

# Characterisation of long-term climate change by dimension estimates of multivariate palaeoclimatic proxy data

R. Donner<sup>1</sup> and A. Witt<sup>1,2</sup>

<sup>1</sup>Nonlinear Dynamics Group, Department of Physics, University of Potsdam, Am Neuen Palais 10, 14 469 Potsdam, Germany

<sup>2</sup>Department of Geography, King's College London, Strand, London WC2R 2LS, England, UK

Received: 18 April 2006 – Revised: 17 July 2006 – Accepted: 17 July 2006 – Published: 21 September 2006

**Abstract.** The problem of extracting climatically relevant information from multivariate geological records is tackled by characterising the eigenvalues of the temporarily varying correlation matrix. From these eigenvalues, a quantitative measure, the linear variance decay (LVD) dimension density, is derived. The LVD dimension density is shown to serve as a suitable estimate of the fractal dimension density. Its performance is evaluated by testing it for (i) systems with independent components and for (ii) subsystems of spatially extended linearly correlated systems.

The LVD dimension density is applied to characterise two geological records which contain information about climate variability during the Oligocene and Miocene. These records consist of (a) abundances of different chemical trace elements and (b) grain-size distributions obtained from sediment cores offshore the East Antarctic coast. The presented analysis provides evidence that the major climate change associated with the Oligocene-Miocene transition is reflected in significant changes of the LVD dimension density. This is interpreted as a change of the interrelationships between different trace elements in the sediment and to a change of the provenance area of the deposited sediment.

## 1 Introduction

Variations of the Earth's climate occur on many very different temporal scales related to different driving forces. The annual cycle is caused by the variations of the solar net radiation. Interannual variability patterns like the El Niño/Southern Oscillation (ENSO) (Battisti and Sarachik, 1995) or the North-Atlantic Oscillation (NAO) (Hurrell et al., 2003) are related to the internal variability of the atmosphere-ocean system. Solar cycles on the decadal scale (as with a

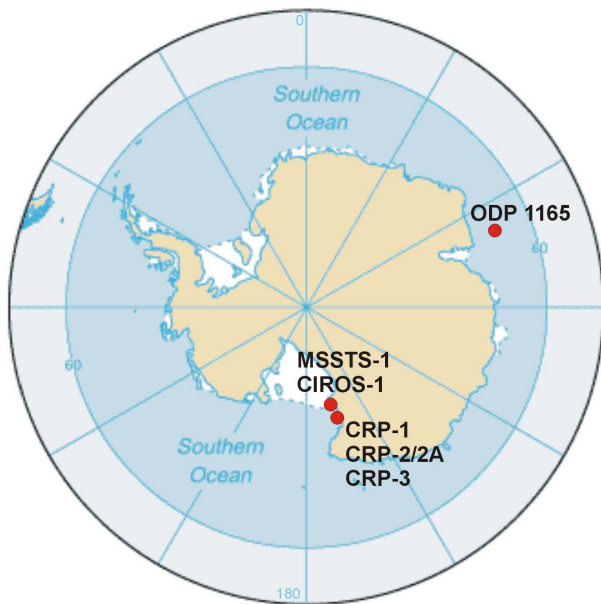
period around 11 years or the Gleisberg cycle with an average period length of 88 years) are known to drive climate variations (Lean et al., 1995). Centennial scale solar oscillations affect ocean circulation. The orbital dynamics of the Sun/Earth system are responsible for climate variations on millennial scales and in particular for glaciation/deglaciation cycles (Paillard, 2001). Plate tectonics, leading to an opening and closure of oceanic gateways and the formation of mountain belts due to continental collisions, directly influence the atmospheric and oceanic circulation on very long time scales (Hay, 1996; Florindo et al., 2005).

Variations of environmental conditions are reflected by variations of parameters obtained from palaeoclimatic proxy data such as physical or chemical properties of sediment and ice cores, or tree rings. Depending on the particular site and measurement, paleoclimatic proxy data are tracers of local or global climate conditions. To achieve an overall picture of the information included in the geological source under investigation, complementary measurements and analyses need to be performed. These analyses involve measurements of physical, chemical, and (in the case of sedimentary sequences) biological or sedimentological parameters. Furthermore, age models that quantify the age-depth relationship of a sediment or ice core incorporating information about their uncertainties must be developed (Buck and Millard, 2004; Telford et al., 2004).

The intention of geological studies is to identify signatures of climate change and relate them to the variability of climatologically meaningful quantities. A direct interpretation in terms of meteorological parameters from the individually measured series is often not possible. Therefore, it is a standard approach to derive variability patterns from multivariate geological time series (Bradley, 1999). These patterns can be associated with changes of climatic observables like temperature, moisture conditions (i.e., seasonal precipitation, snow volumes), vegetation cover, or the strength and location of different atmospheric oscillatory patterns. For this purpose,

---

Correspondence to: R. Donner  
(reik@agnld.uni-potsdam.de)



**Fig. 1.** Map of the Antarctic continent including the locations of the Cape Roberts project drill sites and the preceding MSSTS-1 and CIROS-1 drillings. In addition, the location of ODP site 1165 (Williams and Handwerger, 2005), the site of collection of what is probably the best studied Antarctic offshore marine core, is displayed.

transfer functions are frequently applied (Sachs et al., 1977; Fischer et al., 2004), which may be a potential matter of criticism (Telford and Birks, 2005).

Most palaeoclimatic time series are the result of extensive and expensive measurement campaigns. Surprisingly, it was typically not discussed which complementary information can be extracted from the resultant multivariate records. In our paper, we address this problem and propose the consideration of dimension densities of multivariate records as a measure of the information content. For this purpose, we derive appropriate quantities and describe their major strengths and weaknesses by applying them to data sets generated by stationary model systems. For the considered paleoclimate proxy data, an age-dependent calculation of these measures (i.e., a separate computation for different parts of the climate history) allows one to infer the variability of the appropriately quantified information content.

In this paper, we illustrate the considered problem for data obtained from marine sediment cores collected off the shore of the East Antarctic coast (see Fig. 1). The first data set comprises the abundances of 32 trace elements that are measured for 60 time slices, the second data set contains grain size histograms with respect to 23 size classes which are measured for 129 time horizons. The number of measurement points is too low to apply one of the numerous approaches from nonlinear time series analysis (see the text books of,

e.g., Priestley, 1988; Tong, 1990; Abarbanel, 1996; Kantz and Schreiber, 1997; Diks, 1999, or the specific references given in Sect. 3). To treat such short multivariate data appropriately, we have recently proposed (Donner and Witt, 2006) that one measures the strength of the interrelationships of the individual univariate measurements series of the multivariate record. Our method is based on a statistical decomposition of the covariance matrix of time windows of the related data. In particular, we have studied the remaining variances as a function of the number of eigenmodes and assumed an exponential decay model. We demonstrated that this approach leads to qualitatively robust results even for very short data sets.

This manuscript presents a detailed discussion about the general strengths and weaknesses of the considered approach and its application to a particular palaeoclimatic example. It is organised as follows: In Sect. 3, we review the approach of estimating the dimension density by characterising the decay of eigenvalues of the covariance matrix using Karhunen-Loève decomposition (KLD). This idea was already proposed by Zoldi and Greenside (1997) for estimating the so-called KLD dimension of spatio-temporal chaotic systems. Then, the linear variance decay (LVD) dimension is introduced as a more sensitive estimate. The LVD dimension is, in contrast to the KLD dimension, defined as a continuous measure. The corresponding advantages and disadvantages are discussed in Sect. 4 by comparing the performance of both characteristics for uncorrelated data and for a model system of spatio-temporal chaos. In Sect. 5, we apply our method to real-world geological data sets. Our findings are discussed from a palaeoclimatic point of view. Finally, a brief outlook of possible generalisations and further fields of application is given.

## 2 Description of the data

We analyse data from marine sediments obtained within the framework of the Cape Roberts project off the shore of the East Antarctic coast. This campaign consisted of three scientific drillings at slightly different locations. The main objective of the Cape Roberts project was to carry out a detailed study of glaciation and deglaciation intervals in the Antarctic region in a time interval of between about 17 to 34 million years Before Present (BP). During this time interval, the global mean temperature was significantly higher than today with long-term temperature fluctuations caused by orbital cycles (for more details and a list of references, see Naish et al., 2001). During the same period, tectonic activity caused openings and closures of ocean passages. The resultant dramatic changes in ocean circulation are responsible for changing atmospheric dynamics. In addition to the related signal, the variability on time scales associated to variations of orbital parameters is also preserved in the data.

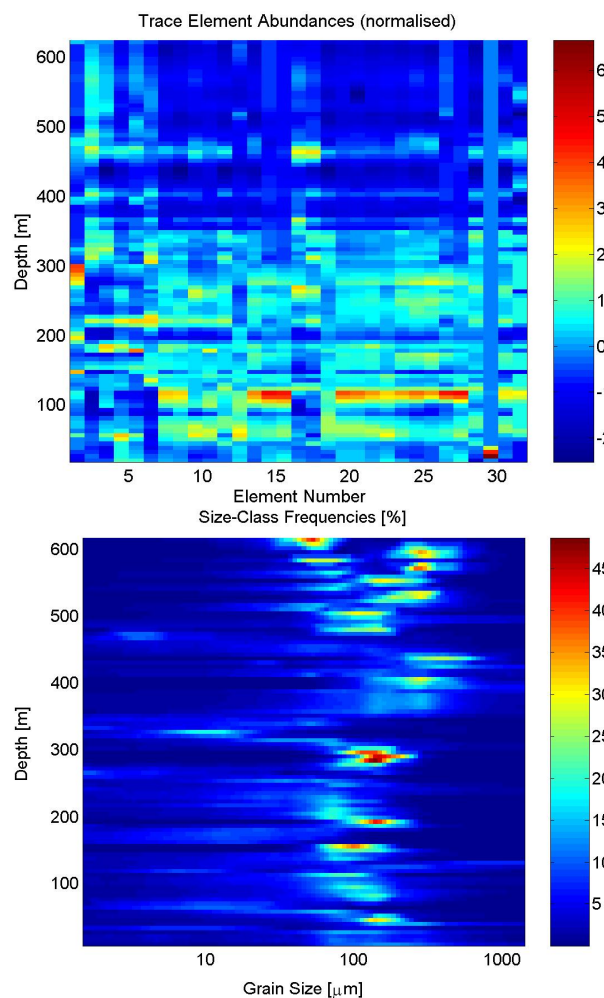
In the sediment core CRP-2/2A (see Fig. 1), the Oligocene-Miocene transition (OMT), the probably most remarkable climatic transition in the considered time interval, is well resolved. The OMT is characterised by an opening of the Drake passage between South America and the Antarctic continent, which led to an intensification of the Antarctic circumpolar current and a successive thermic isolation of the continent. These effects finally caused an enhanced glacial variability in the high latitudes of the Southern hemisphere. Although the core does resolve the OMT well, the data of palaeoclimatic proxies like trace element abundances or grain-size distributions have only been obtained for a very small number of time slices. Thus, the actual mechanism of the transition is not well resolved in the corresponding component time series. In this paper, we aim to improve the explanatory power of the considered data set by considering all components instantaneously. In particular, we focus on the question of whether the different climatic conditions in the older and the younger part of these records (Naish et al., 2001) are reflected by varying strength and changing patterns of the interrelationships between the different palaeoclimatic observables.

In our recent paper (Donner and Witt, 2006), we already started to study a record of trace element abundances from the CRP-2A core (Krissek and Kyle, 2000; Krissek, 2004). The analysed data set consists of 46 parameters measured within 104 slices of the sedimentary core. Due to gaps in the measurement series, we have analysed a homogeneous subset of records of 32 trace elements from 60 time slices. The trace element abundances were measured by X-ray fluorescence (XRF, 19 elements) and instrumental neutron activation analysis (INAA, 13 elements)<sup>1</sup>. As the absolute abundances of these elements (given in parts per million (ppm)) cover several orders of magnitude, we first standardised all component time series to unit variance before subjecting the data set to further analysis.

Complementary information is provided by studying an independent paleoclimatic proxy: a record of grain size distributions. Grain-size distributions can be obtained by sieving (mass-frequency distributions) or more advanced optical measurement techniques applied to the suspended material (number-frequency distributions). For the CRP-2/2A core, the corresponding data were obtained using a SEDI-GRAPH 5100 which measures the absorption of x-rays during the sedimentation of the material. The data (Barrett and Anderson, 2000, 2003) consist of relative frequencies of particles in 23 different size classes, which are equally spaced on a logarithmic scale (phi-scale). In total, measurements have been performed on 119 different time slices.

On the one hand, trace element abundances give important information about the mineralogical composition and, hence, the origin of the sedimented material. On the other hand,

<sup>1</sup>Among the 32 parameters, the abundances of arsenic, thorium and uranium have been measured twice with both methods.



**Fig. 2.** Palaeoclimatic data sets used in this study. Upper panel: Abundances (normalised to zero mean and unit variance) of the 32 trace elements S, Sc, V, Cr, Ni, Cu, Zn, Ga, As (twice), Rb, Sr, Y, Zr, Nb, Sb, Cs, Ba, La, Ce, Sm, Eu, Tb, Yb, Lu, Hf, Ta, Pb, Th (twice) and U (twice) (from left to right). Lower panel: Frequencies of grains with sizes between about 1 and 1000  $\mu\text{m}$ .

grain-size distributions are related to the mode of transportation of the material from its source to the final debris and may be modified by different erosional mechanisms. Hence, this proxy may provide insights into the dominating transport and deposition mechanisms of the sediment (McBride, 1971).

All data sets used in this study are publically available from the PANGAEA data base (<http://www.pangaea.de>). In addition, a complete list of all measurements and some information about possible climatic implications can be found in Krissek and Kyle (2000) (trace element abundances) and Barrett and Anderson (2000) (grain-size distributions), respectively. To get an impression of the contained variability, both data sets are displayed in Fig. 2.

### 3 Methodology

For analysing univariate time series, a number of nonlinear techniques can be applied to characterise the complexity of the data. However, it is awkward to work with palaeoclimatic data. Due to their usually rather coarse temporal resolution with an uneven sampling and the uncertainties of both the measured data and the applied age model, an appropriate analysis requires some especially sophisticated methods. Traditional approaches of time series analysis can thus hardly be applied to such data, which was demonstrated for fractal dimensions estimated with the Grassberger-Procaccia algorithm (Grassberger and Procaccia, 1983) for some marine isotopic records (Mudelsee and Stattegger, 1994).

In the case of multivariate data, several nonlinear methods originally developed for univariate data may be generalised, including the Lyapunov spectrum (Bünner and Hegger, 1999), the scaling of fractal dimensions (Politi and Witt, 1999), or dimension densities based on a normalised Grassberger-Procaccia algorithm (Bauer et al., 1993). However, as shown by Olbrich et al. (1998), the dimensions calculated from data sets depend crucially on the resolution of the observations. To overcome the corresponding difficulties, Raab and Kurths (2001) have proposed a normalisation for approaching large-scale correlation dimension densities.

A multivariate time series can be statistically decomposed by a variety of different approaches: by purely linear methods like the Karhunen-Loève decomposition (KLD) (which is also known as Empirical Orthogonal Function (EOF) method or Principal Component Analysis (PCA) (Jolliffe, 1986; Preisendorfer, 1988)) or by nonlinear approaches like Nonlinear Principal Component Analysis (NLPCA) (Kramer, 1991), Locally Linear Embedding (LLE) (Roweis and Saul, 2000), Isometric Feature Mapping (ISOMAP) (Tenenbaum et al., 2000), or Independent Component Analysis (ICA) (Hyvärinen et al., 2001). All mentioned nonlinear methods require sufficiently long stationary time series and, therefore, they cannot appropriately applied to the considered palaeoclimatic data sets. That is the reason why we hark back to linear statistical decomposition.

In this work, we restrict our considerations to the case of Karhunen-Loève decomposition (KLD). These days, KLD is a standard method for compressing spatiotemporal data by finding the largest linear subspace that contains substantial statistical variations of the data. In the case of observations with  $N$  simultaneously measured variables and  $M$  points in time, the  $M \times N$ -dimensional data matrix  $\mathbf{A}$  (rescaled to zero means for any component time series) is used to define an  $N \times N$ -dimensional symmetric and positive semidefinite scatter matrix  $\mathbf{S} = \mathbf{A}^T \mathbf{A}$ . The matrix  $\mathbf{S}$  can be completely described by its non-negative eigenvalues  $\sigma_i^2$  ( $i=1, \dots, N$ ) and their corresponding eigenvectors, which in the geoscientific community are called empirical orthogonal functions (EOF). Without loss of generality, we consider the  $\sigma_i^2$  of  $S$  in decreasing order  $\sigma_1^2 \geq \dots \geq \sigma_N^2 \geq 0$ . In addition, we will nor-

malise the eigenvalues to unit sum  $\sum_{i=1}^N \sigma_i^2 = 1$  wherever appropriate.

#### 3.1 KLD dimension density

By definition, the KLD is a purely linear method because it characterises the covariance matrix. Zoldi and Greenside (1997) have extended the range of applicability to a quantitative characterisation of spatio-temporal chaotic systems by introducing the KLD dimension. Meanwhile, this approach is commonly applied to quantify the degree of complexity of high-dimensional nonlinear systems (Meixner et al., 2000; Varela et al., 2005).

In this paper, we define the KLD dimension as the number of eigenvalues needed to capture some specified fraction  $0 \leq f \leq 1$  of the total variance  $\sum_{i=1}^N \sigma_i^2$  of the data:

$$D_{\text{KLD}}(f) = \min \left\{ p : \sum_{i=1}^p \sigma_i^2 / \sum_{i=1}^N \sigma_i^2 \geq f \right\}. \quad (1)$$

The limiting cases are  $D_{\text{KLD}}(0)=0$  and  $D_{\text{KLD}}(1)=N$ . This is a modified version of the original definition of the KLD dimension used by Zoldi and Greenside (1997) and Meixner et al. (2000) who defined  $D_{\text{KLD}}$  as the maximum number of eigenmodes describing less than a fraction of  $f$  of the total variance. The advantage of our version of the definition concerns the limiting cases  $f=0$  and  $f=1$ .

In the case of simulations of spatio-temporally chaotic systems, Zoldi and co-workers observed a linear scaling of  $D_{\text{KLD}}$  with the system size  $N$  for any  $f$ . This finding justifies the normalisation of the KLD dimension to the KLD dimension density  $\delta_{\text{KLD}} = D_{\text{KLD}}/N$  (Meixner et al., 2000), which has values within the unit interval.

So far, the KLD dimension has been mainly determined for simulated data of idealised model systems. Recently, Varela et al. (2005) have applied  $D_{\text{KLD}}$  in an investigation of spatiotemporal data from electrochemical oscillator experiments (with  $M \geq 6000$  and  $N=50$ ) and the authors demonstrated that this measure is well suited for quantifying differences between regular and turbulent states.

To adapt the concept of KLD dimension density to non-stationary systems, we propose to analyse the temporal variability of  $\delta_{\text{KLD}}$ . If  $\delta_{\text{KLD}}$  is derived from the scatter matrix  $\mathbf{S}$  with respect to the complete data set, it quantifies the temporal average of the complexity of interrelationships between the different components. A separate computation of the KLD dimension density,  $\delta_{\text{KLD}}$ , for sliding temporal windows (Meixner et al., 2000), however, gives a resolution of the varying complexity down to the scale of  $N$  points in time or even below.

#### 3.2 LVD dimension density

The KLD dimension can be widely applied to large data sets from spatio-temporally chaotic systems. Its direct application to an observational record with rather few observables

(i.e., small  $N$ ) or a low number of measurement points along the time axis (small  $M$ ) is problematic for the following reasons:

1.  $\delta_{\text{KLD}}$  has a range of  $N+1$  different, equally spaced values. As a consequence, this measure cannot respond to small changes of the structure of the interrelationships between the component time series. If  $N$  is rather small, only strong changes within the data lead to a (dramatically) different  $\delta_{\text{KLD}}$ .
2.  $\delta_{\text{KLD}}$  depends on the choice of the cutoff parameter  $f$ . For a fixed cutoff  $f$ ,  $\delta_{\text{KLD}}$  must be considered to be a *relative* rather than an *absolute* dimension density. For applications where only a qualitative detection and description of changes of the complexity of interrelationships within a multivariate data set is requested, this subtle difference is negligible.

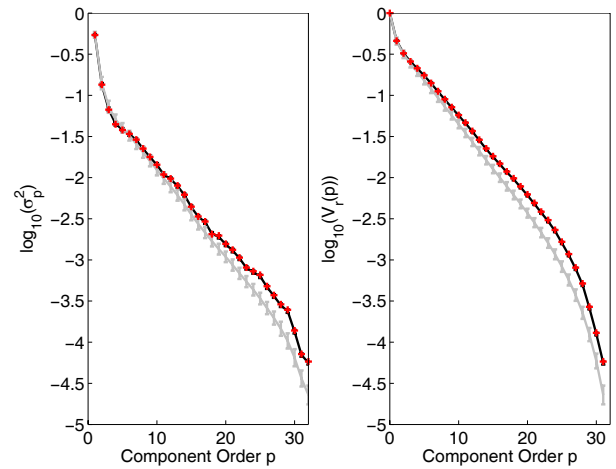
3. Due to the small amount of observational data in time, certain finite-size effects have to be expected which may cause any quantitative interpretation of  $\delta_{\text{KLD}}$  to fail.

These arguments call for further modifications of the original concept. As one possibility, we recently introduced the linear variance decay (LVD) dimension density  $\delta_{\text{LVD}}$  (Donner and Witt, 2006) that leads even in the case of small data sets to reliable results. This measure can be obtained by considering the scaling of  $\delta_{\text{KLD}}$  with the cutoff parameter  $f$  and fitting a suitable parametric function to the resultant curve.

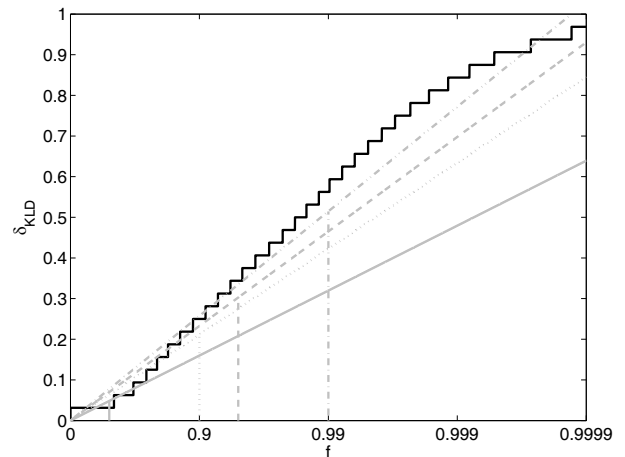
Now, an appropriate model function for the scaling of  $\delta_{\text{KLD}}$  must be introduced. The scaling of the component variances,  $\sigma_i$ ,  $i=1, \dots, N$ , have been investigated for random matrices  $\mathbf{A}$  (Farmer, 1971; Probert-Jones, 1973) as well as real-world geoscientific data (Craddock and Flood, 1969) in terms of logarithmic eigenvalue (LEV) curves (for an overview, see Preisendorfer, 1988). In contrast to the component variances, there are no studies investigating the scaling of the remaining variances. To exploit the scaling behaviour of the remaining variances, one has to be aware that for a given value of  $\delta_{\text{KLD}}(f)=p/N$  ( $p=0, \dots, N$ ),  $1-f$  plays the role of the remaining variances defined as  $V_r(p/N)=1-\sum_{i=1}^p \sigma_i^2$  for  $p=1, \dots, N$  ( $V_r(0)=1.0$ ), where  $p/N$  is the relative number of components considered. An inspection of the corresponding values for both random matrices as well as observational data, shows that the decay that corresponds to the components' largest variances can be well described by an exponential decay (see Fig. 3). As a consequence, we make the following ansatz:

$$V_r(p/N) = 10^{-\frac{p}{N}/\delta_{\text{LVD}}} \text{ for } p \leq p_{\text{max}} < N. \quad (2)$$

The corresponding value for  $\delta_{\text{LVD}}$  can be computed by a simple linear least squares approach. However, if  $N$  is rather small, there are only a few points to which the model function can be fitted. And again, we find only  $N$  possible choices of the threshold  $p_{\text{max}}$  for fitting this function: Since  $V_r(N)=0.0$  by definition, an exponential decay law must be subjected to a certain cutoff at  $p_{\text{max}} < N$ . To overcome this difficulty and define the model function with respect to a continuous cutoff parameter  $f$ , we make use of the relationship between



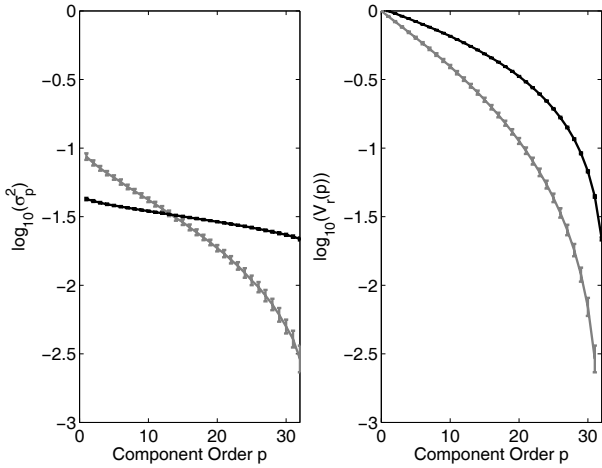
**Fig. 3.** Scaling of the component variances  $\sigma_p^2$  (left panel) and the corresponding remaining variances  $V_r(p)$  (right panel) for the trace element abundance data ( $M=60$  and  $N=32$ , red crosses) discussed in Sect. 5. For comparing the results with those of finite-size random matrices, we additionally computed  $V_r(p)$  for ensembles of 1000 multivariate ( $N=32$ ) surrogate data sets consisting of normally distributed data (with prescribed component variances equal to those of the original data) with length  $M=60$  (gray line) and  $M=1000$  (black line) points in time. The displayed error bars correspond to the standard deviations of the values from the respective surrogate data. The deviation between the black and the gray curve is mainly explained by the small number and non-Gaussian distribution of the observed time series values.



**Fig. 4.** Scaling of the KLD dimension density  $\delta_{\text{KLD}}(f)$  with  $\log(1-f)$  (black line) for the trace element abundance data discussed in Sect. 5. Vertical gray lines indicate the cutoff values of  $f=0.5$  (solid),  $0.9$  (dotted),  $0.95$  (dashed), and  $0.99$  (dash-dotted), whereas the slope of the associated gray diagonal lines correspond to the respective values of  $\delta_{\text{LVD}}(f)$ .

$V_r(p)$  and  $1-f$  which is illustrated in Figs. 3 and 4: reversing the axes in Fig. 4 and multiplying  $\delta_{\text{KLD}}$  by  $N(=32)$ ,





**Fig. 5.** Scaling of the component variances  $\sigma_p^2$  (left panel) (LEV curve) and the corresponding remaining variances  $V_r(p)$  (right panel) for multivariate Gaussian distributed white noises with  $N=32$  and  $M=60$  (gray lines) as well as  $M=1000$  (black lines) points in time, resp. The displayed error bars correspond to the standard deviations of the values from 100 realisations.

one approaches a continuously defined equivalent of the right panel in Fig. 3 (where the illustrated function is defined only for integer values of  $p$ ). Thus, we consider a scaling law of the KLD dimension density corresponding to that of the remaining variances as

$$\delta_{\text{KLD}}(\phi) = -\delta_{\text{LVD}}(f) \log(1 - \phi) \text{ for } \phi \in [0, f]. \quad (3)$$

The value of  $\delta_{\text{LVD}}(f)$  is called the *linear variance decay* (LVD) dimension density. As  $\delta_{\text{KLD}}(f)$  is well-defined for  $f \in [0, 1]$ , this expression allows to calculate  $\delta_{\text{LVD}}(f)$  for any  $f \in (0, 1)$ . The values of  $\delta_{\text{LVD}}$  are calculated by minimising the functional

$$F_\alpha(f) = \int_{\log(1-f)}^0 (\delta_{\text{KLD}}(x) + \alpha x)^2 10^x dx \quad (4)$$

with respect to  $\alpha$  (here, we have made use of the transformation  $x = \log(1 - \phi)$ ).  $F_\alpha(f)$  has (for any value of  $f$ ) a unique global minimum at

$$\alpha_{\min}(f) = -\frac{\int_{\log(1-f)}^0 \delta_{\text{KLD}}(x) x 10^x dx}{\int_{\log(1-f)}^0 x^2 10^x dx} \quad (5)$$

which can be computed easily. This minimum,  $\alpha_{\min}(f)$ , corresponds to the LVD dimension density  $\delta_{\text{LVD}}(f)$ .

The LVD dimension density still depends on the cut-off parameter  $f$ , i.e., gives only a qualitative dimension density estimate. The appealing alternative to obtain a parameter-free measure by, for example, taking the minimum or maximum of  $\delta_{\text{LVD}}$  over all values of  $f$  has severe disadvantages: there is only a local minimum and maximum of  $\delta_{\text{LVD}}(f)$  for  $f$  within the open interval  $(0, 1)$  since

$\log(1 - f) \rightarrow 0$  as  $f \rightarrow 0$  ( $\delta_{\text{LVD}} \rightarrow +\infty$ ) and  $\log(1 - f) \rightarrow -\infty$  as  $f \rightarrow 1$  ( $\delta_{\text{LVD}} \rightarrow 0$ ). Moreover, the local minimum of  $\delta_{\text{LVD}}$  taken over all  $f \in (0, f_{\max})$  always occurs at  $f = 1 - V_r(p)$  for a suitable  $p \in \{1, \dots, N - 1\}$ . Thus, a dynamic characterisation of the record by this local minimum LVD dimension density is not suitable as it may occur at completely different values of  $f$  (possibly even changing discontinuously if the associated value of  $p$  changes with time).

As a consequence, we propose that  $\delta_{\text{LVD}}$  is always considered as a qualitative dimension estimate corresponding to a fixed value of  $f \in (0, 1)$ . Although it still shares this disadvantage with the KLD dimension density,  $\delta_{\text{LVD}}$  is much more sensitive with respect to minor changes in the correlations of the component time series and simultaneously applicable to very small data sets. We will demonstrate this in more detail in the following section.

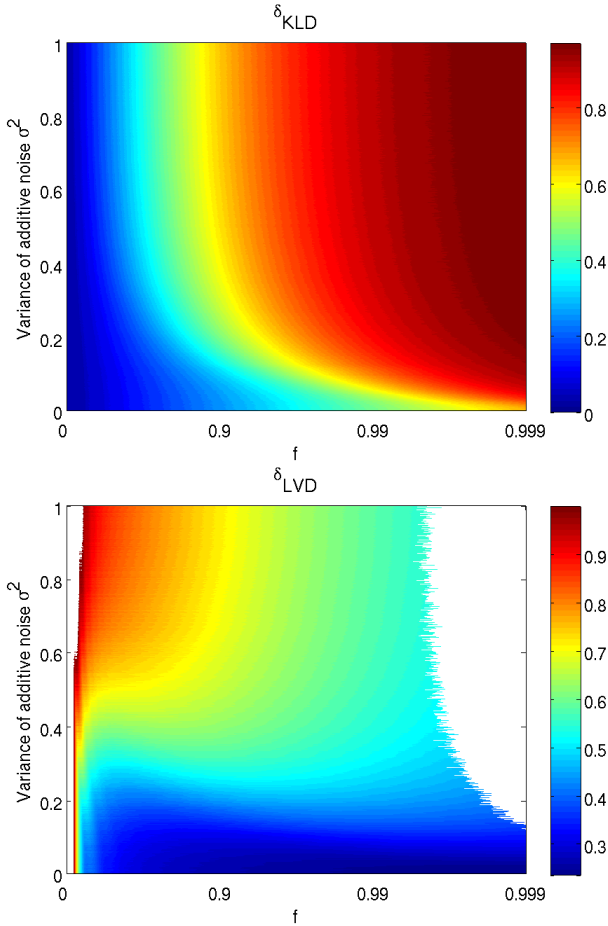
#### 4 Application to model systems

Before we are able to apply our approach to measured palaeoclimatic time series, it is necessary to discuss the performance of  $\delta_{\text{KLD}}$  and  $\delta_{\text{LVD}}$  for short and noisy multivariate data sets. For this purpose, we study both measures for different synthetic data sets in the limit of small data sets (i.e., either  $N$  or  $M$  is comparable to typical geological time series).

##### 4.1 Independent stochastic data

Spectra of eigenvalues of covariance matrices of multivariate Gaussian distributed white noise (i.e., there are no temporal correlations and the components are statistically non-correlated) with a limited length have been extensively studied both, analytically and numerically (for an overview and further references, see chapter 5 of Preisendorfer, 1988). In particular, there are analytic expressions for the eigenvalue spectra of such matrices. The resulting logarithmic eigenvector curves show a quasi-exponential decay of values steepening towards the major components as well as towards the components with the smallest variances. Our numerical calculations of covariance matrices of Gaussian distributed white noise, which are displayed in Fig. 5, resemble the results of Preisendorfer, 1988, p. 240. In addition, we show the corresponding decay curves of the remaining variances  $V_r(p)$  which start to significantly deviate from an exponential decay for lower numbers of considered components than the corresponding LEV curve. Nonetheless, for the high amplitude components, the exponential model appears to be a good approximation.

As a next step, we study the influence of additive Gaussian white noise on the eigenvalues and remaining variances of the covariance matrices. For this purpose, we modify the above setting by considering Gaussian distributed white noise with a component dependent standard



**Fig. 6.** Color-coded representations of  $\delta_{\text{KLD}}$  (upper panel) and  $\delta_{\text{LVD}}$  (lower panel) for Gaussian distributed white noise with  $N=32$  and  $M=60$  and with component variances  $\sigma_p^2 = \exp(-\frac{p}{N}/\delta)$  ( $\delta=0.2$ ) as a function of the cutoff level  $f$  for different additive noise amplitudes  $\sigma^2$  (mean values from 100 realisations). White areas correspond to parameters where  $\delta_{\text{LVD}}$  either could not be computed (very large  $f$ ) or gave artificially high values  $>1$  (very small  $f$ ).

deviation (i.e.,  $\sigma_p^2 \neq 1$ ). As an example, we prescribe the component variances to have exponentially decaying values  $\sigma_p^2 = \exp(-\frac{p}{N}/\delta)$ . In this case, additive noise dominates the decay of the eigenvalues only on scales where  $\sigma_p^2$  is somewhat smaller than the noise variance  $\sigma^2$  (which is in this example related to the fact that both, signal and noise, are the same kind of process). In contrast to the eigenvalues themselves, the remaining variances are much more sensitive to the noise and show remarkable deviations already for  $(V_r(p)/\sigma)^2 \sim 1$ . For component orders  $p$  where the eigenvalues and remaining variances are smaller than these thresholds (whose values are closely related to the specific setting), the noise leads to a significant change of the slope of the corresponding decay curves. Hence, the decay at these minor components is mainly described by the noise.

The different sensitivity of the eigenvalues of the covariance matrix and the corresponding remaining variances is reflected by a larger sensitivity of the LVD dimension density to additive noise compared to that of the KLD dimension which is – as a coarse-grained estimate – much more robust to reasonably small changes of the covariance structure of the data. In Fig. 6, the behaviour of the two measures is systematically studied as a function of both, the cutoff value  $f$  and the noise amplitude  $\sigma^2$ . In particular, one observes that  $\delta_{\text{KLD}}$  generally increases with  $f$ , whereas its values increase relatively slowly as noise amplitude grows. In contrast,  $\delta_{\text{LVD}}$  changes (for sufficiently large  $f$ ) moderately as the cutoff value  $f$  increases, but is still sensitive to changes in the amplitude of the applied noise. Note that in the case of  $\delta_{\text{LVD}}$ , the cutoff value  $f$  has to be chosen large enough to avoid the strong and unbounded increase in the values of this measure for  $f \rightarrow 0$ . Concerning the uncertainty of both dimension estimates, it is found that these are of similar orders of magnitude with maximum values at parameters where the corresponding measures have a large gradient.

#### 4.2 Subsets of large-scale systems

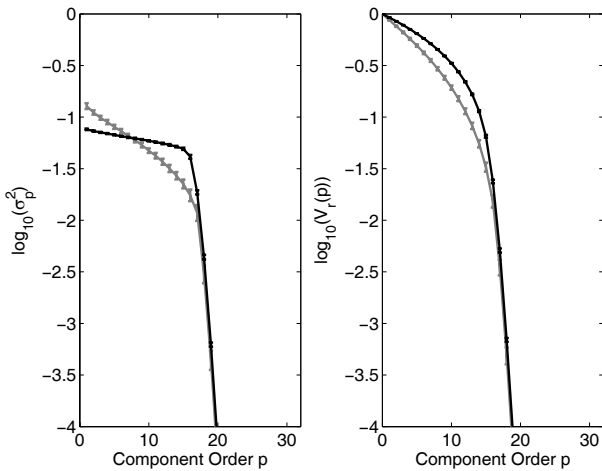
The case of temporarily and spatially uncorrelated component time series discussed so far is rather generic. Observational data from geoscientific systems are likely to have some deterministic, or even high-dimensional chaotic components. To demonstrate the power of KLD-based dimension estimates for such data sets, we propose studying their performance for systems that model the behaviour of spatio-temporal chaos. We utilise a computationally efficient model (Politi and Witt, 1999) of spatio-temporal chaotic dynamics that is constructed for a prescribed value of the dimension density  $d \in [0, 1]$ . This model is based on the Fourier basis vectors  $\{F_1, \dots, F_n\}$  of a sufficiently high-dimensional space (i.e.,  $n$  is large, in our case  $n=1000$ ):

$$F_{kj} = \begin{cases} 1/\sqrt{n}, & \text{if } k = 1, \\ \sqrt{2/n} \cos\left(\frac{2\pi}{n} \left[\frac{k}{2}\right] j\right), & \text{if } k > 1 \text{ and odd,} \\ \sqrt{2/n} \sin\left(\frac{2\pi}{n} \left[\frac{k}{2}\right] j\right), & \text{if } k \text{ even,} \end{cases} \quad (6)$$

(where  $[\cdot]$  denotes the integer part, and  $j=1, \dots, N \leq n$  gives the “spatial” position on a regular one-dimensional lattice). These vectors are used to construct component time series of the multivariate data set as

$$x_{ij} = \sum_{k=1}^{dn} \xi_{ik} F_{kj}. \quad (7)$$

Here,  $\xi_{ik}$  (with  $i=1, \dots, M$  corresponding to the position in time) is a set of random numbers taken from an appropriate distribution. If  $|\xi_{ik}| < 1$ , the values  $x_{ij}$  (that form the  $M \times N$ -dimensional data matrix) are located in a  $dn$ -dimensional hypercube. If  $M$  is sufficiently large, the eigenvalues  $\sigma_p^2$  of the associated covariance matrix show an abrupt decay at the



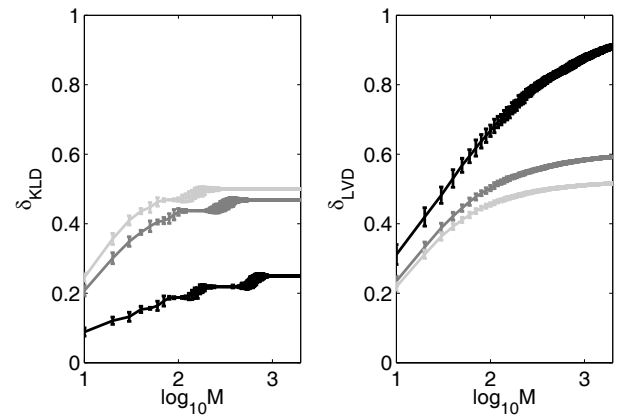
**Fig. 7.** Scaling of the component variances  $\sigma_p^2$  (left panel) (LEV curve) and the corresponding remaining variances  $V_r(p)$  (right panel) for the model system for spatio-temporal chaos ( $d=0.5$ ) with  $N=32$  and  $M=60$  (gray lines) as well as  $M=1000$  (black lines) points in time, resp. The displayed error bars correspond to the standard deviations of the values from 100 realisations.

component index  $dn$ , corresponding to the dimension of the underlying hypercube (see Fig. 7).

In our recent paper (Donner and Witt, 2006), we already used this system with  $\xi_{ik}$  taken from a uniform distribution on  $[-3^{1/3}, 3^{1/3}]$  to demonstrate the power of the KLD-based dimension estimates. By studying several simulations, we could show that  $\delta_{\text{KLD}}$  approaches the true dimension of the system better than  $\delta_{\text{LVD}}$ . For this system, the LVD dimension density shows a different behaviour compared to the case of random matrices: Whereas  $\delta_{\text{KLD}}$  increases with increasing  $f$  by definition,  $\delta_{\text{LVD}}$  decreases for this particular example. Now, we investigate the dependence of the two measures on the size of the system and on the true (prescribed) dimension density  $d$ .

Figure 8 presents simulations of the two quantities for different length of the multivariate time series  $M$ . We find that for suitably long time series ( $M \leq n=1000$ ), the KLD dimension density approaches constant values. Short time series cause poor estimates of the correlation matrix and, consequently,  $\delta_{\text{KLD}}$  is biased and has rather large confidence intervals. The LVD dimension density  $\delta_{\text{LVD}}$  is even more sensitive to small changes of the length of the time series: an increasing size of the record leads to a gradually increasing estimate of the dimension. Only for very long time series  $M \geq n$  do the computed values saturate. For both dimension estimates, it is found that saturation occurs for shorter time series if  $f$  is close to 1. Furthermore, if  $f$  and  $M$  are both chosen sufficiently large, both  $\delta_{\text{KLD}}$  and  $\delta_{\text{LVD}}$  approach the prescribed dimension density  $d$  of the system.

The sensitivity of KLD-based dimension estimates on the observational length allows a comparison of different data



**Fig. 8.** Scaling of the dimension estimates  $\delta_{\text{KLD}}$  (left panel) and  $\delta_{\text{LVD}}$  (right panel) for the model system of spatio-temporal chaos ( $d=0.5$ ) with  $N=32$  components as a function of the length  $M$  of the record for  $f=0.5$  (black),  $f=0.9$  (dark gray) and  $f=0.95$  (light gray). The displayed error bars correspond to the standard deviations of the values from 100 realisations.

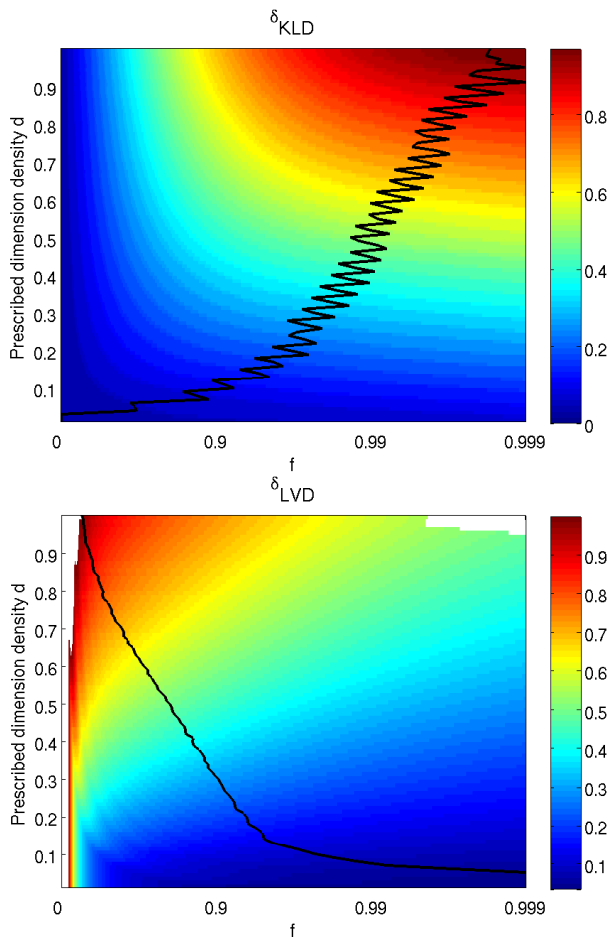
sets only if they have the same length. This restricts the applicability to geoscientific data sets: Sedimentary records are often unevenly sampled in time (age). Instead of comparing time windows of equal length, windows that are related to the same number of measurement data points have to be compared.

When considering variations of the true system dimension  $d$ , the “optimum” truncation level  $f$  to identify  $d$  with the considered technique increases with increasing  $d$  in the case of the KLD dimension density  $\delta_{\text{KLD}}$ , whereas for  $\delta_{\text{LVD}}$ , the opposite behaviour is found. The results shown in Fig. 9 underline our earlier findings (Donner and Witt, 2006) (note that the number values of both dimension density estimates differ from those in the reference due to a modified setting, i.e., a different choice of  $N$  and  $M$ ). As our study presented here is more systematic, we additionally find that the discrete values of  $\delta_{\text{KLD}}$  lead to oscillations of the optimum truncation level  $f$  for changes in  $d$ . In contrast,  $\delta_{\text{LVD}}$  changes with  $d$  in a smooth way, but leads (for our specific setting and “typical” values of  $f$ ) to a clearly worse quantitative estimate of  $d$ . In general, for fixed  $f$ , both dimension density estimates detect changes in the true dimension of the system. As the associated changes of their related values are discrete in the case of  $\delta_{\text{KLD}}$  but continuous for  $\delta_{\text{LVD}}$ , the latter measure is better suited for qualitatively detecting and describing changes of the correlation structure in multivariate data sets.

## 5 Palaeoclimatic example: Cape Roberts Project

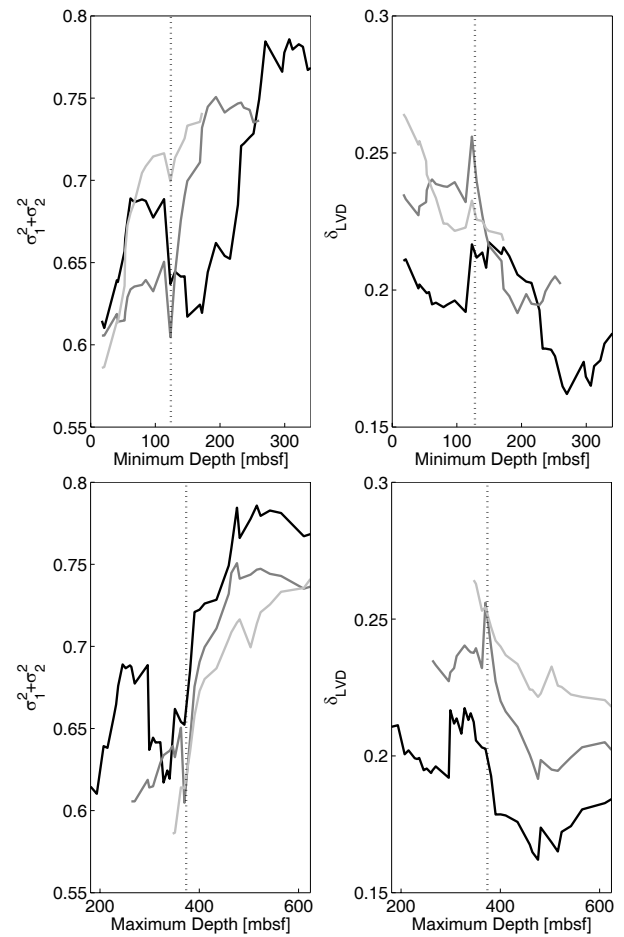
We start our analysis by considering the trace element record of the CRP-2/2A sediment core. Because only  $\delta_{\text{LVD}}$  is sufficiently sensitive to small changes of the correlation structure





**Fig. 9.** Color-coded representations of  $\delta_{\text{KLD}}$  (upper panel) and  $\delta_{\text{LVD}}$  (lower panel) for data from the space-time chaos model with  $N=32$  and  $M=60$  as a function of the cutoff level  $f$  for different prescribed dimension densities  $d$  of the system. The values are averages from 100 realisations of the respective system for each parameter. Black lines correspond to cutoff levels  $f$  for which the prescribed dimension density  $d$  is recovered by the respective dimension estimates. White areas correspond to parameters where  $\delta_{\text{LVD}}$  either could not be computed (very large  $f$ ) or gave artificially high values  $>1$  (very small  $f$ ).

of the data, we calculate this measure to understand temporal variations in the strength of correlation within the paleoclimate records. Significant changes of  $\delta_{\text{LVD}}$  occur when sediments with contributions from below 400 m b.s.f. (meters beyond sea floor) maximum depth or above 130 m b.s.f. minimum depth is considered (see Fig. 10). The interval between these two horizons covers a relatively small time window between about 24.3 and 23.8 Myr BP, which includes the Oligocene-Miocene transition (OMT) that was explained in Sect. 2. The qualitative changes recorded in the trace element data are associated with a change of the provenance of the material (Krissek and Kyle, 2000) and an enhanced vari-

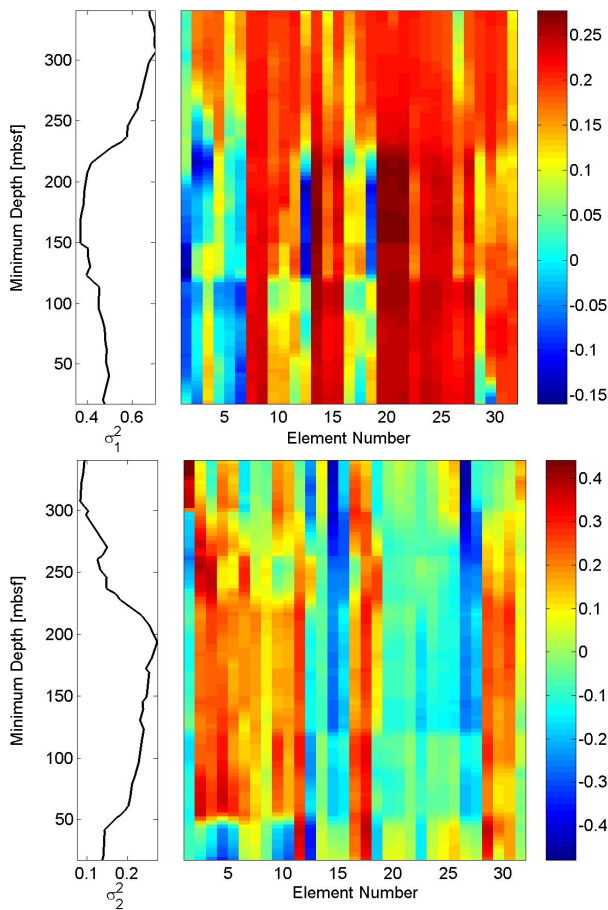


**Fig. 10.** Sum of the first two eigenvalues  $\sigma_1^2 + \sigma_2^2$  (left panels), and the corresponding LVD dimension density  $\delta_{\text{LVD}}$  (right panels) for the normalised trace element abundances from the CRP-2A core. Sliding windows of  $M=20$  (black),  $M=30$  (dark gray) and  $M=40$  (light gray) points in time are considered. The results are presented as a function of the minimum (upper panels) and maximum (lower panels) depth (m b.s.f.) of the sediment layer associated to the respective data subsets. Vertical dotted lines indicate the major transitions recorded in the data.

ability of the sedimentation. The two effects together result in a decrease of the interrelationships between the variability of different trace elements and, consequently, an increase of the dimension of the record.

Although a significant number of components are required to explain a certain fraction of total variance, the temporal variation of the first two leading eigenvalues already shows a pattern similar to the LVD dimension density<sup>2</sup>. For

<sup>2</sup>One can easily show that the corresponding variations are significant by considering the results for different values of  $M$  (cf. Fig. 10). To study the robustness of the measure against small perturbations of the data, we considered small windows with only  $M=20$  points in time, substituted randomly chosen measurement



**Fig. 11.** The first two eigenvalues  $\sigma_i^2$  ( $i=1, 2$ ) (left panels) and their associated eigenvectors (color-coded representation in right panels) for the normalised trace element abundances from the CRP-2A core as a function of the minimum core depth for sliding windows of  $M=20$  points in time.

time windows containing  $M=20$  points, Fig. 11 shows these eigenvalues and the corresponding eigenvectors. One observes that the behaviour of the first two eigenvalues is completely different. For the older part of the record, the first eigenmode clearly dominates the record, whereas the second one becomes increasingly important when considering data resulting from the time interval associated with the OMT. The corresponding changes in the first eigenvector are mainly reflected by the components associated to the trace elements scandium (element number 2), vanadium (3), strontium (12), and barium (18). The last two are also the main recorders of the increase of the first eigenvalue when younger sediment ( $\leq 23.8$  Myr BP) from above 130 m b.s.f. is considered. The onset of the increase of the second eigenvalue is reflected by the eigenvector component associated with sulfur (element

number 1), whereas a number of other components start to change later. These results apparently indicate that the climate change associated with the OMT is particularly pronounced in three elements of the record (S, Sr, Ba), which are also the trace elements with the highest absolute abundances in the record. This result is particularly remarkable as all component time series were standardised to unit variance before our analysis.

Apart from the climate change associated with the OMT, the eigenvector analysis seems to indicate further transitions in the climate system. For example, a qualitative change of the second eigenvector in the youngest part of the sediment is found. To gain a deeper insight into the corresponding age interval, one has to consider data from other sources, e.g., the CRP-1 and CIROS-1 cores. For these locations, however, there are no comparable records of trace element abundances available.

The grain-size distributions that we use as a complementary source of information belong to the class of compositional data which require a specific statistical treatment. In general, for an arbitrary multivariate data set, a transformation dividing the original data by their respective sum at every point in time leads to a set of compositional data. This situation is present in the case of grain-size distributions: As there are no absolute, but rather relative values, the statistically relevant quantities are no longer the component data themselves, but appropriate ratios thereof, as these ratios are invariant under the respective transformation.

Aitchinson (1986) has demonstrated that there are three equivalent ways of considering either pairwise or centred ratios within a compositional vector. Among these, for a data vector  $(x_1, \dots, x_N)$ , the  $N$  centred ratios are defined as  $x_i^* = x_i / g(x_1, \dots, x_N)$ , where  $g(x_1, \dots, x_N)$  is the geometric mean of the vector. We will consider these centred ratios as they do not give particular weight to any fixed component of the original data set. Typically, one uses the corresponding log-ratio transformed data instead of the centred ratios themselves (i.e.,  $\log x_i^*$ ). Because zero “counts” occur frequently in typical grain size histograms, the consideration of logarithms leads to numerical instabilities, and thus, we use the original data.

Considering the results displayed in Fig. 12, one firstly observes a much more detailed variability pattern when compared to the trace element data discussed above, which is an effect of the higher total number of observations. Secondly, the non-transformed data give a more diffuse pattern compared to the transformed ones, which underlines the necessity of a transformation to obtain statistically meaningful results on compositional data. Thirdly, the recorded transitions in the climate system associated with the OMT are consistent with the results from the trace element record, but are better resolved due to the larger number of data. In particular, one observes that the pattern corresponding to the Oligocene is reflected by a strong successive decrease of the LVD dimension, which is followed by an increase when sediment from

horizons by stochastic data, and computed  $\delta_{LVD}$  for the perturbed data set (see Donner and Witt, 2006).

between 350 and 400 m b.s.f. is considered. We attribute the corresponding layer to the onset of the OMT. Furthermore, the decrease of dimension above 130 m b.s.f. is resolved as in the case of the trace elements, which probably determines the end of the transition with a full development of the Antarctic circumpolar current and a resulting provenance change.

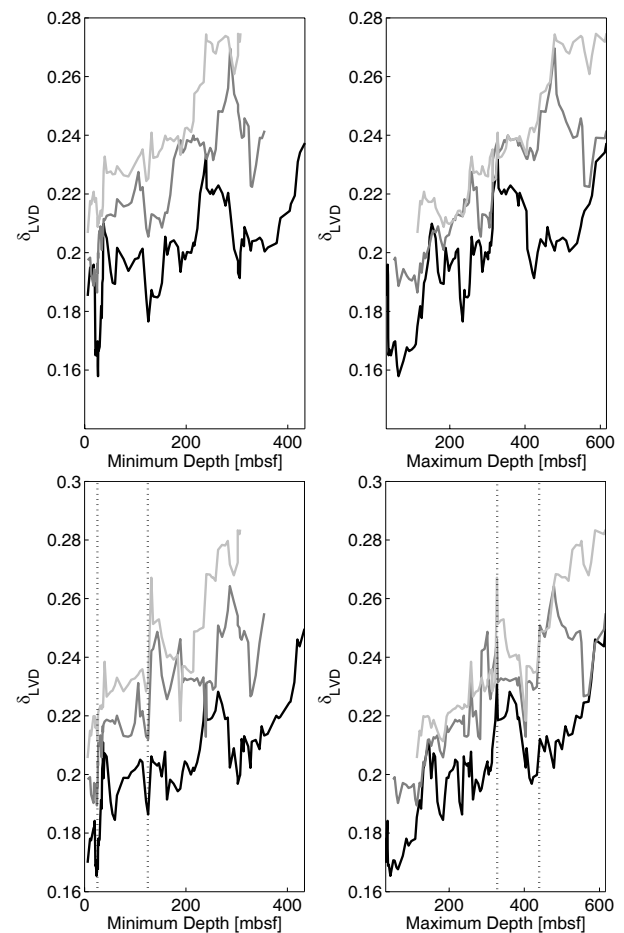
A further decrease of the complexity of interrelationships is observed when sediment from above 50 m b.s.f. is considered. This confirms the findings from the geochemical data and proves their actual relevance. The age associated to the corresponding layer is about 20.3 to 20.4 Myr BP. In another study from Prydz Bay (ODP Site 1165) (Williams and Handwerger, 2005), this age interval was found to correspond to the probably most pronounced layer of ice-rafted debris during the Early Miocene, which indicates that the transition found in the Cape Roberts data is probably related to a major deglaciation event on the Antarctic continent.

## 6 Conclusions

We have introduced a novel method to quantify the complexity of interrelationships of the components from short and noisy multivariate data sets by considering the scaling of component variances gained by an appropriate statistical decomposition of the record. In particular, we have shown that the LVD dimension density  $\delta_{LVD}$ , defined as the exponential decay rate of the remaining variances, serves as a suitable estimate for the fractal dimension of certain model systems and qualitatively detects even weak changes in the correlation structure of the data. Our approach allows one to determine an “optimum” cutoff for dimension reduction because it is related to the number of significant components of a multivariate record. Also, the method of KLD-based dimension estimates yields additional information about the dynamics and the temporal variations of the recorded system.

In this paper, we have characterised multivariate records of chemical and sedimentological properties of marine sediments collected off the shore of the East Antarctic coast in terms of  $\delta_{LVD}$ . In particular, we found evidence that the major climate change associated with the Oligocene-Miocene transition is reflected by significant changes of the interrelationships between the contents of different trace elements in the sediment. A corresponding eigenvector analysis indicated that these changes are mainly associated to the major trace elements sulfur, strontium, and barium. The observed variability pattern is even more pronounced when studying grain-size distributions from the same location. The grain-size distributions resolve another climatic event at about 20.3 Myr BP which is probably related to a major deglaciation of the ice shield on the Antarctic continent.

The eigenvalues of the covariance matrix of the data (computed for small sliding windows in time) yield important information about changes of the correlations between the component time series of the record.  $\delta_{LVD}$  is an aggregated



**Fig. 12.**  $\delta_{LVD}$  computed for grain-size distributions from the CRP-2/2A core without (upper panel) and with (lower panel) a transformation of the data to centred ratios. The results are displayed as a function of both, the minimum (left) and maximum (right) depth of the sediment layer considered for sliding windows of  $M=20$  (black),  $M=30$  (dark gray) and  $M=40$  (light gray) points in time. Vertical dotted lines correspond to common features of the three settings representing major climatic transitions recorded in the data.

parameter which qualitatively indicates long-term climate change. Instead of considering single observables (or heuristically defined parameters derived from), one can use  $\delta_{LVD}$  as a measure of the density of the extended system. Single observables, that are often easily interpretable in terms of meteorological parameters, respond to changes of environmental conditions in different ways, sometimes even with different time lags. So, it may be a promising approach to apply the proposed method to dynamically define new aggregated variables which are related to a single meteorological or hydrological parameter.

The method described in this paper is not restricted to palaeoclimatic studies of records from geological sequences. Under rather general conditions, it is applicable to a wide

range of potential problems in the geosciences, including the characterisation of seismic or hydro-meteorological time series. The proposed technique allows one to assess temporal variations of different spatio-temporally recorded parameters and predict their behaviour in the near future (i.e., in terms of forecasting of weather, floods, or earthquakes).

As a major limitation, our method implicitly requires time windows with a fixed number of data points. Even for an exactly known system with a prescribed correlation structure, the estimated dimension densities depend on the size of the data set, given in terms of both the number of component time series  $N$  and the number of points in time  $M$ .

Possible generalisations of our approach with respect to other methods of statistical decomposition can be considered by using nonlinear techniques instead of the standard linear KLD method. An application of the proposed technique to univariate time series is possible after an appropriate embedding. In particular, we propose the use of Singular System Analysis (SSA) (Vautard and Ghil, 1989) as a straightforward modification of the method presented here for multivariate data.

*Acknowledgements.* R. Donner acknowledges the Volkswagen foundation for funding (project no. I/78218), A. Witt is indebted to the Ministry of Science, Research and Culture of the Federative State Brandenburg (HWP program), and the European Community project “Extreme Events – Causes and Consequences” (NEST contract no. 012975-E2C2) for financial support.

Edited by: M. Thiel

Reviewed by: F. Demory and two other referees

## References

- Abarbanel, H. D. I.: Analysis of Observed Chaotic Data, Springer, New York, 1996.
- Aitchinson, J.: The Statistical Analysis of Compositional Data, Chapman and Hall, London, 1986.
- Barrett, P. and Anderson, J. B.: Grain-size analysis of samples from CRP-2/2A, Victoria Land Basin, Antarctica, *Terra Antarctica*, 7(3), 373–378, 2000.
- Barrett, P. and Anderson, J. B.: Frequency percent in each size class for grain size analysis of sediment core CRP-2/2A (Table 1), PANGAEA database, doi:10.1594/PANGAEA.133964, 2003.
- Battisti, D. S. and Sarachik, E. S.: Understanding and predicting ENSO, *Rev. Geophys.*, 33(S1), 1367–1376, 1999.
- Bauer, M., Heng, H., and Martienssen, W.: Characterization of Spatiotemporal Chaos from Time Series, *Phys. Rev. Lett.*, 71(4), 521–524, 1993.
- Bradley, R. S.: Paleoclimatology: Reconstructing Climates of the Quaternary, 2nd edition (Academic Press, San Diego), 1999.
- Buck, C. E. and Millard, A. R. (Eds.): Tools for Constructing Chronologies: crossing disciplinary boundaries, Springer, London, 2004.
- Bünner, M. J. and Hegger, R.: Estimation of Lyapunov spectra from space-time data, *Phys. Lett. A*, 258(1), 25–30, 1999.
- Craddock, J. M. and Flood, C. R.: Eigenvectors for representing the 500 mb geopotential surface over the Northern Hemisphere, *Quart. J. R. Meteorol. Soc.*, 95, 576–593, 1969.
- Diks, C.: Nonlinear time series analysis: methods and applications, World Scientific, Singapore, 1999.
- Donner, R. and Witt, A.: Temporary Dimensions of Multivariate Data from Paleoclimate Records – A Novel Measure for Dynamic Characterisation of Long-Term Climate Change, *Int. J. Bif. Chaos*, in press, 2006.
- Farmer, S. A.: An Investigation into the Results of Principal Component Analysis of Data derived from Random Numbers, *The Statistician*, 20(4), 63–72, 1971.
- Fischer, H., Kumke, T., Lohmann, G., Flöser, G., Miller, H., von Storch, H., and Negendank, J. F. W.: The Climate in Historical Times – Towards a Synthesis of Holocene Proxy Data and Climate Models, Springer, Berlin, 2004.
- Florindo, F., Harwood, D. M., and Wilson, G. S. (Eds.): Long-term changes in Southern high-latitude ice sheets and climate, the Cenozoic history, *Glob. Planet. Change*, 45(1–3), 1–263, 2005.
- Grassberger, P. and Procaccia, I.: Characterization of Strange Attractors, *Phys. Rev. Lett.*, 50, 346–349, 1983.
- Hay, W. W.: Tectonics and climate, *Geol. Rundsch.*, 85, 409–437, 1996.
- Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. (Eds.): The North Atlantic Oscillation: Climatic Significance and Environmental Impact, AGU, Washington, 2004.
- Hyvärinen, A., Karhunen, J., and Oja, E.: Independent Component Analysis, Wiley, New York, 2001.
- Jolliffe, I. T.: Principal Component Analysis, Springer, New York, 1986.
- Kantz, H. and Schreiber, T.: Nonlinear Time Series Analysis, Cambridge University Press, Cambridge, 1997.
- Kramer, M. A.: Nonlinear Principal Component Analysis Using Autoassociative Neural Networks, *Amer. Inst. Chem. Engin. Journal*, 37(2), 233–243, 1991.
- Krissek, L. A.: Element abundances, loss on ignition, total analyzed abundances, CIA values, and Al-oxide/Ti-oxide ratios of core CRP-2A, PANGAEA database, doi:10.1594/PANGAEA.144432, 2004.
- Krissek, L. A. and Kyle, P. R.: Geochemical indicators of weathering, Cenozoic palaeoclimates, and provenance from fine-grained sediments in CRP-2/2A, Terra Antarctica, 7(4/5), 589–597, 2000.
- Lean, J., Beer, J., and Bradley, R.: Reconstruction of solar irradiance since 1610: Implications for climate change, *Geophys. Res. Lett.*, 23(22), 3195–3198, 1995.
- McBride, E. F.: Mathematical Treatment of Size Distribution Data, in: Procedures in sedimentary petrology, edited by: Carver, R. E., Wiley, New York, 109–127, 1971.
- Meixner, M., Zoldi, S. M., Bose, S., and Schöll, E.: Karhunen-Loève local characterization of spatiotemporal chaos in a reaction-diffusion system, *Phys. Rev. E*, 61(2), 1382–1385, 2000.
- Mudelsee, M. and Statterger, K.: Plio-/Pleistocene Climate Modeling Based on Oxygen Isotope Time Series from Deep-Sea Sediment Cores: The Grassberger-Procaccia Algorithm and Chaotic Climate Systems, *Math. Geol.*, 26, 799–815, 1994.
- Naish, T. R., Woolfe, K. J., Barrett, P. J., et al.: Orbitally induced oscillations in the East Antarctic ice sheet at the Oligocene/Miocene boundary, *Nature*, 413, 719–723, 2001.

- Olbrich, E., Hegger, R., and Kantz, H.: Analysing local observations of weakly coupled maps, *Phys. Lett. A*, 244(6), 538–544, 1998.
- Paillard, D.: Glacial Cycles: Toward a New Paradigm, *Rev. Geophys.*, 39(3), 325–346, 2001.
- Politi, A. and Witt, A.: Fractal Dimension of Space-Time Chaos, *Phys. Rev. Lett.*, 82(15), 3034–3037, 1999.
- Preisendorfer, R. W.: *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, Amsterdam, 1988.
- Priestley, M. B.: *Non-linear and non-stationary time series analysis*, Academic Press, London, 1988.
- Probert-Jones, J. R.: Orthogonal Pattern (Eigenvector) Analysis of Random and Partly Random Fields, *Conference on Probability and Statistics in Atmospheric Sciences*, 3, 187–192, 1973.
- Raab, C. and Kurths, J.: Estimation of large-scale dimension densities, *Phys. Rev. E*, 64, 016216, 2001.
- Roweis, S. and Saul, L.: Nonlinear dimensionality reduction by locally linear embedding, *Science*, 290(5500), 2323–2326, 2000.
- Sachs, H. M., Webb III, T., and Clark, D. R.: Paleocological Transfer Functions, *Ann. Rev. Earth Planet. Sci.*, 5, 159–178, 1977.
- Telford, R. J., Heegard, E., and Birks, H. J. B.: All age-depth models are wrong: but how badly ?, *Quat. Sc. Rev.*, 23, 1–5, 2004.
- Telford, R. J. and Birks, H. J. B.: The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance, *Quat. Sc. Rev.*, 24, 2173–2179, 2005.
- Tenenbaum, J., de Silva, V., and Langford, J. C.: A Global Geometric Framework for Nonlinear Dimensionality Reduction, *Science*, 290(5500), 2319–2323, 2000.
- Tong, H.: *Non-linear time series: a dynamical system approach*, Clarendon Press, Oxford, 1990.
- Varela, H., Beta, C., Bonnefort, A., and Krischer, K.: Transitions to Electrochemical Turbulence, *Phys. Rev. Lett.*, 94, 174104, 2005.
- Vautard, R. and Ghil, M.: Singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series, *Physica D*, 35, 395–424, 1989.
- Williams, T. and Handwerger, D.: A high-resolution record of early Miocene Antarctic glacial history from ODP Site 1165, Prydz Bay, *Paleoceanography*, 20, PA2017, 2005.
- Zoldi, S. M. and Greenside, H. M.: Karhunen-Loève Decomposition of Extensive Chaos, *Phys. Rev. Lett.*, 78(9), 1687–1690, 1997.