ELSEVIER

Marine Geology 184 (2002) 167–187

MARINE GEOLOGY
INTERNATIONAL JOURNAL OF MARINE GEOLOGY, GEOCHEMISTRY AND GEOPHYSICS

www.elsevier.com/locate/margeo

Pitfalls of normalization of marine geochemical data using a common divisor

Cornelis H. Van der Weijden *

Department of Geochemistry, Faculty of Earth Sciences, Utrecht University, Budapestlaan 4, P.O. Box 80.021, 3508 TA Utrecht, The Netherlands

Received 9 April 2001; accepted 14 November 2001

Abstract

Normalization of trace element contents of sediment through division by the content of an immobile element – usually aluminum – is common practice in marine geochemical studies. This might appear to be a simple way of correcting for dilution by sedimentary phases barren of a particular trace element – for instance carbonates or silica – and for comparison with its content in a standard clay or shale. The purpose of this paper is to revive the awareness of the pitfalls of this practice by giving numerical examples of the often-unexpected results of normalization. Taking a statistical perspective, it is first shown that uncorrelated variables acquire spurious correlations when normalized. But normalization can also increase, decrease, change sign of, or even blur the correlations between unmodified variables. Next, a number of realistic scenarios are worked out to show that the correlations between normalized element contents still suffer from the closure effect. Only in a few simple cases it is possible to extract, from the normalized data, realistic estimates of trace element contents in distinct sedimentary phases, such as organic matter. This is ascertained from data on copper contents in sediment from the Black Sea and Arabian Sea. When the coefficient of variation (i.e. standard deviation divided by the mean) of the aluminum data is relatively low and much lower than for the original values of the trace elements, comparisons of correlations can equally well or even better be made on the basis of the unmodified values. Then, the comparison of trace element data with their values for standard shales is discussed. The inherent problem is that the composition of the commonly used standard shale and, consequently, the reference values of normalized elements are not necessarily representative of the local or regional sediments in the study area. Also, this method falls short of identifying the processes responsible for enrichment or diminution in trace element contents. Finally, consideration is given to some aspects of a proper use of *Al* normalization © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Trace element normalization; Ratio correlation; Spurious correlations

1. Introduction

Sediment consists generally of detrital minerals (quartz, feldspars, etc.), clays and carbonates as major constituents, and organic matter and hydrogenous phases (e.g. Fe and Mn oxides, Fe sulfides) as minor components. The percentages of all constituents sum up to 100%, implying that when the content of one component increases, the content of one or more other phases must

* Fax: +31-30-253-5302.
E-mail address: chvdw@geo.uu.nl (C.H. Van der Weijden).

0025-3227/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0025-3227(01)00297-3

become smaller; in other words, each variable must be negatively correlated with at least one other variable. This notorious closure effect is also present in the analyses of mineral-forming oxides. Methods to circumvent the problems of such compositional data have been proposed by Aitchison (1986) and summarized by Rollison (1993). Discussion of the closed sum problem in sedimentary data analysis is, however, not the key issue in this communication, but its effect will play a role in the background.

Bulk chemical analyses of sediment do not naturally reveal which sedimentary phases add to the concentration of a particular element, e.g. the total SiO_2 content has contributions from all (alumino)silicates plus free quartz plus biogenic silica. Such total analyses are insufficient to interpret the sedimentary record in terms of, for instance, (paleo)redox processes or reconstruction of sediment provenance and paleoclimate. However, a commonly adopted procedure is to consider aluminum as only present in clays and detrital aluminosilicates. This is the basis for comparison of *element/Al* values in marine sediment with these ratios in standard shale (Turekian and Wedepohl, 1961; Wedepohl, 1971; Brumsack, 1989; Calvert and Pedersen, 1993; Piper and Isaacs, 1995a,b; Morford and Emerson, 1999). In addition, using this ratio to correct for dilution by sedimentary phases devoid of Al (carbonate, quartz, organic matter, Fe–Mn oxides) seems also an attractive quick and easy way to normalize chemical data and to present Al-normalized element concentrations in sediment profiles. It should be noted that not only Al is used for this purpose, but other elements (e.g. K, Li, Sc, Ga, Zr and Ti) are used as well[1]. Stretching their application even further, normalized element concentrations are often used to demonstrate correlations between elements. Although the geochemical community should be aware of the pitfalls inherent to the use of ratio correlations, this seems not always to be realized (cf. Aitchison, 1986; Berges, 1997). The problem is that three uncorrelated variables, each with a normal distribution, do not retain zero correlations when plotted as ratios with one of them as common divisor (Pearson, 1896; Chayes, 1971; Butler, 1986; Rollison, 1993; Kim, 1999). These non-zero ratio correlations have been coined 'spurious correlations' by Pearson (1896). The null correlation for the case of a plot of Y/Z against X/Z can be calculated by (Kim, 1999):

$$r_{(X/Y)(Y/Z)} = \frac{V_{1/Z}^2}{\sqrt{[V_X^2(1+V_{1/Z}^2)+V_{1/Z}^2][V_Y^2(1+V_{1/Z}^2)+V_{1/Z}^2]}} \quad (1)$$

where V is the coefficient of variation (= standard deviation divided by the mean). For the sake of completeness it has to be mentioned that plots of two or three variables with common numerator or denominator all display spurious results. The effect of creating these spurious correlations is greatest when the V value of the common divisor is greater than the V value of the other variables. With regard to the closed sum problem in chemical data sets, Chayes (1971) argued that the use of ratios may eliminate the initial closure effect, but that the means, variances and co-variances of the original data are in fact still subject to this effect. So, the spurious correlations are also an expression of the closure effect.

In this communication, characterized by reasoning by means of contrived but realistic scenarios, I will first give some statistical examples of the unintentional effects of normalization on artificial profiles as well as on ratio correlations. The discussion will be continued on the basis of geochemical examples of how one easily can be misled by ratio correlations between elements. Next, I will discuss problems with comparisons of normalized trace elements in marine sediment and in standard shale. The intention of the communication is to revive the awareness of the pitfalls in using element ratios for interpretation of geo-

[1] Other ways of data manipulation, frequently used in marine sedimentary data analyses, imply correcting for contributions from included pore water salt and from the carbonate and organic matter content; in essence, such corrections have effects similar to Al normalization, because the bulk trace element content is put into the basket of the detrital phases.

chemical data. Finally and briefly, some reflections regarding good practice of normalization are discussed.

2. Statistics of normalization and their inferences

Four scenarios will be worked out and discussed. The first scenario considers the case of three variables with zero correlations among them. The second scenario is based on a positive, the third on a negative correlation between two variables, but in both cases no correlation exists with the third variable that is used as a common divisor. In the last scenario all three variables are correlated. The 23 values for each variable are chosen in such a manner that they have a normal distribution between the indicated ranges of their values. To make the scenarios geochemically more appealing, let variables *A* and *B* be trace elements and let *Al* serve as the common divisor. The data are listed in Table 1 and the results for the four worked out scenarios are shown in Fig. 1.

Before discussing the results, it needs to be stressed that the order in which the values for the variables are chosen determines the coefficients of correlation among the variables and between the normalized variables – with the exception of scenario 1 for which the spurious correlation can be predicted by Eq. 1. The same holds for the shapes of the regular and normalized profiles. Therefore, Fig. 1 represents numerical examples of the effects of normalization based on a couple of chosen data sets; using different data sets within the given boundary conditions, one can obtain different results. Scenario 1 offers a clear illustration of the spurious correlation generated by normalization of a data set with three uncorrelated variables. Presented in a profile fashion, the uncorrelated single concentrations are converted into highly correlated normalized concentrations. Scenario 2 is an example of a high correlation between two single elements, which becomes even better when normalized by the third uncorrelated element. The high *element/Al* values also appear as peaks in their profiles and might be easily interpreted as sedimentological events. In scenario 3, the uncorrelated Al converts the negative correlation between B and A into a positive correlation after normalization of A and B. It deserves particular notice that the reduced major axis (r.ax.) lines, that are based on minimizing the areas between data points and the best fit line (Rollison, 1993), have slightly different slopes and intercepts. In scenario 4, with three correlated variables, the normalized concentrations have lost all correlation between them. Plotted in a profile manner, again two peaks emerge in the *A/Al* values suggesting events for which no support can be found in the single element plots. It has to be noted that the quite high correlations of normalized variables in scenarios 2 and 3 are for a large part due to two or three out of 23 values. Going by the scatter plots, one is tempted to consider these values as outliers, the removal of which making the correlation between the remaining data considerably lower and probably insignificant. The result would be that one ignores the (positive and negative) correlations between the untreated values.

Table 1
Statistics of the four scenarios, that are visualized in Figs. 1 and 2

	Scenario 1		Scenario 2		Scenario 3		Scenario 4	
	R	*V*	*R*	*V*	*R*	*V*	*R*	*V*
A		0.13		0.12		0.12		0.29
B		0.32		0.32		0.28		0.11
Al		0.36		0.33		0.33		0.29
B vs *A*	0.05		0.82		−0.92		0.86	
A vs *Al*	0.03		0.05		0.06		0.94	
B vs *Al*	0.04		0.12		0.04		0.89	

R represents the coefficients of correlation between the variables, *V* the coefficients of variation (see text). The ranges of the values of the variables are: 15–25 (A), 3–13 (B) and 15–25 (Al); the full data set is given in Appendix 1.

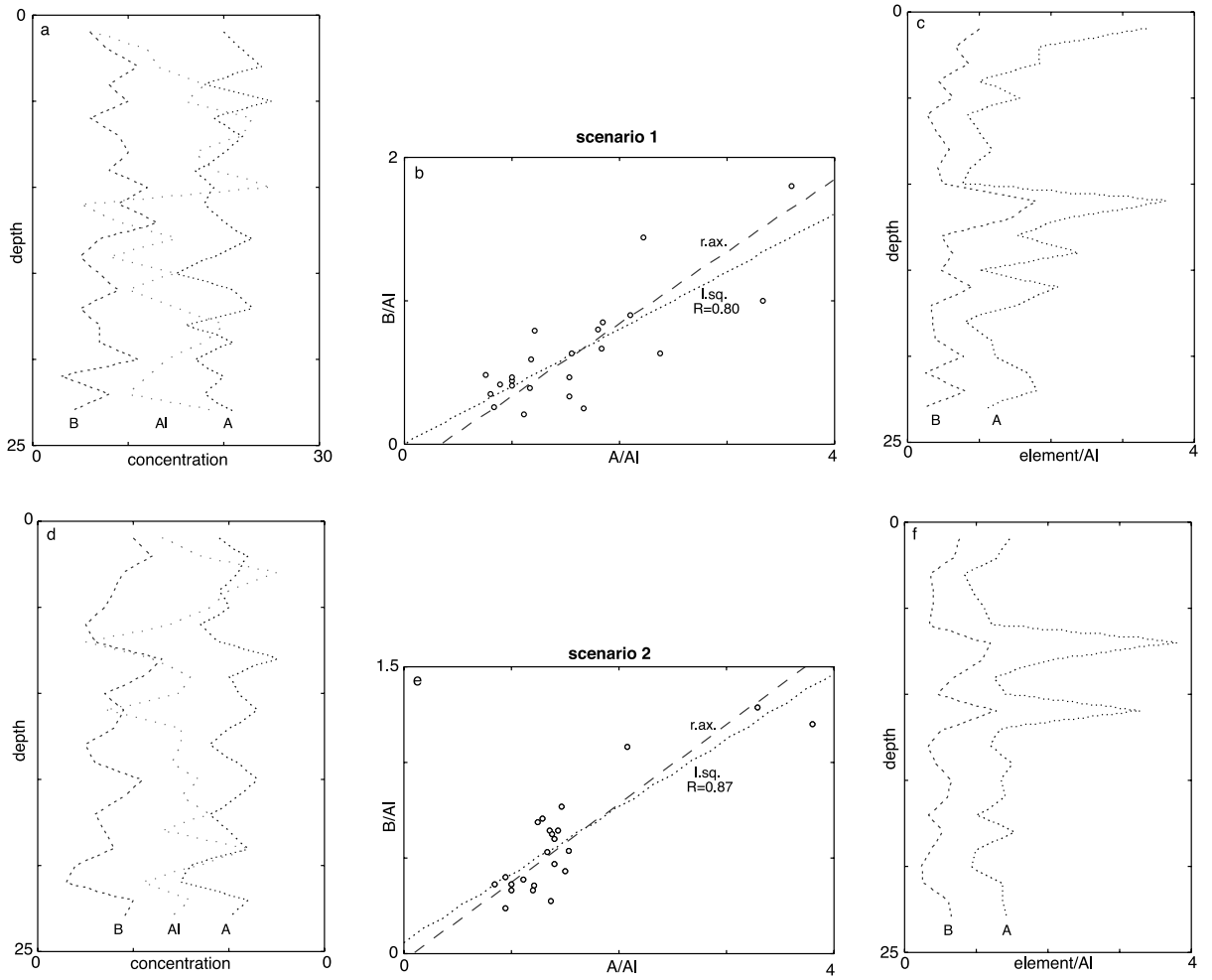

Fig. 1. Results of four scenarios, the basic statistics of which are given in Table 1. The top row represents scenario 1, in which three uncorrelated elements (A, B and Al) are represented. First, all three variables are plotted in a profile fashion in panel a, then A and B as normalized variables in the scatter plot in panel b and as a profile in panel c. The correlation coefficient (R) pertaining to the least-square fit (l.sq.) equals R calculated by Eq. 1. Panels d–f show similar plots for scenario 2, in which A and B are positively correlated, while Al is not correlated with either one of them. Panels g–i show the results for scenario 3, which only differs from scenario 2 by a negative correlation between A and B. The results for scenario 4, with positive correlations among A, B and Al, are displayed in panels j–l. In addition to the least-square fits of y on x, reduced major axis regression fits (r.ax.; Rollison, 1993) are shown in panels b, e and h.

On the basis of the results of this statistical section, I conclude that prudence is in order when interpreting normalized data sets. Using a common divisor with a high coefficient of variation, not only can uncorrelated variables be made to look like highly correlated, but also this practice may enhance, distort or even destroy the correlations existing before the normalization[2]. In addition, peaks in profiles of normalized variables may be interpreted as expressions of sedimentological events for which there is no solid other evidence. When one uses a variable with a low coefficient of variation as common divisor, these effects will be much smaller or not occur at all, but in such a case there is no reason for normalization in the first place.

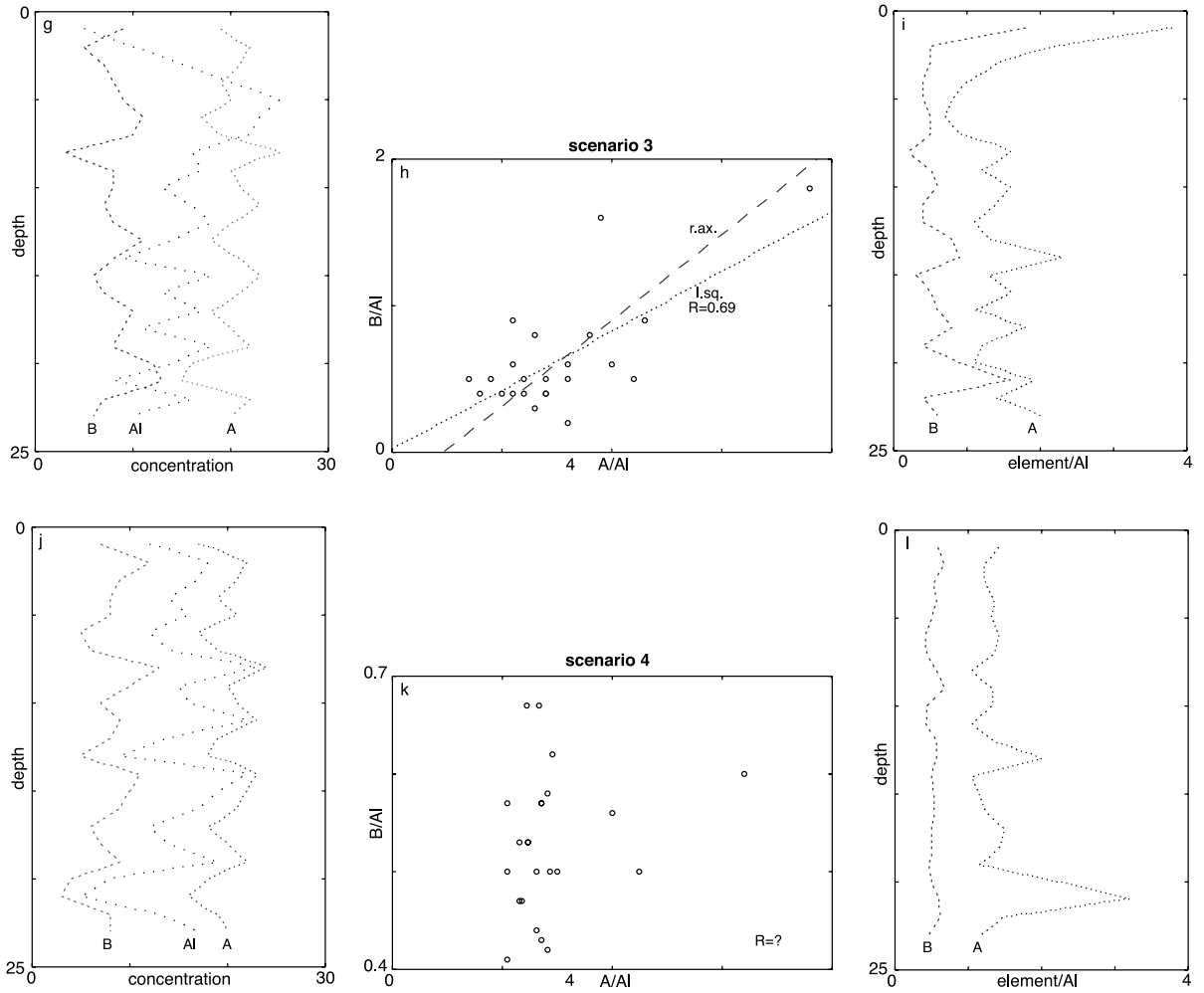

Fig. 1 (*Continued*).

3. Geochemical perspectives

Following up the previous section, an example is given in which a correlation is shown between normalized values of *As* and *S* in a shale, suggesting that As is associated with pyrite. The example is extracted from published work (Brumsack, 1991) and has been chosen for instructive purposes only. The paper concerned did not contain the full data set, only the concentration ranges of Al, S and As. Within these ranges, I made up a data set of uncorrelated values of these variables. The statistics of this data set are given in Table 2.

Fig. 2 shows that the spurious correlation ($R = 0.72$ for least-square regression) between the normalized artificial values comes close to the cor-

[2] By the same token, conversion of element concentrations into element accumulation rates can have similar effects. Sedimentation rate is used as one of the multipliers in this case, so the accumulation rates of A and B will contain a common numerator. Spurious, distorted or destroyed correlations between elements A and B, when plotted as their accumulation rates, can be obtained when, in a sediment profile, variations in the sedimentation rate are large, and the concomitant coefficient of variation is large relative to the coefficients of *A* and *B*. A very instructive example is discussed in Middelburg et al. (1997).

relation ($R = 0.82$) of the measured data. As demonstrated in the previous section, the correlation between the normalized measured data is probably enhanced in comparison with the correlation between unmodified *As* and *S*. The result of reduced major axis regression almost coincides with the published linear regression but differs from the least-square regression (intercepts: 1.11, 1.25, 2.84 ppm/%, and slopes: 9.6, 9.4, 6.9 ppm/% respectively). This means that the best fit through the artificial data is indistinguishable from the fit through the measured data. From a geochemical point of view a relation between *As* and *S* in a shale is very likely, but a scatter plot of their normalized values cannot serve as a proof of this association if there is no correlation between the unamended values.

Next, a number of scenarios will be discussed based on variations in the composition of the sediments caused by different fractions of the constitutive phases clay (cl), amorphous and crystalline silica (si), carbonate (ca), organic matter (OM) and hydrogenous phases (hy). The trace element content of the total sediment can be calculated by summation of the trace element content of each phase multiplied by the fraction of that phase in the bulk. It is common practice to use the Al content of the bulk sediment to correct for dilution by other phases. The mass balance

Table 2
Statistics of scenario 5 with *Al*, *S* and *As* as uncorrelated variables

	Scenario 5 (Posidonia)		
	Range	*R*	*V*
Al (%)	2.6–6.6		0.25
S (%)	1.6–3.6		0.16
Carbonate (%)	15–56		0.47
TOC (%)	1.6–11.8		0.75
As (ppm)	20–42		0.16
S vs *Al*		0.01	
As vs *Al*		0.05	
As vs *S*		0.02	

R and *V* have the same meaning as in Table 1. The full, randomized data set is given in Appendix 2. The scenario is based on a summary of data on the Posidonia shale (from Brumsack, 1991). Spurious correlations between the normalized values of As and S are shown in Fig. 2.

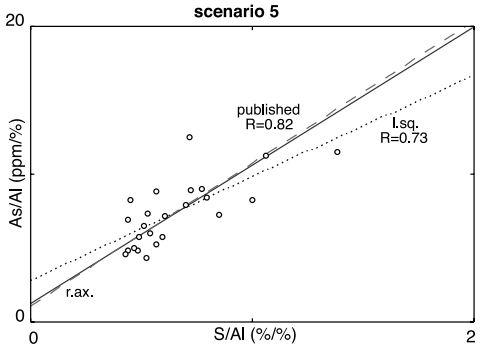

Fig. 2. Scenario 5 generates a spurious correlation between *As*/*Al* and *S*/*Al*, based on a data set (Appendix 2) of uncorrelated Al, S and As concentrations (l.sq. fit, with $R = 0.72$). The statistics of the data are given in Table 2. The linear fit based on reduced major axis regression (r.ax.) almost coincides with the fit published for the original data ($R = 0.82$). Values of slope and intercept are given in Table 6.

equation for the bulk trace element content is:

$$A_{\mathrm{tot}} = aA_{\mathrm{cl}} + bA_{\mathrm{si}} + cA_{\mathrm{ca}} + dA_{\mathrm{OM}} + eA_{\mathrm{hy}} \qquad (2a)$$

or, in a normalized form:

$$\left(\frac{A}{Al}\right)_{\mathrm{tot}} = \frac{aA_{\mathrm{cl}} + bA_{\mathrm{si}} + cA_{\mathrm{ca}} + dA_{\mathrm{OM}} + eA_{\mathrm{hy}}}{Al_{\mathrm{tot}}} \qquad (2b)$$

where *A* is the trace element content (usually in ppm), *Al* is the Al content (commonly in %), and *a*–*e* are the fractions of the phases which make up the bulk sediment (tot).

To keep the scenarios simple, it is assumed that each phase has a fixed trace element content – this is of course seldom the case in the real world, but taking variability into account would make the discussion unnecessarily complicated. Another simplification is that the possible presence of quartz and opal (phases with usually negligible trace element contents but decreasing the total trace element content by dilution) is ignored. It is important to keep in mind that the plots to be shown on the basis of the scenarios appear to a beholder who has no a priori knowledge of the recipe. So, geochemical interpretations have to be made on the basis of apparent correlations between variables[3].

Table 3
Summary of six scenarios with recipes to compose sediments by mixing major (clay and carbonate) and minor (organic matter and hydrogenous phases such as metal oxides or sulfides) sedimentary constituents

A (ppm)	*Al* (%)	Phase	Ranges in % Scenario					
			6	7	8	9	10	11
100	15	clay	90–100	20–90	20–90	19–90	49–75	20–90
0		carbonate		0–75				
50		carbonate			0–75	0–80	20–50	0–75
1 000		organic	0–10	5–10	5–10	1–10	1–5	4–8
10 000		authigenic	0	0	0	0	0	1–2

Aluminum (Al) is only present in a fixed amount in clay, while all constituents have a fixed trace element (A) content. Good to perfect correlations among *Al*, *OM* and *A* exist in all scenarios. The full data set is given in Appendix 3. The correlations between *A*/*Al* and *OM*/*Al* are shown in Fig. 3 and the linear regression results are given in Table 4.

The recipes of the scenarios are given in Table 3. The problem to be solved is defined as the search for a relation between a trace element and organic matter. Eq. 2b can then be recast into:

$$\left(\frac{A}{Al}\right)_{\mathrm{tot}} = \left(a \times \frac{A_{\mathrm{cl}}}{Al_{\mathrm{tot}}} + \frac{cA_{\mathrm{ca}} + eA_{\mathrm{hy}}}{Al_{\mathrm{tot}}}\right) + 0.01 \times A_{\mathrm{OM}} \times \frac{OM}{Al_{\mathrm{tot}}} \quad (3)$$

where *OM* is the organic matter content expressed in %.

When considered as an equivalent of the mathematical description of a straight line by the equation $y = px+q$, a plot of A/Al_{tot} against OM/Al_{tot} seems, at first sight, to be usable for estimation of the trace element content which is contributed by organic matter (from the slope) and for estimation of the contributions of the other phases (from the intercept). Application of this formulation to the scenarios summarized in Table 3 results in scatter plots shown in Fig. 3. Regressions of *y* on *x* are used, because it is reasonable to consider *A* as the dependent (*y*-axis) and *OM* as the independent variable (*x*-axis). In addition, reduced major axis regression lines are shown. With the exception of scenario 10, all fits have high correlation coefficients (*R*) and there is not much difference between the least-square and reduced major axis regression lines. Values of the slopes and intercepts, the inferred values of the trace element contents in clay ($A_{\text{'cl'}}$) and organic matter ($A_{\text{'OM'}}$), and the trace element contents (A_{cl} and A_{OM}) according to the recipes are summarized in Table 4.

Scenario 6 shows the case of trace element A associated only with the clay fraction with the trivial result of a zero slope and with the intercept $a \times (A/Al) = 6.67$ ppm/%, where $a = Al/Al_{\mathrm{cl}}$. Since $Al_{\mathrm{cl}} = 15\%$, $A_{\mathrm{cl}} = 6.67 \times 15 = 100$ ppm, not surprisingly equal to the value assumed in this scenario. In scenario 7 (with carbonate, clay and organic matter as constituents), the slope corresponds with $A_{\text{'OM'}} = 969$ ppm (least square regression) or 1000 ppm (r.ax. regression), and the intercept with $A_{\text{'cl'}} = 105$ ppm (l.sq.) or 101 (r.ax.) ppm, values that compare very well with the recipe values of 100 ppm. It has to be noted that this agreement hinges on the fact that the recipe suggests that $a \rightarrow 1$ when $\mathrm{OM} \rightarrow 0$. Things go wrong in scenario 8, where A is also present in the carbonate fraction. From the equation describing the linear fit in this scatter plot one derives $A_{\text{'OM'}} = 2322$ ppm (l.sq.) or 2347 ppm (r.ax.), compared to the recipe value of 1000 ppm, and a negative value for $A_{\text{'cl'}}$. Scenarios 9 and 10 are other examples of the strange things that can happen with A present in all three sedimentary phases. The scatter plot for scenario 9 makes no sense at all (negative $A_{\text{'OM'}}$), whereas the (poor) fit in scenario 10 gives

[3] In reality, sound interpretations of such correlations are only possible on the basis of geochemical principles.

$A_{`OM'}$ = 150 ppm (l.sq.) or 207 ppm (r.ax.) compared to 1000 ppm (recipe), while $A_{`cl'}$ = 168 ppm (l.sq.) or 165 ppm (r.ax.) compared to 100 ppm (recipe). Small wonder that scenario 11, with four phases containing A, leads to erroneous results: $A_{`OM'}$ = 5330 ppm (l.sq.) or 5404 ppm (r.ax.) instead of 1000 ppm, and $A_{`cl'}$ = negative.

The premise that the relations between normalized values can be treated like a mathematical function describing a linear relation does not hold. First, because the use of *Al* as the common divisor attributes the trace element content of all phases to only the clay. Second, the closed sum ($a+c+d+e \equiv 1$) brings about that the slope in plots of scenarios 7–11 depends on the values of *d* as well as on *a*, *c* and *e*, while the intercept depends on the values of *a*, *c* and *e* and thus also on *d*. For this reason and apart from the statistical problems with regard to correlations introduced by normalization of elements, correlations between sediment components cannot be used in the same manner as are mathematical correlations[4]. One can play numerous games with scenarios similar to the ones presented here, and get different and all erroneous results.

It has to be realized that A_{cl} was calculated using a fixed value of Al_{cl}, whereas in reality there will be a range of *Al* values in clay assemblages or – in a broader sense – in detrital components of a sediment. Another point to make is that extrapolation of the linear fits to $OM/Al \rightarrow 0$ will seldom be justified, because sediments generally do contain organic matter, the value of which is modified by the actual clay content when plotted as a ratio. In cases in which Al_{tot} is high and *OM* low, extrapolation may not prove too bad, but in other cases a zero *x* value is not realistic. A final remark to be made is that the range of *Al* in scenarios 7–11 is, for illustrative purposes, exceptionally large in comparison with ranges in real-world sedimentary situations, so the correction by Al normalization is probably stretched beyond reasonable limits.

It can be objected that most authors, when using *Al* normalization, circumvent the problem of spurious correlation by using scatter plots of *A*/*Al* versus C_{org}. Such plots are useful to demonstrate the contribution of OM to *A*/*Al* in the sediment in excess of *A*/*Al* in reference shale (see last section of this paper). A summary of the results of regression for scenarios 6–11 is given in Table 5. The best fit equations show that the linear fits have negative slopes for scenarios 7 and 8, a positive slope – with a large scatter of the data points, manifest by a poor coefficient of determination – for scenario 10, whereas the data points of scenarios 9 and 11 can only be fitted by a power function. Counterintuitively, four out of the five scenarios show a decrease of *A*/*Al* with increasing *OM*. The use of intercept values of the linear fits to estimate A_{cl} gives worse results than those summarized in Table 4 – even for scenario 7. Admittedly, these results depend on the scenarios used and cannot serve to draw general conclusions. The examples demonstrate, however, that plotting *A*/*Al* versus *OM* may cause problems for geochemical interpretation as well.

4. Geochemical examples

Next, I will discuss three geochemical examples in which ratio correlations are used.

4.1. Posidonia shale

Returning to scenario 5 (Fig. 2, Tables 2 and 6), in which sulfur can be considered to represent an authigenic phase (pyrite), the actual *S* was not lower than 1.6%, so extrapolation of $S \rightarrow 0\%$ stretches this scenario beyond reality. One would obtain $As_{pyrite} = (64/120) \times 960 = 512$ ppm from the slope obtained for reduced major axis regression of the artificial data, about equal to 500 ppm based on the measured data. The concomitant intercept value in scenario 5 yields (with $Al_{detritus} = 12.2 \pm 5.3\%$) $As_{detritus} = 13 \pm 8$ ppm, to be compared with 10 ppm (WSA; Wedepohl, 1971) or 28 ppm (NASC; Gromet et al., 1984)

[4] The relation $y = px+q$ represents the relation between an independent variable (*x*) and a variable *y* dependent on *x*. But making *x* and *y* mutually dependent, as is the case by dividing them by a common denominator, this mathematical condition collapses.

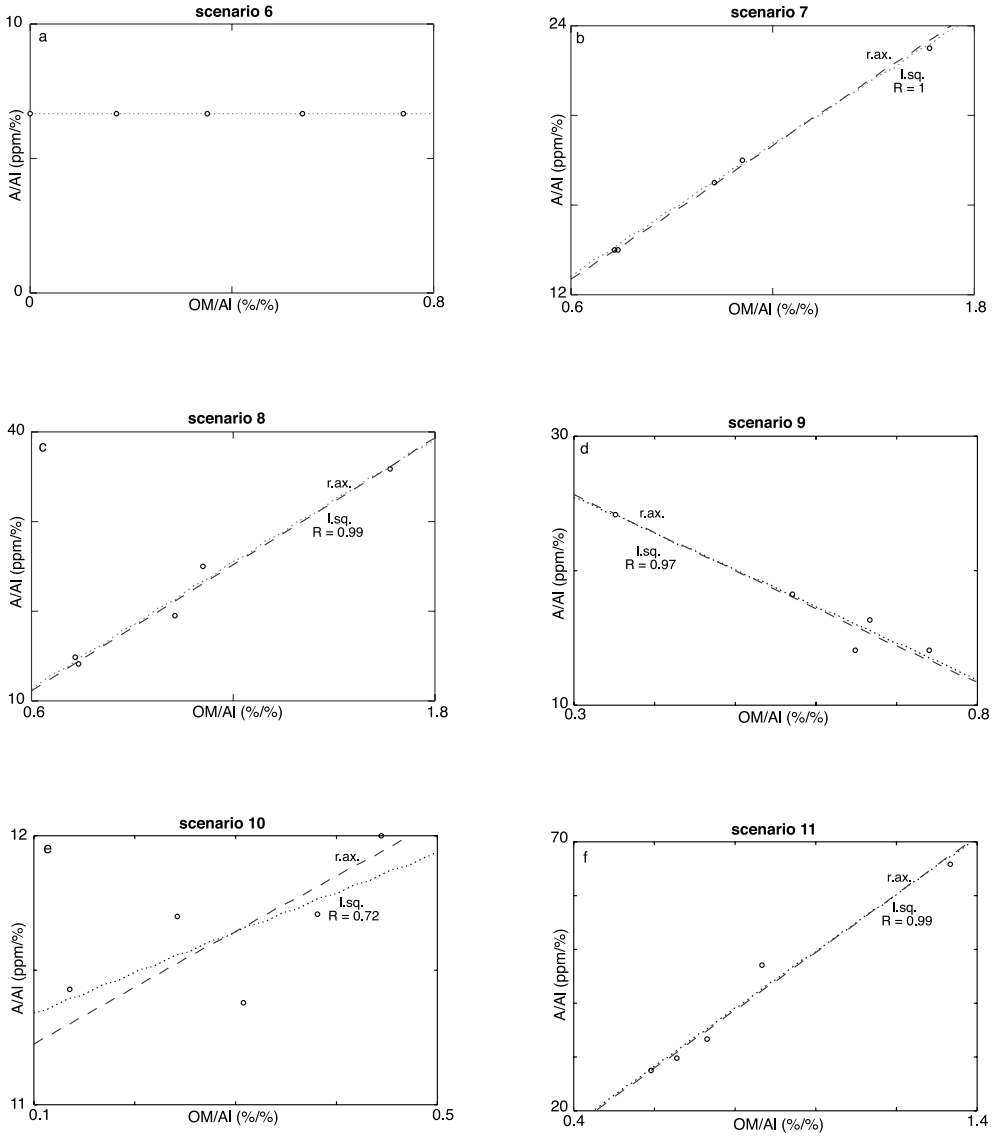

Fig. 3. Graphical results of six scenarios, the basic recipes of which are given in Table 3, the full recipes in Appendix 3. Scenario 6 represents the relation between the normalized values of trace element A (only present in clay) and organic matter. Scenario 7 visualizes this relation for sediment consisting of carbonate, clay and organic matter, with A only present in the latter two phases. Scenarios 8–10 differ from scenario 7 in having A present in all three phases. Scenario 11 depicts the relation between the normalized variables for a case in which – in addition – a hydrogenous phase, hosting A, is present. The values of slopes and intercepts (in ppm/%) pertaining to least-square (l.sq.) and reduced major axis (r.ax.) fits as well as the trace element contents in clay and organic matter calculated from them are summarized in Table 4.

in standard shales (ss). Avoiding the pitfall inherent to the use of *Al* as a common divisor, the author used average values of *Al* (4.6%), *As* (28 ppm) and *S* (2.63%), and (after correction of *As* by adopting $As_{ss} = 5$ ppm) he derived $As_{pyrite} = 420$ ppm. For the subset of average measured data used in scenario 5 and the just derived range of $As_{detritus}$, I calculated $As_{pyrite} = 470$ ppm (200–530 ppm). The two different lines of reasoning (one based on slope and intercept for the fit of the

Table 4
Summary of the results obtained by linear regression applied to scenarios 6–11 (cf. Table 3 and Appendix 3)

Scenario →	6	7	8	9	10	11
Slope (l.sq.)	0	9.69	23.22	−27.1	1.5	−3.94
Slope (r.ax.)	0	10	23.47	−28.0	2.07	−4.53
Intercept (l.sq.)	6.67	6.97	−2.65	33.5	11.19	53.3
Intercept (r.ax.)	6.67	6.72	−2.91	34.1	11.02	54.04
$A_{\cdot cl'}$ (ppm)	100	105	?	503	168	800
$A_{\cdot cl'}$ (ppm)	100	101	?	511	165	811
A_{cl} (ppm)	100	100	100	100	100	100
$A_{\cdot OM'}$ (l.sq.) (ppm)	0	969	2322	?	150	5330
$A_{\cdot OM'}$ (r.ax.) (ppm)	0	1000	2347	?	207	5404
A_{OM} (recipe) (ppm)	0	1000	1000	1000	1000	1000

Standard least-square regression of y on x is indicated by l.sq., reduced major axis regression by r.ax. The trace metal contents (in ppm/%) of organic matter ($A_{\cdot OM'}$) are derived from the slopes, those of clay ($A_{\cdot cl'}$) from the intercepts. The single quotation marks indicate that in these calculated A values contributions from phases other than clay and OM are ignored. For comparison, A_{cl} and A_{OM} used in the recipes are given.

normalized data, the other on average values of the measured variables) do yield similar but not identical results. First, it has to be realized that spurious correlation could hardly be distinguished from the published correlation. Second, values of the unamended variables were not published, so it is not possible to check whether *As* and *S* were correlated. But even then, normalization will have affected the correlation between *As* and *S* as well as the values derived for intercept and slope. Third, with regard to the calculation based on the averaged concentrations, it has to be kept in mind that the result highly depends on the choice of the reference value for As_{ss}. For these reasons, the differences in the results may be partly but not fully attributed to normalization.

4.2. Black Sea

Scenario 12 is based on a high quality data set for OM-rich and OM-poor sediment collected in the Black Sea (Brumsack, 1989). Relevant chemical and statistical information is summarized in Table 6. Scatter plots are shown in Fig. 4 and numerical regression results are presented in Table 7. Can we distinguish between Cu associated with OM (in this case represented by total organic carbon, TOC) and with pyrite (represented by S)? It will be difficult to answer this question, because these two equally likely carrier phases are highly correlated in the sapropels. The high correlation between *Cu* and *Al* in OM-poor sediment is suggestive of an association of Cu with detritus, not with OM or sulfides that in turn show no correlation with *Al*. The other statistical parameters do not look too promising for obtaining clear-cut correlations – with the exception of the one for *Al* in OM-poor samples, all *V* values are high. It has to be noted that the values for the variables are not normally distributed, so a direct compar-

Table 5
Summary of results of regression analysis on scatter plots of *A*/*Al* vs *OM* according to scenarios 7–11 (Table 3)

Scenario ↓	Linear regression (l.sq.)	Linear regression (r.ax.)	Non-linear regression
7	$y = 27.2-1.39x$ ($R = 0.82$)	$y = 29.9-1.76x$	
8	$y = 47.5-3.56x$ ($R = 0.86$)	$y = 51.6-4.13x$	
9			$y = 24.1x^{-0.25}$ ($R = 0.99$)
10	$y = 11.3+0.12x$ ($R = 0.73$)	$y = 11.2+0.16x$	
11			$y = 234x^{-1.07}$ ($R = 0.88$)

L.sq. and r.ax. have the same meaning as in Table 4. For scenarios 9 and 11 a power function fit rather than a linear fit describes the data best.

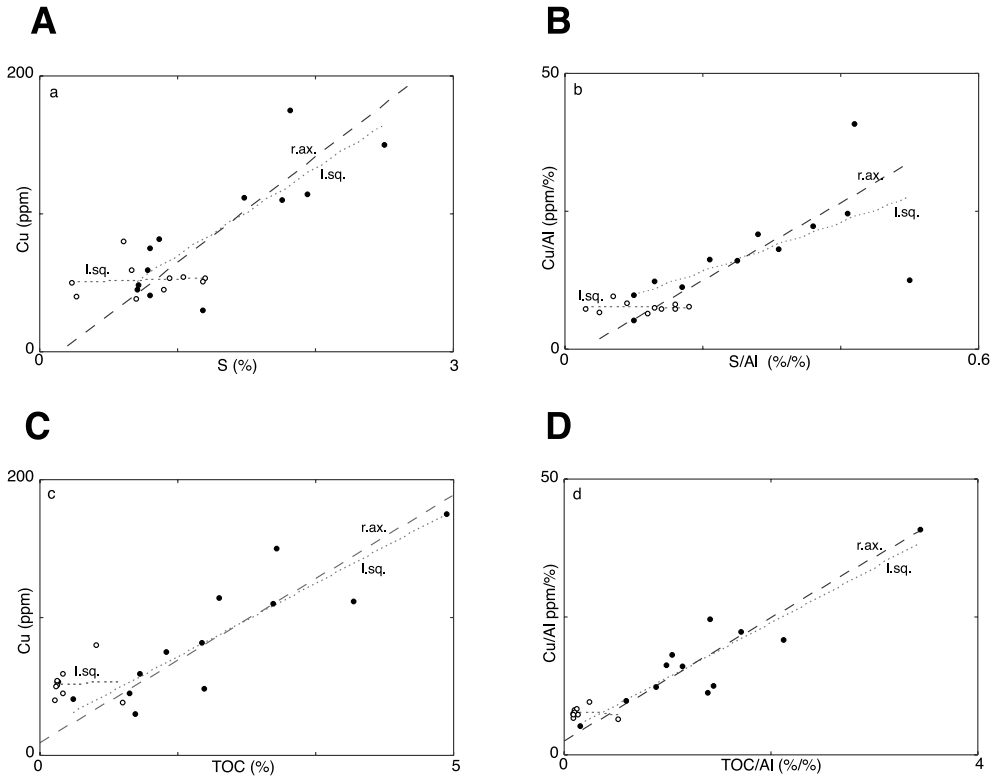

Fig. 4. Scatter plots of unmodified and normalized concentrations of copper (*Cu*) against organic carbon (*TOC*) and sulfide (*S*) respectively. Two sets of data on the composition of sediment samples collected in the Black Sea (Brumsack, 1989) are used: one for sapropels, the other for OM-poor sediment. The statistics for these data sets are summarized in Table 7. Black dots represent sapropels, open symbols OM-poor samples. Values of slopes and intercepts (in ppm/%) of linear fits obtained by least-square (l.sq.) and reduced major axis (r.ax.) regression are summarized in Table 8.

ison with the results discussed in the section on statistical aspects is not warranted.

Fig. 4a,b shows that, ignoring two obvious outliers, the correlation between *Cu* and *S* in the sapropels is high. At first glance, this linear fit may even include the OM-poor samples, but for them a high correlation between *Cu* and *Al* exists (Table 4). Linear fits with even better correlation coefficients seem to encompass the sapropels and OM-poor samples (Fig. 4c,d). However, one has to realize that the *V* value for *Al* in the sapropels is quite high and higher by a factor 3 than in the case of the OM-poor samples (Table 6). Although the *V* values for the other variables are even higher, it still means that *Al* normalization affects the ratio correlations (Fig. 4d) more for the sapropels than for OM-poor samples. It remains uncontested, however, that the *V* values for *TOC* and *Cu* are governing the correlation coefficients.

The scatter plots in Fig. 4a,b indicate that Cu is strongly associated with sulfides in the sapropels. The near-zero correlation between *Cu* and *S*, either unamended or normalized, indicates that sulfides in the OM-poor sediment do not host Cu. When $S \rightarrow 0$ in the OM-poor sediments, *TOC* is small but still positive (about 0.4–0.5%, estimated from scatter plots of *S* vs *TOC*).

Ignoring a contribution of Cu in OM, the intercept values for the OM-poor sediment, $a \times Cu_{cl'} = 50.3$ ppm and $a \times Cu_{cl'}/Al = 7.7$ ppm/% (Table 7), are attributed to the detrital fraction, which is supported by the high correlation between *Cu* and *Al* (Table 6). Similarly, extrapola-

Table 6
Summary of results of regression analysis on the scatter plot of *As*/*Al* versus *S*/*Al* (Fig. 2)

	Linear regression		
		Slope	Intercept
Artificial data set (Appendix 2)	l.sq.	6.94	2.84
	r.ax.	9.60	1.11
Measured data	l.sq.	9.38	1.25

Slope and intercept values are in ppm/%. L.sq. and r.ax. have the same meaning as in Table 4. The artificial data set is given in Appendix 2; the regression line for the measured data is estimated from the scatter plot in Brumsack (1991).

tion of $TOC \rightarrow 0$ (Fig. 4c,d) gives intercept values $a \times Cu_{\text{`cl'}} = 51.2$ ppm and $a \times Cu_{\text{`cl'}}/Al = 7.8$ ppm/% (Table 7). It can be calculated (cf. Table 4) that $a = 0.80 \pm 0.03$ and $Al_{\text{detritus}} = 8.4 \pm 0.3$, from which it follows that in the OM-poor sediment $Cu_{\text{`cl'}} = 63 \pm 2$ ppm (Fig. 4a,c) and $Cu_{\text{`cl'}} = 65 \pm 2$ ppm (Fig. 4b,d). These values are equal, but significantly higher than Cu_{ss} (28 ppm, Brumsack, 1989, or 35 ppm, Piper and Isaacs, 1995a). Slopes and correlation coefficients (Table 7) do not allow the estimation of Cu_{OM} and Cu_{pyrite} for the OM-poor sediment.

When using the results of least-square regression in Fig. 4a,b and of reduced major axis regression in Fig. 4c,d for the plots of sapropel data (Table 7), in combination with data in Table 6 (from which can be estimated: $a = 0.51 \pm 0.07$, $Al_{\text{detritus}} = 10.2 \pm 1.4$), the following results can be obtained for Cu_{detritus} (in ppm): 12 ± 2 (Fig. 4a), 57 ± 8 (Fig. 4b), 15 ± 3 (Fig. 4c), 21 ± 4 (Fig. 4d). The first two values differ significantly, owing to a poor least-square fit of the data in comparison with reduced major axis regression. The last two values are overlapping. In any case, it seems that Cu_{detritus} is lower in the OM-rich than in the OM-poor sediment. For the sapropels, Cu_{OM} can be estimated from the slopes (Table 7) for the reduced major axis regression and assuming $TOC/OM = 0.36$ (Redfield ratio). The results are: $Cu_{\text{OM}} = 433$ ppm (Fig. 4c) or $Cu_{\text{OM}} = 399$ ppm (Fig. 4d). Visually, the linear best fits seem also to represent the data on the OM-poor sediment. Similarly, assuming that S represents the pyrite content, it can be derived from the slopes (Table 7) that $Cu_{\text{pyrite}} = 4080$ ppm (Fig. 4a) or $Cu_{\text{pyrite}} = 3770$ ppm (Fig. 4b).

The relation between normalized total *Cu* and *Cu* contributed by the most plausible carriers detritus, OM and pyrite (py), is given by a reduced form of Eq. 2b:

$$\left(\frac{Cu}{Al}\right) = \frac{aCu_{\text{detritus}} + dCu_{\text{OM}} + eCu_{\text{py}}}{Al} \tag{4}$$

To substitute the relations among the normalized values of *Cu*, *TOC* and *S* shown in Fig. 4b,d, Eq.

Table 7
Summary of analytical and statistical data for sediment samples collected in the Black Sea; two types of sediment are distinguished: sapropels and OM-poor sediments

	Scenario 12 (Black Sea)					
	Sapropels			OM-poor sediment		
		R	*V*		*R*	*V*
Al (%)	5.2 (1.8)		0.34	6.7 (0.7)		0.11
Carbonate (%)	32 (15)		0.46	16 (6)		0.35
TOC (%)	6.5 (3.8)		0.59	1.0 (0.8)		0.78
S (%)	1.3 (0.6)		0.47	0.8 (0.3)		0.44
Cu (ppm)	87 (46)		0.53	52 (12)		0.22
TOC vs *Al*		0.04			0.01	
S vs *Al*		0.10			0.09	
Cu vs *Al*		0.22			0.95	
S vs *TOC*		0.65			0.12	

R and *V* have the same meaning as in Table 1. It has to be noted that the values of the variables in the data set (from Brumsack, 1989) are not normally distributed. Scatter plots of original and normalized data are shown in Fig. 4 and the regression results are given in Table 8.

4 can be recast into:

$$\left(\frac{Cu}{Al}\right) = 0.01 \times \left[\left(\frac{100}{Al_{\text{detritus}}}\right) \times Cu_{\text{detritus}} + \left(\frac{OM}{TOC}\right)_{\text{Rr}} \times \frac{TOC}{Al} \times Cu_{\text{OM}} + \left(\frac{^{M}\text{FeS}_2}{2M_{\text{S}}}\right) \times \frac{S}{Al} \times Cu_{\text{py}}\right] \quad (5)$$

where Al, OM, TOC, S are in % and Cu in ppm; Rr is the Redfield ratio (bracketed ratio = 2.8), and M the molecular weight (bracketed ratio = 1.875).

Eq. 5 offers a possibility to test the values derived for Cu_{OM} and Cu_{pyrite} in the sapropels. The following ratios can be calculated from the data set (Table 6)

$$S/Al = 0.25 \pm 0.14\%/\%$$
$$TOC/Al = 1.7 \pm 1.2\%/\%$$
$$Cu/Al = 16.7 \pm 12\ ppm/\%$$

Additionally, the next values can be used: $Al_{\text{detritus}} = 10.2 \pm 1.4\%$ (just derived), $Al = 5.2 \pm 1.8\%$ (Table 6) and $Cu_{\text{detritus}} = 17.5 \pm 7.5$ ppm (just derived). Substitution of these values gives the following equation:

$$16.7 \pm 12 = (18.5 \pm 6.5)/(10.2 \pm 1.4) + (0.048 \pm 0.034) \times (416 \pm 17) + (0.0047 \pm 0.0026) \times (3925 \pm 155) = (1.8 \pm 0.7) + (20 \pm 14.2) + (18 \pm 10) = 40 \pm 17 \quad (6)$$

Numerically, this test does not falsify the results obtained by the graphical methods, but one cannot push aside the impression that the value obtained on the right side of this equation is too high[5]. The logical explanation is that the values obtained for Cu_{OM} and Cu_{pyrite} are too high, the reason for which is that the slopes, from which these values were derived, still contain contributions of Cu_{OM} (Fig. 4a,b) and of Cu_{S} (Fig. 4c,d). Other pathways have to be followed (e.g. microprobe analysis of pyrite) to overcome this problem of mutual dependency.

4.3. Arabian Sea

Scenario 13 uses data on two sediment cores collected in the Arabian Sea, one core (463) from within the oxygen minimum zone (OMZ) with extremely low oxygen concentrations ($[O_2] \leq 2$ μM), the other (464) from underneath the OMZ.

Information on concentrations and statistics is presented in Table 8. The correlations of the original and normalized values of the concentrations of the variables Cu and TOC are shown in Fig. 5. The main differences between the scenarios 13 (Arabian Sea) and 12 (Black Sea) are: (1) in the OMZ of the Arabian Sea – though oxygen concentrations are extremely low – full-fledged sulfidic conditions do not exist like in the deep Black Sea; (2) the ranges of the Al concentrations are smaller and the concomitant V values consequently lower; (3) the data set is much larger; (4) the sedimentary regimes at stations 463 and 464 are – apart from depth and dissolved oxygen levels – very similar, in contrast to the regimes in the shallow and deep parts of the Black Sea.

Fig. 5 shows that the Cu concentration in core 464 (below OMZ) is less dependent on the OM content than in core 463 (within OMZ). Oxic diagenesis of OM in sediment at station 464 may have diminished the association between these variables, but diagenetic enrichment of Cu under the low BWO conditions is a more likely cause of the higher Cu contents in core 463. The numerical results of linear regression pertaining to Fig. 5 are given in Table 9. The correlation coefficients between Cu and TOC are low for both cores, those for 464 being lower than for 463. Although it is tempting to attach most significance to the ratio correlations, it has been demonstrated that this may – for statistical reasons – not be warranted.

[5] The combination of errors in this calculation has been made on the assumption that the errors are unrelated, which they are not. However, the here calculated ranges suffice to make my point.

Table 8
Summary of results of linear regression (Fig. 4) of unamended and *Al*-normalized values of *Cu* versus *S* and *Cu* versus *TOC* for sapropels and OM-poor sediment in the Black Sea

		Least-square regression			Reduced major axis regression	
		Slope	Intercept	*R*	Slope	Intercept
Cu vs *S*	sapropels	63.2	5.9	0.83	76.53	−11.16
	OM-poor	2.6	50.3	0.07		
Cu/*Al* vs *S*/*Al*	sapropels	44.3	5.6	0.64	70.77	−1.57
	OM-poor	−0.57	7.7	0.03		
Cu vs *TOC*	sapropels	10.6	18	0.88	12.02	9.03
	OM-poor	1.1	51.2	0.07		
Cu/*Al* vs *TOC*/*Al*	sapropels	10.1	3.8	0.91	11.08	2.47
	OM-poor	−1.1	7.8	0.17		

The values of *Cu* are in ppm, those of *Al* and *TOC* in % of bulk sediment.

The *V* values of *Al* are slightly higher for samples from core 463 than from core 464 (Table 8), but in both cases the decrease (463) and increase (464) of the correlations of the normalized values compared to the original values of *Cu* may be caused by statistics. It is important to keep in mind that (1) relations between the normalized values are also apparent between the original values, and (2) normalization by *Al*, because of its low *V* values, will hardly change these relations. On the other hand, calculations on the basis of these relations to estimate the Cu content of individual phases will, because of the rather low *R* values, not lead to very accurate results.

Notwithstanding this caveat, an attempt can be made to get semi-quantitative information from these relations. Assuming that when $TOC \rightarrow 0$ the sediment has only two significant components left, i.e. detrital material and carbonate, and considering that Cu_{ca} is negligibly low, the intercept values in both graphs may be used to estimate the copper content of the clay fraction. The coefficients of determination (R^2) for the least-square regressions are low. For that reason, preference is given to the results of reduced major axis regression, although the scatter of data around the best fits remains large. When this is done for Fig. 5a and when the carbonate contents of 'OM-free' sediment are estimated to be $41 \pm 6\%$ (463) and $45 \pm 6\%$ (464) (cf. Table 5), the values of $Cu_{detritus}$ are 23.4 ± 2.4 ppm (463) and 37 ± 4 ppm (464). Combination of the intercept values

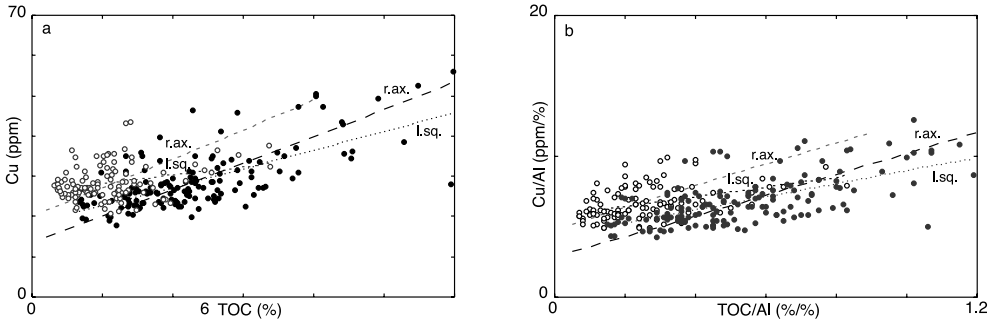

Fig. 5. Scatter plots of unmodified and normalized concentrations of copper (*Cu*) against organic carbon (*TOC*) in two piston cores collected at Murray Ridge in the northern Arabian Sea. Information about the coring stations and the statistics of the chemical data are summarized in Table 9. Black dots represent samples from core 463; open symbols the ones from core 464. Values of slopes and intercepts (in ppm/%) of linear fits obtained by least-square (l.sq.) and reduced major axis (r.ax.) regression are summarized in Table 10.

Table 9
Summary of analytical and statistical data for sediment samples from piston cores collected at Murray Ridge in the northern Arabian Sea

	Scenario 13 (Arabian Sea)					
	Core 463 (within OMZ)			Core 464 (below OMZ)		
		R	*V*		*R*	*V*
Al (%)	4.2 (0.6)		0.14	4.3 (0.5)		0.10
Carbonate (%)	39 (6)		0.11	44 (5)		0.12
TOC (%)	2.3 (1.1)		0.48	1.1 (0.6)		0.55
Cu (ppm)	29 (7)		0.25	29 (5)		0.16
TOC vs *Al*		0.32			−0.25	
Cu vs *Al*		0.20			0.14	

The contrast between these geographically nearby cores lies mainly in the bottom water oxygen concentrations in which the sediments have been deposited. Core 463 contains sediments deposited within a very intensive ($[O_2]$ down to ≤ 2 μM) OMZ, whereas core 464 represents sediments from below the OMZ. Water depth at station 463 (22°33.60′N, 064°03.25′E) is 1 km, at station 464 (22°15.00′N, 063°34.70′E) 1.5 km. Both cores contain two full glacial cycles, spanning records of 205 ka (463) and 225 ka (464). *R* and *V* have the same meaning as in Table 1. Scatter plots of the original and normalized data are shown in Fig. 5.

of the plots of the normalized values (Fig. 5b) with $Al_{\text{detritus}} = 7.1 \pm 0.4\%$ (463) and $7.65 \pm 0.2\%$ (464) gives: $Cu_{\text{detritus}} = 21 \pm 1.2$ ppm (463) and 36.5 ± 1 ppm (464). The Cu_{detritus} values obtained in both ways for the two cores agree very well, but for core 463 this value is lower than and for core 464 equal to $Cu_{\text{ss}} = 31 \pm 4$ ppm (Brumsack, 1989; Piper and Isaacs, 1995a).

With respect to the slopes of the graphs, it first has to be noted that I have no data on sulfide-S in these cores. This means that Cu in sulfidic phases such as pyrite is not separately considered here but is hidden in the contribution attributed to OM (see discussion in previous section). Using the values of the slopes for reduced major axis regression, it follows that Cu_{OM} is equal to 240 ppm (463) and 257 ppm (464) for Fig. 4a, and 263 ppm (463) and 274 ppm (464) for Fig. 4b. The agreement between the values obtained for each core is quite good and for core 464 much better than when the parameters for least-square regression are used.

The conclusion drawn from these results is that the relations between the normalized variables do not greatly differ from those between the unmodified variables. The main reason is that the *V* values of *Al* are low in comparison to those for *TOC* and *Cu*. In essence, normalization does not improve the diagnostic prospects in such cases.

5. Comparison with standard shale values

In the foregoing section, I have already preluded on the use of and comparison with standard shale values of variables. In this short section I will dwell a bit more on this quite popular practice.

Table 10
Summary of results of linear regression (Fig. 5) of unamended and *Al*-normalized values of *Cu* versus *TOC* for sediment collected at Murray Ridge in the Arabian Sea

	Station	Least-square regression			Reduced major axis regression	
		Slope	Intercept	*R*	Slope	Intercept
Cu vs *TOC*	463	4.61	18.3	0.69	6.7	13.6
	464	2.13	26.1	0.30	7.19	20.2
Cu/*Al* vs *TOC*/*Al*	463	4.64	4.33	0.63	7.36	2.9
	464	3.37	5.83	0.45	7.66	4.77

More information is given in the legends of Table 8. Values of *Cu* are in ppm, those of *Al* and *TOC* in % of bulk sediment.

Starting with the mass balance (Eq. 2a), the first term on the right-hand side can be considered to represent the presence of a standard marine shale type mixed phase and its contribution to the total content of A. This can be inserted into Eq. 6 by expressing fraction a as: $a = 0.01 \times (100/Al_{ss}) \times Al_{tot}$, with Al in weight percent. As a consequence, A_{cl} is converted into A_{ss}. Substitution into Eq. 6 gives:

$$\left(\frac{A}{Al}\right)_{tot} = \left(\frac{A}{Al}\right)_{ss} + \frac{bA_{si} + cA_{ca} + dA_{OM} + eA_{hy}}{Al_{tot}} \quad (7)$$

This relation can be used to estimate, for instance, the free silica content of sediment, as follows:

$$SiO_{2free} = SiO_{2tot} - \left(\frac{SiO_2}{Al}\right)_{ss} \times Al_{tot} \quad (8)$$

or, in general:

$$A_{excess} = A_{tot} - \left(\frac{A}{Al}\right)_{ss} \times Al_{tot} \quad (9)$$

Another way of expressing this is by introducing an enrichment factor (EF):

$$EF = \frac{(A/Al)_{tot}}{(A/Al)_{ss}} =$$

$$1 + \frac{(bA_{si} + cA_{ca} + dA_{OM} + eA_{hy})/Al_{tot}}{(A/Al)_{ss}} \quad (10)$$

The second term on the right-hand side comprises all contributions supposedly not yet considered in the standard shale, normalized by the total Al content and relative to the contribution of the normalized content of A in the standard shale[6].

How about pitfalls – other than discussed in the foregoing sections – when one uses one of the last four equations? The first problem lies in the key role of Al_{ss} in all of them. Adoption of a published value may be the wrong choice for a particular sedimentary area or region. To avoid this error, an attempt can be made to analyze the detrital matter carried to the studied deposition area. However, the actual sources of detritus are not necessarily representative for those in the geological past. As examples: the dominant provenance of detrital material in sediments in the Mediterranean and in the northern Arabian Sea has changed in response to variations in the monsoon (Calvert and Fontugne, 2001; Reichart et al., 1997). Consequently, it cannot be expected that this material had a constant Al content. Another example is presented by Calvert et al. (1996), who related the difference in the chemical composition of the terrigene fraction of a shale between laminated and bioturbated intervals to the variation in mineralogical composition. A second problem with the use of the EF is that it is not possible to identify the carrier phase(s) of A. In the previous section I have shown that it is all but impossible to distinguish between the contributions of organic matter and a hydrogenous phase such as pyrite to the trace element content of sediment. It is tempting to link enrichment with diagenesis, for instance when a strong gradient exists between the concentrations of an element in bottom and pore water or along a pore water profile. However, there are examples to show that this may be wrong, e.g. Mn, Cd, or Ba may also be present in carbonates, V in organic matter or in ferric oxides, Ge in silica and in organic matter, etc. It depends on the relative magnitude of the trace element content in these phases how much of a problem it is. The gist of this is that one has to have a good knowledge of the natural behavior of elements in the aquatic environment in order to identify the processes that have led to enrichment relative to the standard shale. In this context, there exists a great need to assess what can be learned from microanalysis of individual solid phases in terms of their trace element content, or from sequential extractions of sediments to identify the carrier phases of an element and to quantify their contribution to the total content of an element.

Returning to the problem of the composition of standard shale, it is worth recalling that standard shale does not consist only of clay minerals. We-

[6] Needless to say the EF relation does not hold if the standard shale is devoid of component A.

depohl (1971) gives the following mineralogical composition: illite and expanded clay minerals 50 ± 5%, quartz 20%, feldspars 12.5 ± 2.5%, chlorite and kaolinite 14%, calcite and dolomite 3%; the OM content is low. So, standard shale already includes phases that are separately considered in the second term of the right-hand side of Eqs. 7 and 10. As a consequence, Al_{ss} is lower than the Al content of a mixture of only clay minerals, for which one would expect a range from 9% (chlorite) to 28% (kaolinite). Reported Al contents in average shales from different sedimentary provinces range from 8.0% to 9.9% (Gromet et al., 1984; Turekian and Wedepohl, 1961; Wedepohl, 1971). Deep-sea clays have an average Al content of 8.4% and fall within this shale range, but in many cases their trace element contents differ significantly from those reported for shales (Turekian and Wedepohl, 1961). There is also a significant difference between the average trace element content of bituminous and normal non-bituminous shale, the latter having lower concentrations of redox-sensitive elements such as V, Cu, Mo and U (Wedepohl, 1971). This means that reference shales may include a diagenetic component.

Summarizing, comparison of the composition of marine sediment and standard shale is not unbiased. Depending on the area or region of interest, the detrital part of the sediment may have an average mineralogical and chemical composition that differs from those published for average shale. It has to be realized that the practice of normalization of element contents by Al_{ss} can introduce errors of up to 20% because of the observed range of Al contents in shales.

6. Good practice

The emphasis on the pitfalls of normalization in the foregoing sections might have left the impression that normalization is always a bad practice. This would discredit many papers in which normalization was used in a well-considered way. Usually, *Al* normalization is applied to correct for dilution and to assess the enrichment of a trace element (A) by non-detrital components. The results are graphically (profiles, scatter plots) or numerically (EF, *A* in biogenic, hydrogenic or authigenic phases) presented. Among many other ones, I can refer to papers by Calvert and Pedersen (1993), Calvert et al. (1996), Piper and Isaacs (1995a,b), Shimmield and Pedersen (1990), Thomson et al. (2001), Von Rad et al. (1999). Similarly, normalization on a carbonate-free basis when the coefficient of variation for the carbonate content is relatively low can be a useful tool in comparison and interpretation of trace element profiles (Thomson et al., 1993).

A solid understanding of trace element geochemistry is a prerequisite for a proper interpretation of trace element patterns in marine sedimentary settings. When *Al* is used as a divisor, *A* is attributed to the aluminosilicate fraction. By comparison of $(A/Al)_{tot}$ with $(A/Al)_{reference}$, it is possible to estimate A_{excess} and the EF. Note the use of 'estimate' instead of 'determine', because (1) an educated choice has to be made for a $(A/Al)_{reference}$ value (e.g. for standard shale, upper crust, deep-sea sediment, provenance dependent regional sediment), (2) dilution by quartz will decrease *Al* and increase $(A/Al)_{tot}$ and, as a consequence, $(A/Al)_{excess}$ attributed to biogenic and hydrogenic phases will be too high.

For that reason, calculation of $Al_{detritus}$ on a carbonate-, OM-, and salt-free basis, taking into account the statistical uncertainties, is recommended for comparison with $Al_{reference}$.

Systematic patterns or trends in *A*/*Al* profiles are often interpreted in terms of changes in sedimentological and/or climatological regimes. These patterns can be the result of changes in $(A/Al)_{detritus}$ and/or by $(A/Al)_{excess}$. Examples are $(Zr/Al)_{detritus}$ or $(Ti/Al)_{detritus}$ as provenance recorders (e.g. Shimmield, 1992; Schenau et al., 1999), $(Ba/Al)_{excess}$ as a proxy for productivity (e.g. Gingele et al., 1999; Schenau et al., 2001), and $(Mo/Al)_{excess}$ and $(U/Al)_{excess}$ (e.g. Morford and Emerson, 1999) as proxies for redox conditions. When used as qualitative indicators, such profiles are helpful for reconstruction of sedimentological conditions. For use in a quantitative way, the comments made in the previous paragraph apply. Furthermore, one has to be aware that, in cases where an element is present in more than one carrier phase, the relative amount con-

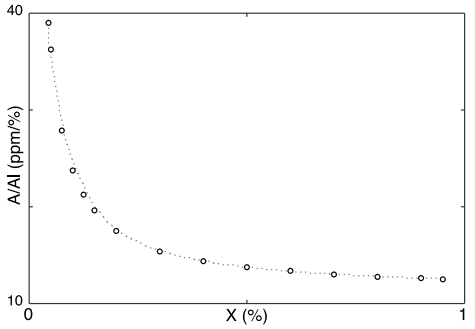

Fig. 6. Plot of scenario with two phases, the first with $Al = 8\%$ and $A = 100$ ppm, the second devoid of Al but with $A = 10$ ppm. X is the fraction of the first phase in the mixture.

tributed by each phase is important. One example is that the difference between about equal values of $(A/Al)_{tot}$ and $(A/Al)_{ss}$ ($\approx (A/Al)_{excess}$) has a high degree of uncertainty. Fig. 6, showing the possible effect of a large compositional range on A/Al, represents another example.

Most of the concerns, however, demonstrated in the foregoing sections are about ratio correlations. The erroneous and/or misleading results have been amply discussed. When two trace elements are present in one phase only, dilution by other phases will affect them to the same degree and their correlation will not change upon dilution. However, when there is only one carrier for the two traces and when detrital and other phases are in relatively different amounts responsible for dilution, Al normalization will change the original correlation. It is advisable to start with inspection of the correlation between elements prior to normalization. Next, one should take into account the V value of Al in comparison with the V values of the elements (or components such as TOC). When the former value is small relative to the values of the other variables, the correlations between the variables will not be seriously affected, in which case it is hardly advantageous to make comparisons on an Al-normalized basis. Caution is recommended when the V value of Al is relatively large, i.e. when the detrital fraction has a large range. In that case, correlations that were present in the original data may be lost through Al normalization. By the same token, interpretations of correlations between variables normalized by other common divisors with a large compositional range should be made with great reserve.

7. Conclusions

It has been shown and discussed that Al normalization of element concentrations in marine sediment:

- can introduce spurious correlations between variables, the more so when the V value of Al is large relative to the V values of the other variables;
- does not circumvent the problem of the closure effect;
- is not an unbiased method for comparison with standard shale compositions;
- is generally a valuable method to estimate the enrichment of an element relative to a reference sediment, but often falls short of identifying and quantifying contributions by sediment components other than the detrital fraction.

Acknowledgements

For their corrections and suggestions that helped to improve this paper, I thank J.J. Middelburg, G.J. Reichart, A. Rutten, and J. Thomson.

Appendix 1

Data sets used for scenarios 1–4 (Fig. 1).

nr	*A*	*B*	*Al*	*A*	*B*	*Al*	*A*	*B*	*Al*	*A*	*B*	*Al*
1	20	6	6	19	10	13	19	9	5	17	7	12
2	22	8	12	22	12	17	22	5	10	22	12	18
3	24	11	13	21	9	25	21	7	15	21	9	17
4	18	8	18	19	8	20	19	8	20	19	8	14

Appendix 1 (*Continued*).

nr	*A*	*B*	*Al*	*A*	*B*	*Al*	*A*	*B*	*Al*	*A*	*B*	*Al*
5	25	10	16	20	7	18	20	9	25	21	8	16
6	19	6	23	17	5	14	17	11	23	17	5	12
7	22	9	22	19	6	5	19	10	22	19	6	14
8	20	10	17	25	13	12	25	3	16	24	13	23
9	17	8	19	20	11	16	20	8	17	20	10	15
10	19	12	25	21	7	15	21	8	13	21	7	16
11	18	9	5	23	9	7	23	7	16	23	9	22
12	20	13	9	20	8	15	20	8	18	19	8	14
13	23	7	15	18	5	15	18	11	14	18	5	9
14	19	5	8	21	6	14	21	8	9	23	11	22
15	15	7	15	23	11	17	23	6	18	22	10	19
16	21	9	10	21	9	15	21	7	13	21	9	17
17	23	5	15	18	6	18	18	10	17	18	6	12
18	16	7	20	20	7	13	20	9	11	20	7	14
19	21	7	18	22	8	22	22	8	18	22	9	19
20	17	11	14	16	4	17	16	12	14	18	4	8
21	20	3	12	15	3	11	15	13	8	16	3	5
22	18	8	10	22	10	16	22	7	16	19	8	13
23	21	4	19	20	9	14	20	6	10	20	8	17

Appendix 2

Randomized concentrations of Al, S and As in Posidonia shale used in scenario 5 (Fig. 2).

Al (%)	*S* (%)	*As* (ppm)
4.6	2.4	20
2.9	2.9	24
5.1	2.3	42
4.8	2.1	33
3.1	2.4	28
5.4	2.6	26
6.6	2.9	32
5.7	2.8	33
6.1	2.6	28
3.4	1.8	25
5.3	3	28
4.3	3	34
4	3.4	29
5	2.7	30
4.3	2.6	31
4.9	2.5	32
2.6	3.6	30
3.9	3.1	33
6.2	2.9	31
3.2	3.4	36
4.7	2.8	27
2.8	2	35
4.4	2.5	39
3.6	2.6	32

Appendix 3

Data used in scenarios 6–11 (Fig. 3); cl, ca, OM, hy, and tot are abbreviations for clay, carbonate, organic matter, hydrogenous phases and bulk sediment; A(X) stands for the trace element content of X.

cl (%)	*Al* (%)	*ca* (%)	*OM* (%)	*hy* (%)	*A*(cl) (ppm)	*A*(ca) (ppm)	*A*(OM) (ppm)	*A*(hy) (ppm)	*A*(tot) (ppm)
Scenario 6									
100	15.00	0	0	0	100	0	0	0	100
97.7	14.63	0	2.5	0	100	0	0	0	97.7
95	14.25	0	5	0	100	0	0	0	95
92.5	13.88	0	7.5	0	100	0	0	0	92.5
90	13.50	0	10	0	100	0	0	0	90
Scenario 7									
90	13.5	0	10	0	100	0	1 000	0	190
73	10.95	19	8	0	100	0	1 000	0	153
52	7.8	40	8	0	100	0	1 000	0	132
30	4.5	65	5	0	100	0	1 000	0	80
20	3	75	5	0	100	0	1 000	0	70
Scenario 8									
90	13.5	0	10	0	100	50	1 000	0	190
73	10.95	19	8	0	100	50	1 000	0	163
52	7.8	40	8	0	100	50	1 000	0	152
30	4.5	65	5	0	100	50	1 000	0	113
20	3	75	5	0	100	50	1 000	0	108
Scenario 9									
90	13.5	0	10		100	50	1 000	0	190
72	10.8	21	7		100	50	1 000	0	153
50	7.5	45	5		100	50	1 000	0	123
35	5.25	62	3		100	50	1 000	0	96
19	2.85	80	1		100	50	1 000	0	69
Scenario 10									
75	11.25	20	5		100	50	1 000	0	135
70	10.5	26	4		100	50	1 000	0	123
65	9.75	31	4.5		100	50	1 000	0	125
55	8.25	42	3		100	50	1 000	0	106
49	7.35	50	1		100	50	1 000	0	84
Scenario 11									
90	13.5	0	8	2	100	50	1 000	10 000	370
66	9.9	26	6.5	1.5	100	50	1 000	10 000	294
52	7.8	41	5.7	1.3	100	50	1 000	10 000	260
30	4.5	65	3.9	1.1	100	50	1 000	10 000	212
20	3	75	4.0	1	100	50	1 000	10 000	198

References

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman and Hall, London, 416 pp.

Berges, J.A., 1997. Ratios, regression statistics, and 'spurious' correlations. Limnol. Oceanogr. 42, 1006–1007.

Brumsack, H.-J., 1989. Geochemistry of recent TOC-rich sediments from the Gulf of California and the Black Sea. Geol. Rundsch. 78, 851–882.

Brumsack, H.-J., 1991. Inorganic geochemistry of the German 'Posidonia Shale': paleoenvironmental consequences. In: Tyson, R.V., Pearson, T.H. (Eds.), Modern and Ancient Continental Shelf Anoxia. Geol. Surv. Spec. Publ. 58, The Geological Society, London, pp. 353–362.

Butler, J.C., 1986. The role of spurious results in the development of a komatiite alteration model. J. Geophys. Res. 91 (B13), E275–E280.

Calvert, S.E., Pedersen, T.F., 1993. Geochemistry of Recent

oxic and anoxic sediments: Implications for the geological record. Mar. Geol. 113, 67–88.

Calvert, S.E., Bustin, R.M., Ingall, E.D., 1996. Influence of water column anoxia and sediment supply on the burial and preservation of organic carbon in marine shales. Geochim. Cosmochim. Acta 60, 1577–1594.

Calvert, S.E., Fontugne, M.R., 2001. On the late Pleistocene–Holocene sapropel record of climatic and oceanographic variability in the eastern Mediterranean. Paleoceanography 16, 78–94.

Chayes, F., 1971. Ratio Correlation. The University of Chicago Press, Chicago, IL, 99 pp.

Gingele, F.X., Zabel, M., Kasten, S., Bonn, W.J., Nürnberg, C.C., 1999. Biogenic barium as a proxy for paleoproductivity: Methods and limitations of application. In: Fischer, G., Wefer, G. (Eds.), Use of Proxies in Paleoceanography: Examples from the South Atlantic. Springer-Verlag, Berlin, pp. 345–364.

Gromet, L.P., Dymek, R.F., Haskin, L.A., Korotev, R.L., 1984. The 'North American shale composite': Its compilation, major and trace element characteristics. Geochim. Cosmochim. Acta 48, 2469–2482.

Kim, J.-H., 1999. Spurious correlation between ratios with a common divisor. Stat. Probab. Lett. 44, 383–386.

Middelburg, J.J., Soetaert, K., Herman, P.M.J., 1997. Empirical relationships for use in global diagenetic models. Deep-Sea Res. I 44, 327–344.

Morford, J.L., Emerson, S., 1999. The geochemistry of redox sensitive trace metals in sediments. Geochim. Cosmochim. Acta 63, 1735–1750.

Pearson, K., 1896. On a form of spurious self-correlation which may arise when indices are used in the measurements of organs. Proc. R. Soc. London 60, 489–502.

Piper, D.Z., Isaacs, C.M., 1995. Geochemistry of Minor Elements in the Monterey Formation, California: Seawater Chemistry of Deposition. US Geol. Surv. Prof. Paper 1566. US Government Printing Office, Washington, DC, 41 pp.

Piper, D.Z., Isaacs, C.M., 1995b. Minor elements in Quaternary sediment from the Sea of Japan: A record of surface-water productivity and intermediate-water redox conditions. Geol. Soc. Am. Bull. 107, 54–67.

Reichart, G.J., Den Dulk, M., Visser, H.J., Van der Weijden, C.H., Zachariasse, W.J., 1997. A 225 kyr record of dust supply paleoproductivity and the oxygen minimum zone from the Murray Ridge (northern Arabian Sea). Palaeogeogr. Palaeoclimatol. Palaeoecol. 134, 149–169.

Rollison, H., 1993. Using Geochemical Data. Longman, Harlow, 352 pp.

Schenau, S.J., Antonarakou, A., Hilgen, F.J., Lourens, L.J., Nijenhuis, I.A., Van der Weijden, C.H., Zachariasse, W.J., 1999. Organic-rich layers in the Metochia section (Gavdos, Greece): Evidence for a single mechanism of sapropel formation during the past 10 My. Mar. Geol. 153, 117–135.

Schenau, S.J., Prins, M.A., De Lange, G.J., Monnin, C., 2001. Barium accumulation in the Arabian Sea: Controls on barite preservation in marine sediments. Geochim. Cosmochim. Acta 65, 1545–1556.

Shimmield, G.B., 1992. Can sediment geochemistry record changes in coastal upwelling paleoproductivity? Evidence from northwest Africa and the Arabian Sea. In: Summerhayes, C.P., Prell, W.L., Emeis, K.C. (Eds.), Upwelling Systems: Evolution since the Early Miocene. Geol. Soc. Spec. Publ. 64, pp. 29–46.

Shimmield, G.B., Pedersen, T.F., 1990. The geochemistry of reactive trace metals and halogens in hemipelagic continental margin sediments. Aquat. Sci. 3, 255–279.

Thomson, J., Higgs, N.C., Croudace, I.W., Colley, S., Hydes, D.J., 1993. Redox zonation of elements at an oxic/post-oxic boundary in deep-sea sediments. Geochim. Cosmochim. Acta 57, 579–595.

Thomson, J., Nixon, S., Croudace, I.W., Pedersen, T.F., Brown, L., Cook, G.T., MacKenzie, A.B., 2001. Redox-sensitive element uptake in north-east Atlantic Ocean sediments (Benthic Boundary Layer Experiment sites). Earth Planet. Sci. Lett. 184, 535–547.

Turekian, K.K., Wedepohl, K.H., 1961. Distribution of the elements in some major units of the Earth's crust. Geol. Soc. Am. Bull. 72, 175–192.

Von Rad, U., Schulz, H., Riech, V., Den Dulk, M., Berner, U., Sirocko, F., 1999. Multiple monsoon-controlled breakdown of oxygen-minimum conditions during the past 30,000 years documented in laminated sediments off Pakistan. Palaeogeogr. Palaeoclimatol. Palaeoecol. 152, 129–161.

Wedepohl, K.H., 1971. Environmental influences on the chemical composition of shales and clays. Phys. Chem. Earth 8, 307–333.