.e., also the observations related to samples contained in either the grain-size or the geochemical data set. Because  $P^T$  and  $Q^T$  are both orthonormal, it holds that  $(P^T)^{-1} = (P^T)^T = P$  and  $(Q^T)^{-1} = (Q^T)^T = Q$ . This, in turn, means that the scores of all observed grain-size distributions and geochemical compositions on their corresponding bases can be derived by matrix product:

$$\begin{aligned} T &= X_c P \\ U &= Y_c Q \end{aligned} \quad (\text{A.5})$$

Subsequently, the shared signal in both datasets of rank  $k$  can be derived by means of a reduced-rank approximation. This implies that we use only the first  $k$  columns of both the loadings  $P$  and  $Q$ , and the scores  $T$  and  $U$ :

$$\begin{aligned} (\hat{X}_c)_k &= T_k P_k^T \\ (\hat{Y}_c)_k &= U_k Q_k^T \end{aligned} \quad (\text{A.6})$$

The 'residuals' are obtained by subtracting the common variability in both data sets from the input data:

$$\begin{aligned} (\hat{X}_c)_k &= X_c - (\hat{X}_c)_k \\ (\hat{Y}_c)_k &= Y_c - (\hat{Y}_c)_k \end{aligned} \quad (\text{A.7})$$

Finally, we may add the mean again to form  $(\hat{X})_k$  and  $(\hat{Y})_k$ , so that the residual signals centre around the mean of their corresponding raw data matrix.

## References

- Aitchison, J., 1986. The Statistical Analysis of Compositional Data, 1st Edition. Chapman and Hall, London.
- Amante, C., Eakins, B.W., 2008. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. National Geophysical Data Center, NESDIS, NOAA, U.S. Department of Commerce. (map publicly available at: <http://www.ngdc.noaa.gov/mgg/global/global.html>).
- Amorosi, A., Centineo, M.C., Dinelli, E., Lucchini, F., Tateo, F., 2002. Geochemical and mineralogical variations as indicators of provenance changes in Late Quaternary deposits of SE Po Plain. *Sedimentary Geology* 151, 273–292.
- Basu, A., 2003. A perspective on quantitative provenance analysis. In: Valloni, R., Basu, A. (Eds.), *Quantitative Provenance Studies in Italy, Memorie Descrittive della Carta Geologica dell'Italia*, Vol. 64, pp. 11–22.
- Boyle, E.A., 1983. Chemical accumulation variations under the Peru Current during the past 130,000 years. *Journal of Geophysical Research* 88 (C12), 7667–7680. <http://dx.doi.org/10.1029/JC088iC12p07667>.
- Calvert, S.E., Pederson, T.F., 2007. Elemental proxies for palaeoclimatic and palaeoceanographic variability in marine sediments: interpretation and Application. In: Hillaire-Marcel, C., de Vernal, A. (Eds.), *Proxies in Late Cenozoic Paleoclimatology*. Developments in Quaternary Research, vol. 1. Elsevier Science, Amsterdam, pp. 567–644.
- de Jong, S., 1993. SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* 18, 251–263.
- De Pol-Holz, R., Ulloa, O., Lamy, F., Dezileau, L.P.S., Hebbeln, D., 2007. Late Quaternary variability of sedimentary nitrogen isotopes in the eastern South Pacific Ocean. *Paleoceanography* 22, pA2207.
- deMenocal, P.B., Ruddiman, W.F., Pokras, E.M., 1993. Influences of high- and low-latitude processes on African terrestrial climate: Pleistocene eolian records from equatorial Atlantic ocean drilling program Site 663. *Paleoceanography* 8, 209–242.
- Dinelli, E., Tateo, F., Summa, V., 2007. Geochemical and mineralogical proxies for grain size in mudstones and siltstones from the Pleistocene and Holocene of the Po River alluvial plain, Italy. In: Arribas, J., Critelli, S., Jonhsson, M. (Eds.), *Sedimentary Provenance and Petrogenesis: Perspectives from Petrography and Geochemistry*: Geological Society of America Special Papers, pp. 25–36.
- Gac, J.Y., Kane, A., 1986. Le fleuve Senegal: I. Bilan hydrologique et flux continentaux de matieres particulaires a l'embouchure. *Sciences Géologiques Bulletin* 39, 99–130.
- Geisser, S., 1993. *Predictive Inference: An Introduction*. Chapman and Hall, New York.
- Goldberg, E.D., Arrhenius, G.O.S., 1958. Chemistry of pacific pelagic sediments. *Geochimica et Cosmochimica Acta* 13, 153–212.

- Haslett, S.K., Davies, C.F.C., 2006. Late Quaternary climate-ocean changes in western North Africa: offshore geochemical evidence. *Transactions of the Institute of British Geographers* 31, 34–52.
- Haslett, S.K., Smart, C.W., 2006. Late Quaternary upwelling off tropical NW Africa: new micropalaeontological evidence from ODP Hole 658C. *Journal of Quaternary Science* 21, 259–269.
- Hebbeln, D., cruise participants, 2001. PUCK, Report and Preliminary Results of R/V Sonne cruise 156, Valparaiso (Chile)–Talcahuano (Chile), March 29–May 14, 2001. Univ. Bremen, Bremen, Germany.
- Hollander, M., Wolfe, D.A., 1973. *Nonparametric Statistical Methods*. Wiley.
- Holz, C., Stuut, J.-B.W., Henrich, R., 2004. Terrigenous sedimentation processes along the continental margin off NW Africa: implications from grain-size analysis of seabed sediments. *Sedimentology* 51, 1145–1154.
- Johnsson, M.J., 1993. The system controlling the composition of clastic sediments. In: Johnsson, M.J., Basu, A. (Eds.), *Processes Controlling the Composition of Clastic Sediments*: Geological Society of America Special Paper, 284, pp. 1–19.
- Kaiser, J., Schefuss, E., Lamy, F., Mohtadi, M., Hebbeln, D., 2008. Glacial to Holocene changes in sea surface temperature and coastal vegetation in north central Chile: high versus low latitude forcing. *Quaternary Science Reviews* 27, 2064–2075.
- Kendall, M.G., Stuart, A., 1973. *The Advanced Theory of Statistics, Volume 2: Inference and Relationships*. Griffin, London.
- Klump, J., Hebbeln, D., Wefer, G., 2000. The impact of sediment provenance on barium-based productivity estimates. *Marine Geology* 169, 259–271.
- Koopmann, B., 1981. Sedimentation von Saharastaub im subtropischen Nordatlantik während der letzten 25.000 Jahre. *Meteor Forschungsergeb., Reihe C*, 35, pp. 23–59.
- Lamy, F., Hebbeln, D., Wefer, G., 1998. Terrigenous sediment supply along the Chilean continental margin: modern regional patterns of texture and composition. *Geologische Rundschau* 87, 477–494.
- Lamy, F., Klump, J., Hebbeln, D., Wefer, G., 2000. Late Quaternary rapid climate change in northern Chile. *Terra Nova* 12, 8–13.
- Martens, H., Naes, T., 1989. *Multivariate Calibration*. Wiley, Chichester.
- McLennan, S.M., Hemming, S., McDaniel, D.K., Hanson, G.N., 1993. Geochemical approaches to sedimentation, provenance and tectonics. In: Johnsson, M., Basu, A. (Eds.), *Processes Controlling the Composition of Clastic Sediments*: Geological Society of America, Special Papers, 285, pp. 21–40.
- Meggers, H., cruise participants, 2002. Report and Preliminary Results of Meteor cruise M 53/1, Limasol – Las Palmas – Mindelo, 30.03–03.05.2002. Univ. Bremen, Bremen, Germany.
- Moreno, T., Querol, X., Castillo, S., Alastuey, A., Cuevas, E., Herrmann, L., Mounkaila, M., Elvira, J., Gibbons, W., 2006. Geochemical variations in aeolian mineral particles from the Sahara-Sahel Dust Corridor. *Chemosphere* 65, 261–270.
- Mulitza, S., cruise participants, 2006. Report and Preliminary Results of R/V METEOR Cruise M65/1, Dakar–Dakar, 11.06.2005–01.07.2005. Univ. Bremen, Bremen, Germany.
- Mulitza, S., Prange, M., Stuut, J.-B., Zabel, M., von Dobeneck, T., Itambi, A.C., Nizou, J., Schulz, M., Wefer, G., 2008. Sahel megadroughts triggered by glacial slowdowns of Atlantic meridional overturning. *Paleoceanography* 23, PA4206. <http://dx.doi.org/10.1029/2008PA001637>.
- Murray, R.W., Leinen, M., 1996. Scavenged excess aluminum and its relationship to bulk titanium in biogenic sediment from the central equatorial Pacific Ocean. *Geochim. Cosmochim. Acta* 60, 3869–3878.
- Nesbitt, H.W., Young, G.M., 1984. Prediction of some weathering trends of plutonic and volcanic rocks based on thermodynamic and kinetic considerations. *Geochimica et Cosmochimica Acta* 48, 1523–1548.
- Nesbitt, H.W., Young, G.M., 1996. Petrogenesis of sediments in the absence of chemical weathering: effects of abrasion and sorting on bulk composition and mineralogy. *Sedimentology* 43, 341–358.
- Pe-Piper, G., Triantafyllidis, S., Piper, D.J.W., 2008. Geochemical identification of clastic sediment provenance from known sources of similar geology: the Cretaceous Scotian basin, Canada. *Journal of Sedimentary Research* 78, 595–607.
- Pe-Piper, G., Karim, A., Piper, D.J.W., 2011. Authigenesis of titania minerals and the mobility of Ti, new evidence from pro-deltaic sandstones, cretaceous Scotian Basin, Canada. *Journal of Sedimentary Research* 81, 762–773.
- Pettijohn, F.J., Potter, P.E., Siever, R., 1987. *Sand and Sandstone*, Second Edition. Springer, New York.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1994. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Second Edition. Cambridge University Press, Cambridge.
- Rea, D.K., 1994. The paleoclimatic record provided by Eolian deposition in the deep sea: the geologic history of wind. *Geophysics* 32, 159–195.
- Richter, T.O., Van der Gaast, S., Koster, B., Vaars, A., Gieses, R., de Stigter, H., de Haas, H., van Weering, T.C.E., 2006. The Avatech XRF core scanner: technical description and applications to NE Atlantic sediments. In: Rothwell, R.E. (Ed.), *New Techniques in Sediment Core Analysis*. Special Publication, Vol. 267. Geological Society, London, pp. 39–50.
- Romero, O.E., Kim, J.H., Donner, B., 2008. Submillennial-to-millennial variability of diatom production off Mauritania, NW Africa during the last glacial cycle. *Paleoceanography* 23, PA3218. <http://dx.doi.org/10.1029/2008PA001601>.
- Sarnthein, M., 1978. Sand deserts during glacial maximum and climatic optimum. *Nature* 272, 43–45.
- Sarnthein, M., Tetzlaff, G., Koopmann, B., Wolter, K., Pflaumann, U., 1981. Glacial and interglacial wind regimes over the eastern subtropical Atlantic and north-west Africa. *Nature* 293, 193–196. <http://dx.doi.org/10.1038/293193a0>.
- Schneider, R.R., Price, B., Müller, P.J., Kroon, D., Alexander, I., 1997. Monsoon related variations in Zaire (Congo) sediment load and influence of fluvial silicate supply on marine productivity in the east equatorial Atlantic during the last 200,000 years. *Paleoceanography* 12, 463–481.
- Solano-Acosta, W., Dutta, P.K., 2005. Unexpected trend in the compositional maturity of second-cycle sand. *Sedimentary Geology* 178, 275–283.
- Spears, D.A., Kanaris-Sotiriou, R., 1976. Titanium in some Carboniferous sediments from Great Britain. *Geochimica et Cosmochimica Acta* 40, 345–351.
- Stuut, J.-B.W., Lamy, F., 2004. Climate variability at the southern boundaries of the Namib (southwestern Africa) and Atacama (northern Chile) coastal deserts during the last 120,000 yr. *Quaternary Research* 62, 301–309.
- Stuut, J.-B., Zabel, M., Ratmeyer, V., Helmke, P., Schefuss, E., Lavik, G., Schneider, R., 2005. Provenance of present-day Eolian dust collected off North-Africa. *Journal of Geophysical Research* 110. <http://dx.doi.org/10.1029/2004JD005161>.
- Stuut, J.-B., Kasten, S., Lamy, F., Hebbeln, D., 2007. Sources and modes of terrigenous sediment input to the Chilean continental slope. *Quaternary International* 161, 67–76.
- Tjallingii, R., Röhl, U., Kölling, M., Bickert, T., 2007. Influence of the water content on X-ray fluorescence core-scanning measurements in soft marine sediments. *Geochemistry, Geophysics, Geosystems* 8. <http://dx.doi.org/10.1029/2006GC001393>.
- Tjallingii, R., Claussen, M., Stuut, J.-B., Fohlmeister, J., Jahn, A., Bickert, T., Lamy, F., Röhl, U., 2008. Coherent high- and low-latitude control of the northwest African hydrological balance. *Nature Geoscience* 1 (10), 670–675.
- Tolosana-Delgado, R., von Eynatten, H., 2008. Grain-size control on petrographic composition of sediments: compositional regression and rounded zeros. *Mathematical Geosciences* 41, 869–886.
- Tolosana-Delgado, R., von Eynatten, H., 2010. Simplifying compositional multiple regression: application to grain size controls on sediment geochemistry. *Computers and Geosciences* 36, 577–589.
- Vekemans, B., Janssens, K., Vincze, L., Adams, F., van Espen, P., 1994. Analysis of X-ray spectra by iterative least squares (AXIL): new developments. *X-ray Spectrometry* 23, 278–285.
- von Eynatten, H., Barceló-Vidal, C., Pawlowsky Glahn, V., 2003. Modelling compositional change: the example of chemical weathering of granitoid rocks. *Mathematical Geology* 35 (3), 231–251. <http://dx.doi.org/10.1023/A:1023835513705>.
- Weltje, G.J., 2004. A quantitative approach to capturing the compositional variability of modern sands. *Sedimentary Geology* 171, 59–77.
- Weltje, G.J., Brommer, M.B., 2011. Sediment-budget modelling of multi-sourced basin fills: application to recent deposits of the western Adriatic mud wedge (Italy). *Basin Research* 23, 291–308.
- Weltje, G.J., Prins, M.A., 2003. Muddled or mixed? Inferring palaeoclimate from size distributions of deep-sea clastics. *Sedimentary Geology* 162, 39–62.
- Weltje, G.J., Tjallingii, R., 2008. Calibration of XRF core scanners for quantitative geochemical logging of sediment cores: theory and application. *Earth and Planetary Science Letters* 274, 423–438.
- Weltje, G.J., von Eynatten, H., 2004. Quantitative provenance analysis of sediments review and outlook. *Sedimentary Geology* 171, 1–11.
- Wien, K., Wissmann, D., Külling, M., Schulz, H.D., 2005. Fast application of X-ray fluorescence spectrometry aboard ship: how good is the new portable Spectro Xepos analyzer? *Geo-Marine Letters* 25–4, 248–264.
- Wold, S., Martens, H., Wold, H., 1982. The multivariate calibration problem in chemistry solved by the PLS method. In: Ruhe, A., Kagstrom, B. (Eds.), *Proc. Conf. Matrix Pencils*. Springer Verlag, Heidelberg, pp. 286–293.
- Xiong, S., Ding, Z., Zhu, Y., Lu, H., 2010. A ~ 6 Ma chemical weathering history, the grain size dependence of chemical weathering intensity, and its implications for provenance change of Chinese loess-red clay deposit. *Quaternary Science Reviews* 29, 1911–1922.
- Yarincik, K.M., Murray, R.W., Peterson, L.C., 2000. Climatically sensitive eolian and hemipelagic deposition in the Cariaco Basin, Venezuela, over the past 578,000 years: Results from Al/Ti and K/Al. *Paleoceanography* 15, 210–228.
- Young, G.M., Nesbitt, H.W., 1998. Processes controlling the distribution of Ti and Al in weathering profiles, siliciclastic sediments and sedimentary rocks. *Journal of Sedimentary Research* 68 (3), 448–455.
- Zabel, M., Bickert, T., Dittert, L., Häse, R.R., 1999. Significance of the sedimentary Al:Ti ratio as an indicator for variations in the circulation patterns of the equatorial North Atlantic. *Paleoceanography* 14 (6), 789–799.