Vrije Universiteit Brussel
Faculteit Toegepaste Wetenschappen
Departement ELEC
Pleinlaan 2, B – 1050 Brussels, Belgium

Identification of the Time Base in Environmental Archives

Fjo De Ridder

Promotoren:
Prof. dr. ir. Rik Pintelon
Prof. dr. Frank Dehairs

Proefschrift ingediend tot het behalen van de academische graad van doctor in de toegepaste wetenschappen.

December 2004

F. DEHAINS

Identification of the Time Base in Environmental Archives

Fjo De Ridder

Voorzitter:
Prof. dr. ir. *Philippe Lataire* (Vrije Universiteit Brussel, Belgium)

Vice-voorzitter:
Prof. dr. ir. *Jean Vereecken* (Vrije Universiteit Brussel, Belgium)

Secretaris:
Prof. dr. ir. *Johan Schoukens* (Vrije Universiteit Brussel, Belgium)

Jury:
Prof. dr. *Luc André* (Musée royal de l'Afrique centrale; Université Libre de Bruxelles, Belgium)
Prof. dr. *André Berger* (Université Catholique de Louvain, Belgium)
Dr. *Guy Munhoven* (Université de Liège, Belgium)
Dr. *Didier Paillard* (Centre d'Etudes de Saclay, Orme des Merisiers, France)
Prof. dr. ir. *Joos Vandewalle* (Katolieke Universiteit Leuven, Belgium)
Prof. dr. *Willy Baeyens* (Vrije Universiteit Brussel, Belgium)
Prof. dr. ir. *Alain Barel* (Vrije Universiteit Brussel, Belgium)

To Anouk

TABLE OF CONTENTS

Table of Contents

Table of Contents

Table of Contents

CHAPTER 1

PREFACE

This work is situated on the interface between paleo climatology and system identification, between proxy data and data models. We have tried to apply the mathematical framework developed in the engineering community to some particular questions raising the paleo-climatology community.

1.1 Where is the story set?

What is system identification and where can it be situated according to other sciences? To come to an answer, I would like to quote Ilya Prigogine, who saw his scientific work as a dialog with nature. When we started working on proxy records, which are data-records (read: numbers), the never-ending quest was to find out how much information was hidden in these records. What are the data actually telling? Soon it became clear the answer was strongly dependent on the assumptions made by the investigator. Assumptions, which can mathematically be expressed in models. It is at this stage that system identification, parameter estimation, signal processing, etc.... comes into the game, because these tools are necessary to find out which model matches on a given record, what is the complexity of these raw data, which assumption are supported by the experiments and which are not, what is noise and which variation tells a story about nature and which are artefacts, belonging to the instrumental set-up? To summarize, system identification can be used as a tool to automate and objectify the interpretation of data. Expressed in I. Prigogine's words, System Identification outlines the rules in our negotiation with nature: it provides the scientific story that we want and the price is paid in assumptions. The stronger the set of assumptions is, the less secrets we can unravel from the data and vise versa. Practically, System Identification can be decomposed into several steps:

1. how are the measurements collected?

2. one or several models have to be proposed;
3. the model parameters have to be tuned to the data in an optimization step; and
4. a model selection criterion is used to decide which model matched best.

For the latter it is important to make a separation between two cases: if the uncertainties on the measurements are known, we would like to find a model which is able to describe all variation, which can not be related to measurement noise. Unfortunately, most proxy records are less complete and the uncertainty bounds on the measured records is unknown or hardly known. Consequently, it becomes impossible to select a 'good' model (describing all significant variation) and the best we can do is selecting the 'best' model within a given model set. So, to conclude, System Identification is an abstract tool, positioned on the interface between measurements and models, helping scientists in extracting significant information from noisy observations. As Max Planck once said: "Experiments are the only means of knowledge at our disposal. The rest is poetry, imagination." It is time for that imagination to unfold.

1.2 Summary

The second chapter handles the measurement. These are often collected with a Laser Ablation Inductively Coupled Mass Spectrometer (LA-ICP-MS) and the first problem we encountered was the calibration of this instrument. It is used to measure trace elements with an incredibly high spatial resolution. Of course, it has also some major disadvantages. One of these is that the laser, which is used to ablate the solid substrate, has a variable intensity. In addition several other internal instrumental parameters seem to vary over time in an unpredictable manner. Therefore, people use internal standards for more than twenty years. Such internal standards are elements whose concentration does not vary over the sample. Consequently, all variation detected in these internal standards must be due to artificial variation in the instrument. Because the drift of the instrument is reflected in the measurement of the internal standard, all other concentrations can be calibrated according to this internal standard. This simple idea seems to work quite well. Unfortunately, the calibration algorithm does not take into account measurement noise in the signal of the internal standards. This is exactly the improvement that we made: imagine two internal standards, each consisting of the same drift pattern, and disturbed by a different noise pattern. A better approximation of the true drift pattern can be reached by the arithmetic mean of both internal standards. Herein the drift remains and the noise cancels out. This idea is further refined, so that the uncertainty on the internal standards can be taken into account as well, and an a priori and an a posteriori verification of the model used is given. Further, the myth of mass dependent internal standards is ruled out and finally a real world example is processed. An additional important property of the proposed calibration is that an internal quality check is performed, which warns the investigator if some artefacts or problems occurred during the measurement.

The remaining part of this work is devoted to the reconstruction of time series. One of the major problems with data-processing of proxy records (e.g. stable isotope ratios of oxygen or carbon, or trace elements in shells, sponges, corals, sediment cores, etc....) is the dating of individual observations. All these proxy records are measured as function of a distance, while generally the time series are desired.

Preface

Due to variations and differences in accretion rate, each record has its unique distance series, which cannot be compared with other records or models. Therefore, distance series are transformed into time series. However, this is only possible if additional information about the accretion rate is available. Unfortunately, this is mostly not the case and thus additional assumptions are necessary.

Such assumption can be made about the signal and formally written down in the signal model. In addition, the concept of a time base distortion is introduced. As will be shown, this enables us to estimate variations in accretion rate. For this, we started from a previously estimated time base (if this is unknown, we initialize the time base assuming a constant accretion rate). Next, we allow this base to be distorted due to nonlinear accretion rates or hiatuses and we will show how such a distortion can be estimated.

In the third chapter, we have illustrated how the time series can be reconstructed, assuming a periodic proxy-record. Therefore, the concept of time base distortions is introduced. First, we propose a model for variations in accretion rate. Next, we illustrate how this accretion rate can be identified and estimated in the Fourier spectrum of the proxy. This approach is compared with the widely used anchor point method in two manners: the assumptions made are compared and next the sensitivity of both methods w.r.t. stochastic noise is illustrated on a simulation. Three case studies are incorporated to validate the method and to illustrate its usefulness. First the Mg-proxy in a shell, i.e. a *Isognomon ephippium,* is used. Herein, the correlation with temperature is used to compare this method with the anchor point method. Next, the stable oxygen isotopes measured in three *Saxidomus giganteus* clams are used. The three clams were sampled from the same site and the correlation between these time series is used as validation. Finally, the stable oxygen isotopes measured in a coral are processed. Herein the signal-to-noise ratio was used as criterion. In these three different case studies, the time base distortion method outperformed.

However, two disadvantages can be pointed out: the growth rate is biased at the borders, due to the Gibbs phenomenon, and the proposed method is not objective, because the user has to tune some parameters. These problems are circumvented in two technical chapters.

In chapter four, a technical solution is proposed to avoid the Gibbs phenomenon. This approach could only be implemented if parametric models for the signal and for the time base were used. First, these were implemented. We have used periodic models to reconstruct the growth or accretion rate. Therefore, we implicitly assume that the model is continuous at the borders of the measurement frame, while the indirectly measured growth rate is not. This contradiction causes an error (Gibbs phenomenon or ringing) To overcome this, we have proposed to use a model frame, which is larger than the actual measurement frame. The extra space that is created outside the measurement frame is used to decrease the Gibbs phenomenon where both frames overlap. Next, the conceptual details are discussed, a simulation is used to validate the approach, a method to determine the size of the model frame is proposed and finally a real world example is processed. Although the problem and its solution are technical, I like this part of my work particularly, because it demonstrated a fast and workable solution for a particular problem. This is precisely what engineering is all about.

Another disadvantage of the proposed method is that the variations in growth rate are estimated employing a spectral window. So far, the size of this window is chosen by the investigator. Consequently, two different users can come up with different conclusions from the same data, if they would have used different window sizes. Applied on parametric models, this problem can be translated to an estimation of the model complexity: how many harmonics are present in the measured proxy and how complex is the time base? In this fifth chapter, we have refined model selection criteria, which can answer these questions. The particular problem to solve was that proxy records are often fairly short and the number of parameters is relatively large. Therefore, classical model selection criteria have the tendency to select overly complex models. We have proposed adapted model selection criteria, which seems to select good models, even if the data records are short. First, we show were and why the classical rules fail, next these rules are refined. The new rules are validated and compared with the old rules on simulated data. Therefore, we have used three completely different models: first we have used the type of models which are used in this work (signal modeling example), next, a system identification model is used and thirdly, a noise modeling example is processed. Finally a real world example is processed.

Preface

This chapter is the key chapter is this work, because it enables to determine in a objective manner, how much information is actually carried by the measurements. Without model selection, any model can be matched on the measurements and potentially wrong conclusions can be drawn. When an automated model selection rule is used, the investigator can be statistically sure that the best model within the model set is selected, if no noise information is used, or that a good model is used, if model information is available. Therefore, it is our opinion that this kind of criteria should be used much more often in environmental sciences, in order to validate the ideas of the user, based on judgement of the data themselves. Personally, it was a great achievement to give a feedback to the identification community, starting from data derived in the paleo climate community, where it was our first goal to go in the opposite direction.

In the sixth chapter, an overview is given of the numerical issues involved in the implementation of the model. When the two previous chapters were implemented, several smaller problems raised. The optimization algorithm seemed to converge sometimes toward local minima. The best solution for this problem is evidently the implementation of global optimization methods, but nowadays these methods would still need geological time scales to converge. Therefore, a procedure to avoid local minima is proposed: models with a different complexity were used to generate initial values for a particulate complexity. Another problem was that some data can consist of e.g. a signal model with two harmonics i.e. the first and third, while the second is absent. The model, as proposed so far, would consists of three harmonics and consequently the model selection could be biased. An algorithm to reduce the model complexity is implemented, which removes redundant parameters Finally, a potential problem of time inversion is briefly discussed. Three examples which were already treated in previous chapters are processed again.

In the seventh chapter, this method is compared with those discussed in literature. Therefore, we have subdivided the time series reconstruction methods in to two classes: those who assume that the profile of the time series is known in advance and those who use other assumptions. For the first class, the anchor point method, the correlation maximization and the method proposed by Martinson *et al.* [1982] are discussed. For the second class, we have discussed a time domain method, our non-parametric approach (Chapter three) and the parametric approach (Chapter six). For each method, the basic ideas, the assumptions, tuning

parameters, limitations, sensitivity to stochastic noise and to model errors are discussed. For the latter, the presence of overtones in an harmonic signal, amplitude modulations, trends and hiatuses are discussed.

So far, we have limited ourselves to periodic signals. Of course, more complicated signals, with trends or other periodic behavior are often of much more interest. In order to illustrate in which direction these models can be expanded, a eighth chapter is added, where multiple non-harmonically related periodic signals are discussed. A first example is El Niño, which appears with a frequency of approximately seven years and which can appear additionally to the yearly variations in some proxies. We received this record from Miriam Pfeiffer (Leibniz Institut fur Meereswissenschaften, Germany), with the question if an El Niño was present and if so, how it could be separated from the large annual signal. Therefore, the algorithm was adapted so that multiple non-harmonically related signals can be identified.
Another example is the 40 kyear periodicity, appearing in addition to the 100 kyear cycle in the temperature profile, derived from the deuterium record in the Vostok ice core record. Petit *et al.* [1999] claimed that this periodicity is present in this record, while Broersen *et al.* [2003] claimed that even the 100 kyear cycle is not significant. The same algorithm is applied and it appeared that the 40 kyear cycle is insignificant, while the 100 kyear cycle could be identified. To illustrate this, we have optimized a time base starting from the GT4 time base. However, it is important to notice here that this approach cannot distinguish a frequency modulation from a time base distortion. Therefore, it is especially well possible that the GT4 time base was indeed correct, but that the frequency of the 100 kyear cyclicity was modulated. This is a standard example of an identification problem. However, in none of both scenarios the 40 kyear cyclicity is significant. Once more, this identification problem illustrates the power and the needs of identification tools in environmental sciences.

Brussels, September 2004

Fjo De Ridder

REFERENCES

[1] Martinson D.G., Menke W. and Stoffa P., 1982. An inverse approach to signal correlation. Journal of geophysical research 87(B6), 4807-4818.

[2] Petit J., Jouzel J., Raynaud D., Barkov N., Basile I., Bender M., Chapellaz J., Davis J., Delaygue G., Demotte M., Kotlyakov V., Legrand M., Lipenkov V., Lorius C., Pépin L., Ritz C., Saltzman E. and Stievenard M., 1999. 420,000 years of climate and atmospheric history revealed by the Vostok deep Antarctic ice core, Nature, 399, 429-436.

[3] Broersen P.M.T., de Waele S., Bos R., 2003. Estimation of Autoregressive Spectra with randomly missing data. IMTC 2003 - Instrumentation and Measurement Technology Conference, Vail, CO, USA, 20-22 May 2003, pp. 1154-1159.

1.3 Acknowledgements

First of all, the ELEC staff created an environment and facilitates in which these ideas could grow. Therefore, I would like to thank

Rik and Johan; Once upon a time, you told me that your only task was to create an environment in which I could work happily. Well, I shall always remember my ELEC years as wonderful years.

Yves; special thanks for encouraging me, especially after my first paper review.

Alain, Ann, Jean-Pierre, Jenny, Wilfried, Wim; thanks for all the problems I had not noticed. Because of you I have had much more time and energy to put in my scientific work.

Where ELEC provided the mathematical and logistic framework, the problems statements and data came from the CALMARS group. I really enjoyed the collaboration and discussions with all of you. I would especially like to thank

Frank; your scientific curiosity and enthusiasm is admirable. Although our approach toward scientific problems is often orthogonal, I hope that this diversity is our strength and that this symbiose will be continued in the future.

Luc, Philippe and Willy: our cooperation was more fragmented, but without each of you, this work would not have reached its current level.

I enjoyed to work on smaller projects with Anouk De Brauwere, Anouk Verheyden, Bert Desmedt, David Gillikin, Guillemette Joly, Nele Schmitz and Miriam Pfeiffer.

I would like to thank the members from the jury, whose thoughtful comments led to significant improvements in this manuscript.

Anouk, all this is for you.

CHAPTER 2

AN IMPROVED MULTIPLE INTERNAL STANDARD NORMALIZATION FOR DRIFT IN LASER ABLATION INDUCTIVELY COUPLED MASS SPECTROMETRY MEASUREMENTS

Abstract - One of the primary factors limiting precision in LA-ICP-MS measurements is the instrument fluctuation. To overcome this limitation one usually employs a normalization based on a single internal standard. This chapter reports a new approach for the normalization of LA-ICP-MS measurements, based on multiple internal standards. The novelty of the approach is that it (i) takes into account the measurement uncertainties, (ii) generates uncertainty bounds on the normalized measurements and (iii) allows to check and dismiss the mass dependency of the normalization procedure. As a result, it is shown that the improvement in normalization factor increases proportionally to the square root of the number of internal standards (e.g. the use of two internal standard leads to a decrease of the standard deviation by 30%, three to 42% and four up to 50% compared with the classical normalization). This is illustrated on the NIST 610 standard with a 266 nm UV LA-ICP-MS. An extra advantage of multiple internal standards is that it allows the detection of artifacts in the measurements (internal quality check).

2.1 Context

This chapter covers the first year of my research. All aspects of the identification approach are already present here: gathering measurements, the proposition of models, automated model selection, parameter optimization and feedback to the investigator.

Summarized, internal standardization, normalization of calibration of the measurement instrument are necessary, because the measurement process is unstable, due to fluctuating instrumental parameters, like laser power, gas flux, etc.... which all influence the measurement signal. In order to estimate this drift of the instrument, internal standards are used, whose signals should be constant over the total measurement. The drift causes fluctuations in these internal standard signals. These fluctuations can consequently be used to make a correction. Usually, this is done by employing only one internal standard.

As is already pointed out in the introduction, the basic idea, proposed here, is very simple: measure the internal standard several times and average out the noise. This results in a remaining drift pattern, which is estimated much more precisely.

The first assignment, which we have examined, is the use of multiple internal standards. First, by mass-dependent calibrations, later by simple averaging. If the individual uncertainties are unknown, the averaging of multiple internal standards can formally be written in a least squares estimator. However, some internal standards can be measured with a better precision than others. Instead of using only the best internal standard, we have used all available information, extracted from all internal standards. This resulted in a weighted least squares estimator. The model used in this estimator is based on a singular value decomposition. This shows that besides the drift pattern, only stochastic noise differentiates the internal standards.

2.2 Introduction and main goals

Maximal accuracy and precision with minimal effort and cost are the major goals in the development of any analytical procedure. Inductively coupled plasma mass spectrometry (ICP-MS) seems to meet both these goals and is used as a powerful analytical technique in geological, biological, environmental, metallurgical and nuclear applications. For the ablation of solids, a laser is often employed. The ability of focused laser radiation to volatilize almost all materials has made this so-called laser ablation (LA)-ICP-MS technique a powerful and one of the most flexible analytical techniques for direct isotopic analysis of solids, without any sample preparation prior to analysis [Gunther *et al.*, 2000; Gray, 1985; Jeffries *et al.*, 1998].

The major sources of errors that limit precision and accuracy are the linear dynamic range, drift and matrix effects [Cheatham *et al.*, 1993]. The linear dynamic range is limited by the hardware of the apparatus and can be improved by a dead time correction [Vanhaecke *et al.*, 1998]. Internal and external standards are used to improve precision and accuracy of the measurements. Overall, the level of precision that can be achieved lies roughly between 5 and 25% [Gunther *et al.*, 2000].

The drift caused by instrument fluctuations mainly limits this precision. In LA-ICP-MS measurements, drift is an artificial and common pattern appearing in the raw data of all nuclides, even in those of the internal standards. This fact is exploited to extract the drift pattern from the signal of one single internal standard [Jackson *et al.*, 1992; VG Elemental Manual]. This approach assumes that the drift on all elements is the same and that the disturbing noise on the internal standard is negligible. Alternatively, a more refined normalization takes also the mass of the elements into account [VG Elemental Manual]. Again, the disturbing noise on the internal standards is neglected.

2.2.1 Remarks on these classical normalization methods and the outline of this chapter

Four problems can be pointed out:

1. Neither of these two normalization strategies take the measurement noise into account. If it is demanded that the variation of the internal standard is zero, without separation of this noise, it will be propagated into the signal of the normalized elements. So it would be better if we could, maybe partially, separate the noise and the drift pattern starting from these counts. If so, we could estimate the drift more precisely.
2. Some elements can be measured more precisely than others, but these differences in precision are not taken into account. Consequently, a weighing by their precision would emphasize the more precise measurements, which would improve the estimation of the normalization factor.
3. If more internal standards are available and the raw measurements are normalized with the single internal standard method, then the different internal standards lead to different results (vide infra). So, when these normalizations are used, without uncertainty bounds, this could easily lead to wrong interpretations of the results. For that reason, it is not only important to improve the precision, but also to quantify it.
4. It is our experience that the fluctuations caused by drift in the laser ablation are much larger than the difference in chemical behavior between different elements. Therefore, the mass and/or element dependencies in the normalization are all neglected in the rest of this article.

Formally, the intensity (counts per second) for the nth ($n \in \{1, \ldots, N\}$) measurement of internal standard m ($m \in \{1, \ldots, M\}$) is modeled as

$$I(n, m) = \alpha(m)p(n) + e(n, m) \tag{2-1}$$

or

$$I = p\alpha^T + e \tag{2-2}$$

where the latter is written in matrix notation, with $\alpha = [\alpha(1), \ldots, \alpha(M)]^T$ is a scale factor (counts per second), which is a function of the specific internal standard and

$p = [p(1), \ldots, p(M)]^T$ is the drift pattern (no units, normalized to one, i.e. $|p|^2 = 1$ or to the first measurement, i.e. $p(1) = 1$), which is a function of the measurement number, but not of the internal standards (T is the matrix transpose [Golub and Van Loan, 1990]) and e is the noise on the measurements. This new normalization approach focuses on finding the best estimation of the drift pattern from multiple internal standards in the presence of noise. In other words: when more internal standards are used, it is possible to calculate a most probable drift pattern, i.e. the pattern which explains most of the variation in all internal standards. Therefore, it takes the respective precision of each internal standard into account and generates uncertainty bounds. Furthermore, an a priori and an a posteriori verification for the simplification of the model in remark 4 is given. This normalization model is general and can thus be applied to any ICP-MS instrument and to any material.

First, improvement and conflict in the classical normalization are shown. Next, the general mathematical framework of this normalization method is explained. Further, this framework is experimentally tested on a NIST 610 silicate standard and on a sclerosponge (coral sponge, consisting of biological aragonite) with a 266nm UV LA-ICP-MS set-up. Finally, the results of both, the classical and the new approach are compared and some remarks are added:

1. it is established that this new approach is a generalization of the classical normalization; and
2. the internal quality control is discussed.

All the mathematics used in this article can be found in [Pintelon and Schoukens, 2001] and some are explicitly given in the appendices 2.A, 2.B and 2.C.

2.3 Experimental set-up and representation of the results

2.3.1 Sample

Because LA-ICP-MS is a spatially sensitive technique, a homogeneous material is required to demonstrate and evaluate the precision of this new normalization approach. A NIST 610 silicate glass reference material was used as sample, because it is a well-described, homogeneous material [Chen *et al.*, 1997; Chen *et al.*, 2000]. No sample preparation prior to the analysis was done. Ten different elements were measured, i.e. ^{11}B, ^{25}Mg, ^{43}Ca, ^{59}Co, ^{65}Cu, ^{86}Sr, ^{90}Zr, ^{135}Ba, ^{138}Ce and ^{207}Pb. These ten were chosen, because they do not suffer too much from interference and fractionation[1], (except Co, Cu and Pb) and their masses are spread over the whole range of detectable masses.

2.3.2 Instrumentation

The LA-ICP-MS analyses were made with a frequency-quadrupled ultraviolet-wavelength (266 nm) ND-YAG (neodymium-doped yttrium aluminium garnet) Fisons-VG microprobe and a Fison-VG Plasma Quad II+ mass spectrometer. The laser probe was operating in the Q-switched mode, the flash-lamp voltage was 630-650 V, the laser output power was 3-4 mJ/ pulse and the repetition rate was 10Hz. The carrier gas flew at a rate of 1.06 l/minute, the auxiliary gas at a rate of 1.4 l/minute and the cooling gas at a rate of 13.5 l/minute. Data were acquired during 145 s using Time Resolved Analysis. Hence, data can be collected from a selected range of elements and be viewed in real time. The laser ablation was stopped after 75 s to clean the ICP-MS instrument and to measure the blanks after each ablation process. A series of 75 craters were made, randomly chosen over the NIST 610 standard's surface. Each ablation crater obtained had an average diameter of 100 μm. The data were treated off-line employing house written software.

1. The relative fractionation index with respect to ^{43}Ca as defined in reference [Fryer *et al.*, 1995] was respectively: 0.87, 1.04, 1.00, 1.19, 1.24, 0.97, 1.05, 1.03, 1.03, 1.26 for the first ablation measurement.

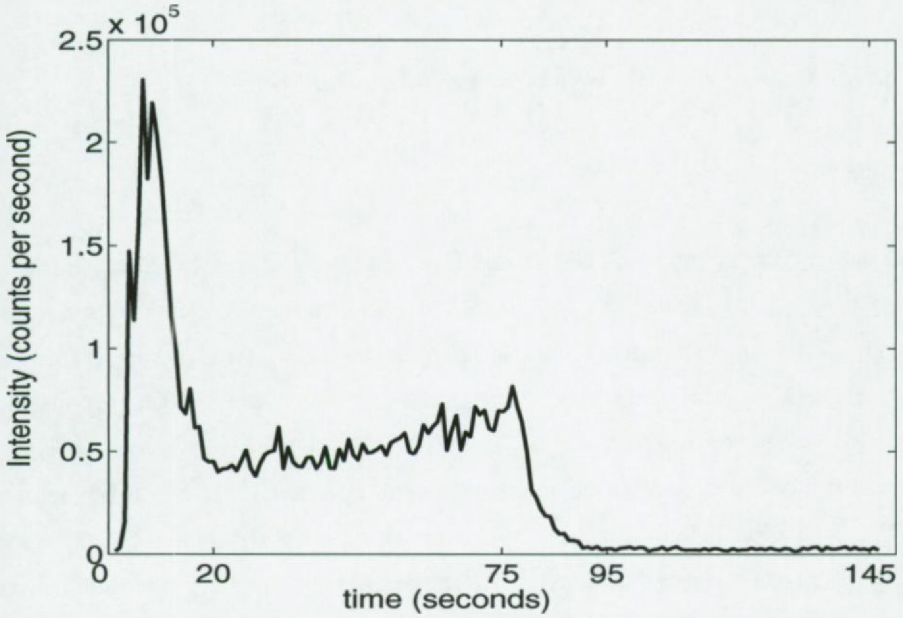

Figure 1: A Time Resolved Analysis of Ca. A mean signal and blank with the corresponding uncertainty bounds are abstracted from these plots for all measurements of all elements.

2.3.3 Data acquisition

Figure 1 shows an example of the Time Resolved Analysis of Ca during one ablation process. This type of plot is typical for a laser ablation system and the measurements from other elements or other ablations are described by the same pattern. The Time Resolved Analyses profile can be separated into 4 parts:

1. In the beginning of the ablation process a large and irregular peak is detected (0-20 s), indicating that a large volume of material is initially ablated at the surface of the sample.
2. Next a nearly flat time signal appears, which is used for measuring the concentrations (20-75s).
3. Then, after 75 s, the laser was stopped, while the ICP-MS continued measuring. During this time the ICP-MS was flushed and cleaned (75-95 s), which can be seen in the exponential decrease of the signal.
4. After 20 s the signal became stable again, so a good estimation of the blank measurements

was possible (95-145 s).

Finally, after 145 s, the apparatus was ready for a new laser shot.

2.3.4 Data reduction

The data collected from these Time Resolved Analyses were reduced, by calculating not only a mean, as is usually done, but also a standard deviation of the signal during ablation, and a mean and standard deviation of the blank signals. It was assumed that within one ablation all trace elements' signals were steady, so one ablation of one element can be represented as 66 repeated measurements in the interval 20 to 75 s. The average signal is a measure for the particular concentration, while the uncertainty on this mean number of counts is proportional to the standard deviation over these 66 measurements (68% uncertainty bound), denoted $\sigma_{ablation}$. From the signals measured between 95 and 145 s the same procedure delivers a mean blank and an estimation of the noise on this blank, σ_{blank}.

Subsequently, a blank correction is carried through by subtracting the mean blank from the mean number of counts at each ablation position and an estimation of the noise on this corrected mean number of counts, σ_I, is given by:

$$\sigma_I^2(n, m) = \sigma^2_{ablation}(n, m) + \sigma^2_{blank}(n, m) \qquad (2\text{-}3)$$

These corrected data are stored in two matrices, referred to as "the raw data": one with the means, $I(n, m)$, and one with the standard deviations, $\sigma_I(n, m)$. In these matrices each column describes a particular element, $m \in \{1, \ldots, P\}$, while each row describes a particular ablation measurement, $n \in \{1, \ldots, N\}$.

2.4 Least-squares approach of the normalization

Let us now turn to the formulation of the new normalization. When more internal standards are combined to estimate the drift pattern, this can be done in many different ways. We have chosen for a cost function-based approach, because a framework exists which enables us to handle the noise and estimate the drift pattern employing statistical considerations [Pintelon and Schoukens, 2001].

Briefly, the drift pattern is extracted from multiple internal standards at once. This is done by modelling the raw data of the internal standards by a bilinear model, as is done in equations (2-1) or (2-2), i.e. the raw data are modelled as a combination of a drift pattern, called p, and a scale factor, α. The pattern describes the drift of the apparatus as a function of the measurement number and not of the specific internal standard, while the scale is a relative measure of the amount of counts for the internal standards, which is not a function of the measurement number. Next, this model is fitted to the raw data by minimization of a weighted least squares cost function, accounting for the quality of the raw data. This weighting will put more emphasis on more precise measurements with respect to the others.

2.4.1 An a priori motivation for the bilinear model

The problems solved in this paragraph are:

1. Is a bilinear model a good choice?
2. What is the lower bound of precision with this model?
3. Is it useful to take a mass-dependency or element-dependency of the drift pattern into the normalization model?

These three questions are answered at once by a singular value decomposition of a scaled raw-data matrix. Those who are not familiar with singular value decompositions can find an introduction in Golub and Van Loan [1990] or Jennings [1977].

To start, we introduce two matrices:

1. the $N \times M$-matrix with the raw data of the internal standards, I (N: the number of measurements and M the number of internal standards – as defined in equation (2-1) or (2-2)); and

2. the column-covariance matrix, C_I, which is calculated with the uncertainty information of the raw data, as follows:

$$\begin{aligned} C_I &= E\{(I - E\{I\})^T (I - E\{I\})\} \\ &= \text{diag}\left(\sum_{n=1}^{N} \sigma_I^2(n,1), \ldots, \sum_{n=1}^{N} \sigma_I^2(n,M) \right) \end{aligned} \tag{2-4}$$

where $\sigma_I^2(n,m) = \text{var}(e(n,m))$ with $e(n,m)$ defined in equations (2-1) and (2-2), $E\{\ \}$ is the expectation value, and $\text{diag}(a, b, \ldots, c)$ is a diagonal matrix with main diagonal elements $a, b, \ldots, c$. It is assumed that the noise is independent over n and m, i.e. from measurement to measurement and from element to element, so the off-diagonal elements are all zero. Additionally, we used the column-covariance matrix, thus the diagonal elements were filled with the respective sum of all variances over the 75 measurements. Next, we normalized the raw data matrix, I, by right division with the square root of the column-covariance matrix, $C_I^{1/2}$ (with $C_I^{T/2} C_I^{1/2} = C_I$)[1]:

$$I_{\text{scaled}} = I C_I^{-1/2} \tag{2-5}$$

where I_{scaled} is the scaled data matrix. Because of this scaling all columns have now the same noise level [Rolain *et al.*, 1997]. Note that each measurement is not scaled by its respective uncertainty, but by the total standard deviation (uncertainty) of that column. The singular values of I_{scaled} allow one to immediately estimate the significance of each singular value, i.e. singular values equal to 1 contain only stochastic variations, while larger values contain significant information compared to the noise level (see appendix 2.A). Now we can tackle the 3 problems (1, 2 and 3) outlined in the beginning of this paragraph:

---

1. The calculation of a square root of a matrix can be found in [Golub and Van Loan, 1990].

2.4.1.1 Is a bi-linear model a good choice?

If so, all information of the raw data should be confined in the first singular value. Indeed, when we neglect all the other singular values, the raw data matrix would simplify to this bilinear model, as follows (see appendix 2.A, equation (2-26)):

$$I_{\text{scaled}} = U\Sigma V^T \cong \sigma_1 u_1 v_1^T$$
$$= \sigma_1 \hat{p}\hat{\beta} = \sigma_1 \begin{bmatrix} \hat{p}(1) \\ \dots \\ \hat{p}(N) \end{bmatrix} [\hat{\beta}(1), \dots, \hat{\beta}(M)] \tag{2-6}$$

where Σ is a diagonal matrix, containing the ordered singular values (from large to small), U and V are orthonormal matrices, σ_1 is the largest singular value, u_1 and v_1 are respectively the first column of U and the first column of V, $\hat{p}$ is the estimated drift pattern (no units), defined in equation (2-1) or (2-2), and $\hat{\beta} = \hat{\alpha} C_e^{-1/2}$ (no units). Note that u_1 is already a very good estimate of the pattern, $\hat{p}$, and v_1 is a linear function of the scale factor, α. May this assumption be made? Figure 2 shows a bar plot of the singular values. It can be seen that over 97% of the useful information is stored in the first singular value. Only a marginal influence (< 2%) of the second singular value and a negligible third singular value (< 0.5%) on the model can be found. The model error is given by the following equation (see appendix 2.A):

$$\text{Model errors} = \sqrt{\frac{\sum_{m=2}^{M} (\sigma_m^2 - 1)}{\sigma_1^2 - 1}} \approx 2.4\% \tag{2-7}$$

where $\sigma_m^2 - 1$ is the square of the mth singular value, corrected for the noise contribution. The magnitude of the model errors is 2.4%. So, we may conclude from these singular values that a bilinear model is a good choice to model the raw data. A physical explanation follows from Figure 3. Here the raw data for five elements (Mg, Ca, Cu, Sr, Ba) are shown. It seems that all elements have a common pattern in time. This pattern has several origins:

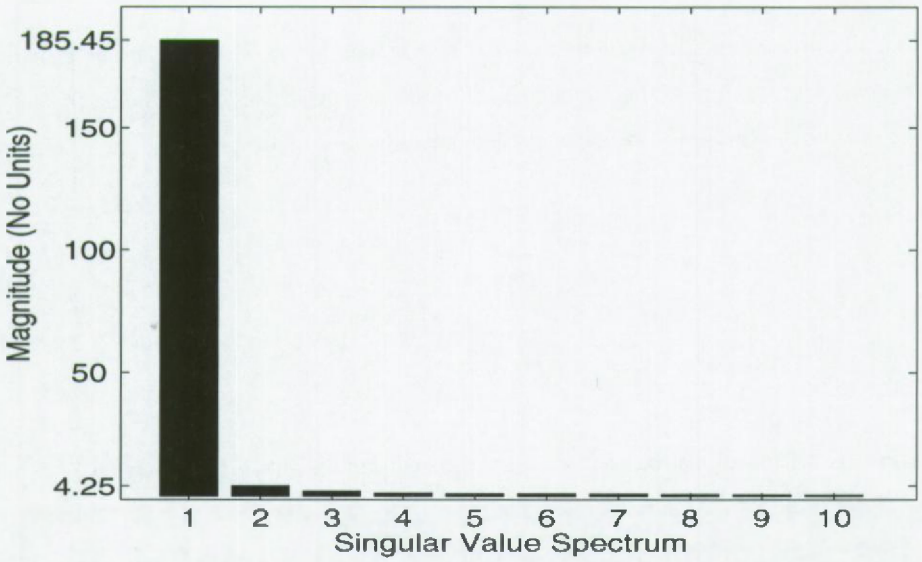

Figure 2: Singular values: more than 97% of the total variation of the raw data can be explained by the drift only. Other influences, like similar chemical properties and noise are less than 2.5%.

1. the fluctuation of the laser intensity, which can differ from measurement to measurement,
2. matrix effects (sample composition and surface structure can influence the ablation process),
3. transport efficiency during LA-ICP-MS analysis, etc....

The major distinction between the patterns of different elements is their scale, which is primarily due to differences in concentration. These observations confirm the conclusion that the measurements could be correctly represented by an measurement dependent pattern and an internal standard dependent scale, e.g. a bilinear model.

2.4.1.2 What is the lower bound of precision?

The singular values are scaled to their noise level, so the singular values 4 to 10 contain no information besides the stochastic noise contribution, because they are close to one (Figure 2). The first singular value has a magnitude of 185.4, so the relative error, due to disturbing noise on this model is given by the following equation (see appendix 2.A- σ_1 is the largest singular value)

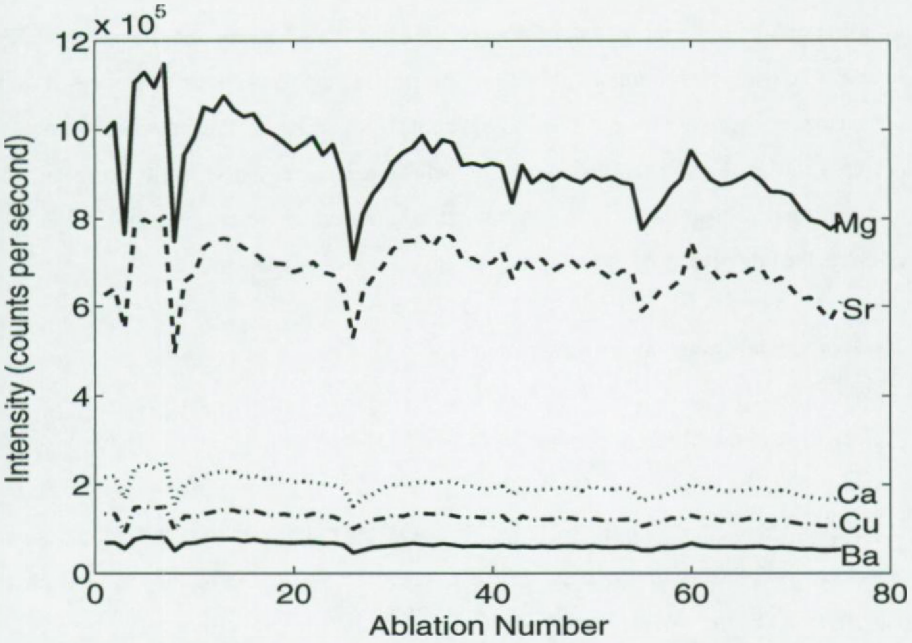

Figure 3: The raw data for five elements (Mg (solid line) – Sr (dashed line) - Cu (dotted line) – Ca (dash-dotted line) – Ba (lower solid line)) as a function of the subsequent measurements are shown. Notice the common drift pattern in all four elements.

$$\text{Noise errors} = \sqrt{\frac{1}{\sigma_1^2 - 1}} \approx 0.6\ \% \tag{2-8}$$

and is less than 0.6% for ten internal standards. When more internal standards would be used, the noise error will converge to zero. And as a result, the hypothetical lower bound of precision for this normalization and this instrument is limited to 97.5%, which is only due to model errors.

2.4.1.3 Is it useful to take a mass-dependency of the internal standards into account?

If a mass dependent correlation exists, it will appear in the model in the second and third singular values, because the first singular value describes the common variation in all internal standards. This influence is only marginal (2.4%), so neglecting it will cause only a very small error.

To conclude, the singular value decomposition of the scaled raw data matrix provided us information about the modelling of the raw data, the model errors and the stochastic errors. In addition, when this singular value decomposition is applied to more complicated samples, whose trace elements are not all homogeneously distributed, the inhomoginities will be detected in the singular values $\sigma_2, \sigma_3, \ldots$ which will differ significantly form one. These higher order contributions will describe the variations in concentration and allows to estimate the complexity of the biogeochemical processes in the sample.

2.4.2 How to match the model to the measurements?

A close look at the different patterns in Figure 3 shows that they are not truly identical, but they are disturbed by noise. The vectors u_1 and $C_I^{T/2} v_1$ are already good estimates of respectively the pattern and the scale factor. However, adding a weight to each measurement individually to further improve the estimate would destroy the structure of the scaled raw-data matrix. For that reason, a cost function approach is employed to refine the estimate of the drift pattern and the scale factor. To extract both from the disturbed measurements, two important assumptions about the noise are made:

1. the measurement noise $e(n, m)$ on the different measurements is independently distributed over n (measurement number) and m (internal standard number),
2. with a mean of zero, i.e. the stochastic noise will not introduce a systematic error. Under these conditions, it can be proven that the estimate of the drift pattern converges to the true values as the number of internal standards increases to infinity [Pintelon and Schoukens, 2001].

The matrix, I, defined in equation (2-1) or (2-2) must be written as a vector, S, to fit it in a cost function approach. Storing the columns of the matrix I on top of each other, it can be found after some calculations that

$$S = H(\alpha)p + \Delta_s \tag{2-9}$$

where S and Δ_s are $NM \times 1$-vectors, containing the columns of respectively I and e, stored on top of each other. For example

$$S = \begin{bmatrix} I(1,1) \\ \cdots \\ I(N,1) \\ \cdots \\ I(1,M) \\ \cdots \\ I(N,M) \end{bmatrix} \tag{2-10}$$

$H(\alpha)$ is given by

$$H(\alpha) = \begin{bmatrix} \alpha(1)I_N \\ \alpha(2)I_N \\ \cdots \\ \alpha(M)I_N \end{bmatrix} \tag{2-11}$$

with I_N the $N \times N$ identity matrix. The vector p remains the same, as defined in equation (2-1) or (2-2) and the structure of Δ_s is the same as this of S.

Once a model is expressed, we need to fit it to the measurements. To quantify the difference between the model and the measurements, a least squares cost function was chosen. Minimizing the cost function will match the model to the measurements. The part of the measurements, which does not fit into the model is identified as noise, more specifically: stochastic noise and model errors. Because of this separation, a better estimation of the drift pattern is made.

The choice for a least squares cost function is motivated by the following arguments:

1. minimizing a least squares estimator is less involved than alternative estimators,
2. normally distributed disturbing noise leads, naturally, to least squares estimators, and
3. many properties of this least squares estimators are already proven and described in the literature [Pintelon and Schoukens, 2001].

By introducing a weighting matrix, C^{-1} (with C the covariance matrix of the raw data vector S: $C = \mathrm{cov}(S)$), in the least squares cost function, K_{nwls}, highly uncertain measurements are suppressed and lowly uncertain measurements are emphasized:

$$K_{\mathrm{nwls}} = \frac{1}{2}(S - H(\alpha)p)^T C^{-1}(S - H(\alpha)p) \tag{2-12}$$

Now, to estimate the pattern and scale one must find the arguments, α and p, that minimize equation (2-12) with respect to both:

$$(\hat{p}, \hat{\alpha}) = \underset{p,\,\alpha}{\arg\min} K_{\mathrm{nwls}}(p, \alpha) \tag{2-13}$$

The cost function is a non-quadratic function of one of the parameters, p or α. The analytic solution of this minimization problem is not known, so an iterative algorithm is needed to calculate the pattern or scale [Pintelon and Schoukens, 1996]. We have chosen for a Newton-Gauss algorithm to find these parameters, which is described in appendix 2.A. Good starting values for this iteration algorithm are the vectors u_1 and $C_e^{T/2}v_1$, for respectively the pattern and the scale, calculated with the singular value decomposition, explained in paragraph 2.4.1. Note that multiplying α by λ and dividing p by λ does not change the cost function (2-12). This parameter ambiguity is removed by choosing one entry of α or p, for example $p(1) = 1$.

2.4.3 An a posteriori verification of the bilinear model

An a posteriori verification for the model choice can be executed within a statistical framework. First, assume that model errors are absent and that the noise is normally distributed, then the expected value and variance of the cost function (2-12) are given by the following formulas [Pintelon and Schoukens, 2001]:

$$\begin{aligned} E\{K_{\mathrm{nwls}}(\hat{p}, \hat{\alpha})\} &= \frac{1}{2}(N-1)(M-1) \equiv K_{\mathrm{noise}} \\ \mathrm{var}(K_{\mathrm{nwls}}(\hat{p}, \hat{\alpha})) &= \frac{1}{2}(N-1)(M-1) \end{aligned} \tag{2-14}$$

with N the number of measurements and M the number of internal standards. Comparison of equation (2-14) with the actual value of the cost function K_{nwls} allows to verify whether or not model errors are present. With 95% confidence the following test can be achieved [de Brauwere *et al.*, in press]

$$\begin{aligned} & K_{\text{noise}} - 2\sqrt{K_{\text{noise}}} \le K_{\text{nwls}} \le K_{\text{noise}} + 2\sqrt{K_{\text{noise}}} \\ & \quad \Rightarrow \text{no model errors can be detected} \\ & K_{\text{nwls}} > K_{\text{noise}} + 2\sqrt{K_{\text{noise}}} \\ & \quad \Rightarrow \text{model errors can be detected} \\ & K_{\text{nwls}} < K_{\text{noise}} + 2\sqrt{K_{\text{noise}}} \\ & \quad \Rightarrow \text{wrong noise variance} \end{aligned} \tag{2-15}$$

An estimate of the model errors is then given by

$$\begin{aligned} K_{\text{model errors}} &= K_{\text{nwls}} - K_{\text{noise}} \qquad K_{\text{nwls}} > K_{\text{noise}} \\ &= 0 \qquad\qquad\qquad\quad \text{otherwise} \end{aligned} \tag{2-16}$$

Furthermore, a quantification for the information covered by the model is given by the following equation:

$$K_{\text{model}} = \frac{1}{2}(H(\hat{\alpha})\hat{p})^T C^{-1}(H(\hat{\alpha})\hat{p}) \tag{2-17}$$

With these values the model error and stochastic error can be computed employing the following equations

$$\begin{aligned} \text{Model error} &= \sqrt{\frac{K_{\text{model errors}}}{K_{\text{model}}}} \\ \text{Stochastic error} &= \sqrt{\frac{K_{noise}}{K_{\text{model}}}} \end{aligned} \tag{2-18}$$

Using the same measurement data as in section 3.1 ($N = 75$, $M = 10$), the actual value of the cost function is $K_{\text{nwls}} = 954$. Applying formulas (2-14) to (2-17) then gives $K_{\text{noise}} = 333$, $K_{\text{model error}} = 621$, $K_{\text{model}} = 1.125 \times 10^6$, from which can be computed

that the model error and stochastic error are respectively 2.3% and 1.7%. From these errors, we may conclude that

1. both the singular value decomposition and the cost function approach estimate a total error of 3%;
2. the stochastic contribution in the latter is larger, probably due to the more refined weighting by the uncertainties.

On the whole, both approaches lead to the same levels of model and stochastic model errors. So, this cost function examination also establishes that the bilinear model is very well suited to represent the measurements of the internal standards.

2.4.4 The new normalization and its uncertainty

Once the common drift pattern is estimated, the data of all elements (internal standards and others), $\{1, \ldots, P\}$, can be normalized. In general, these are not internal standards and their raw data are represented by the $N \times P$-matrix I. These raw data are corrected for the drift by multiplication with a normalization factor, like in the classical normalization. The new normalization factor is given by the inverse of the estimated pattern (normalization factor $= 1/\hat{p}(n)$, where n indicates the measurement number). Within this algorithm it was demanded that the first value of the pattern is equal to one, $\hat{p}(1) = 1$, which made a scale in the normalization factor redundant. So, the raw data for measurement n of element i ($i \in \{1, \ldots, P\}$) are normalized by the following equation:

$$\tilde{I}(n, i) = \frac{I(n, i)}{\hat{p}(n)} \tag{2-19}$$

where $\tilde{I}(n, i)$ represents the normalized data. This normalization suppresses the total variation in all internal standards and corrects the other elements for the drift.

The covariance matrix of the estimated pattern is given by

$$\mathrm{cov}(\hat{p}) = (J_p^T C^{-1} J_p)^{-1} \text{ with } J_p = \left.\frac{\partial(S - H(\alpha)p)}{\partial p}\right|_{\hat{p}, \hat{\alpha}} \tag{2-20}$$

where $C = \mathrm{cov}(S)$. The standard deviation on the pattern can be found in the diagonal elements of the covariance matrix, i.e. $\sigma_p^2(n) = \mathrm{cov}(\hat{p}(n)) =$ nth diagonal element of the covariance matrix in equation (2-20). Once the pattern and its uncertainty are known, the uncertainty on the normalized data are given in a first order approximation of equation (2-19) by

$$\sigma_{\tilde{I}}^2(n,i) = \frac{1}{\hat{p}^2(n)}\sigma_I^2(n,i) + \frac{I^2(n,i)}{\hat{p}^4(n)}\sigma_p^2(n) \qquad (2\text{-}21)$$

where $I(n,i)$ is no internal standard and $\sigma_{\tilde{I}}(n,i)$ represents the standard deviation on the normalized data of measurement n of element i, $\sigma_I(n,i)$ is the respective standard deviation on the raw data and $\sigma_p(n)$ is the standard deviation on the estimated pattern. Notice that an additional cross-term is needed to calculate the uncertainty on the normalized data of an internal standard [Pintelon and Schoukens, 2001].

2.5 The results of the raw data normalizations

First the classical, and next the generalized normalization are validated and discussed. Five elements, i.e Mg, Ca, Cu, Sr, Ba, are chosen to explain the improvements and conflicts of the normalizations. Figure 3 shows the raw data for these five elements. The precision is primarily limited by the drift of the apparatus, which can be seen in the large, common variation of the signals of all the elements. This precision is quantified by the relative standard deviation (RSD), which is respectively 9.7%, 10.2%, 9.2%, 8.7% and 12.5%. The mean and standard deviation (both in counts per second) over the 75 subsequent measurements are employed to calculate this RSD.

2.5.1 The validation of the classical normalization

In a first attempt, this drift can be eliminated by a normalization based on a single internal standard as is described in paragraph 2.2.1. All elements have a constant concentration throughout the NIST 610 standard. So, one element can be chosen as an internal standard, while the four others can be used as a validation set, because we know that the variation in the pattern of the validation elements should hypothetically approach zero. In this example Ba was chosen as internal standard, but any another element would have led to the same conclusions.

The result of this single internal standard normalization is shown in Figure 4. Large improvements are made for the precision, as is illustrated by the new RSD values: 3.8%, 3.5% 5.6%, 7.6% and 0%, which are significantly smaller than the previous ones. Although the variation is reduced, a large fraction remains. This has two reasons:

1. the noise on the measurements of the internal standard appears now in all other elements, because it is neglected in the single internal standard normalization and is consequently a part of the estimated drift pattern;
2. the true and estimated drift pattern can still differ quite a lot. By way of illustration: if another element is chosen as internal standard, the profile of the validation elements would

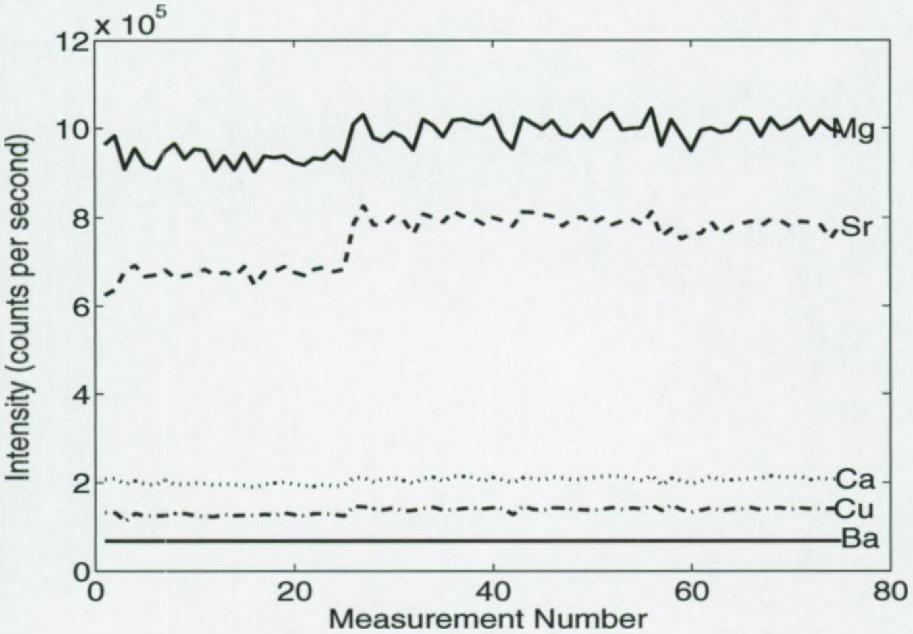

Figure 4: The normalized patterns five elements (Mg (solid line) – Sr (dashed line) - Cu (dotted line) – Ca (dash-dotted line) – Ba (lower solid line)) as a function of the subsequent measurements are shown. The raw data were normalized with the single element normalization employing Ba as an internal standard.

have been different, as can be seen in Figure 5: the five different profiles of Cu, which appear after normalization with each of the five elements used as an internal standard. Differences of 10 to 40% are not uncommon. Consequently, it can become hard to distinguish patterns in trace elements from ordinary artificial patterns, due to the drift or drift correction.

Summarized, different internal standards do not always lead to the same normalization factors.

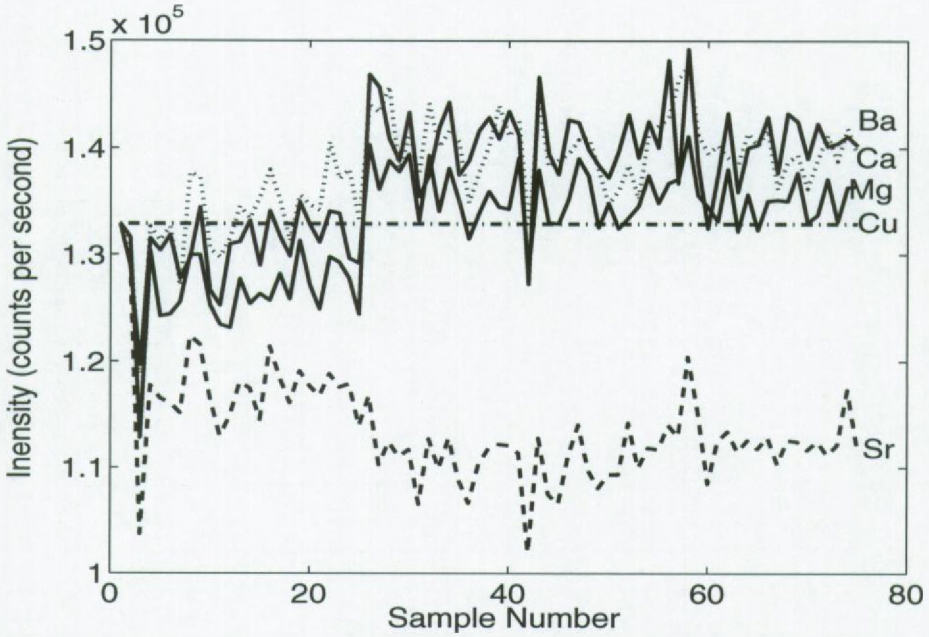

Figure 5: The five different profiles of Cu after normalization with subsequent each of the five elements, i.e. Mg (solid line) – Sr (dashed line) - Cu (dotted line) – Ca (dash-dotted line) – Ba (lower solid line), used as an internal standard.

2.5.2 The validation of this new approach and the influence of the number of internal standards

To illustrate this new normalization, we used the same measurements as utilized in the validation of the classical normalization (75 measurements taken from a NIST 610 standard as described above). A typical example is given by choosing Cu as a validation element and Ba, Ca, Sr and Mg as internal standards. To illustrate the influence of the number of internal standards, the signal of Cu is subsequently normalized with one to four internal standards. The results of these normalizations are shown in Figure 6. The first plot shows the normalized signal of Cu with respect to Ba. Large variations in the signal are noticed, quantified in the RSD: 5.6%, which is the same as in the classical single element normalizations. In fact, the classical normalization is a special case of this one, as will be shown in paragraph 2.6.1.

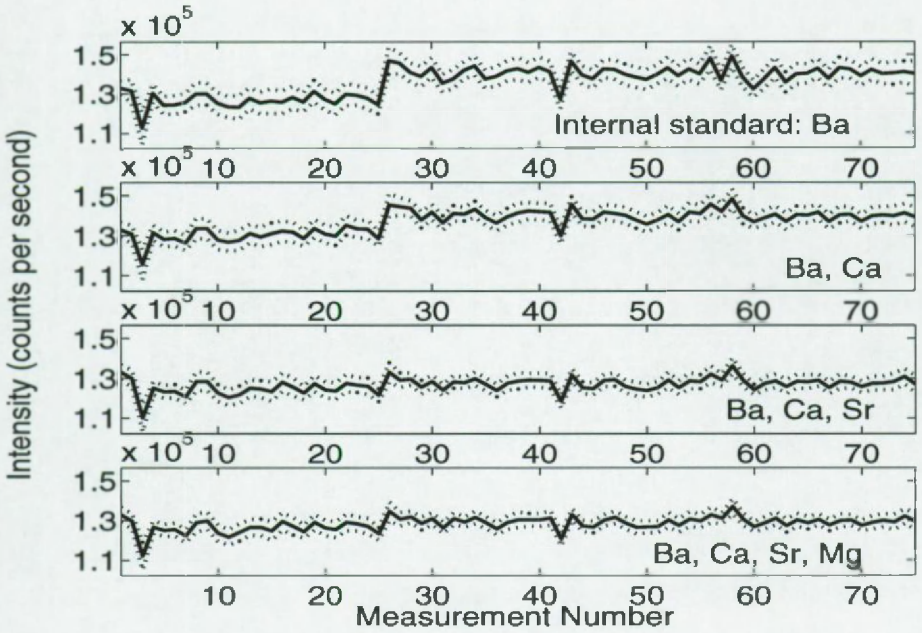

Figure 6: These four figures show the normalized profiles of Cu as a function of the subsequent measurements. In descending order: first only Ba is used as internal standard, then respectively Ca, Sr and Mg are added to the group of internal standards. The gray bounds visualize the 68% uncertainty bounds on the normalized data.

Regarding the normalization by two internal standards (Ba and Ca), the variation has dropped significantly as is seen on the second plot of Figure 8. Its RSD is 4.29%, which is an decrease with one fourth, compared to the single internal standard normalization.

When more internal standards are introduced, the RSD decreases further (two plots at the bottom of Figure 6). The RSD values for three (Ba, Ca and Sr) and four (Ba, Ca, Sr and Mg) internal standards are respectively 2.9% and 2.7%. This is a decrease of respectively 49% and 51% compared to the classical normalization. We may conclude from this example that the use of more internal standards will improve the RSD. This is quantified by a relation between both: the variance on these patterns is inversely proportional to the number of internal standards

$$\sigma_p^2(n) \sim \frac{1}{M} \tag{2-22}$$

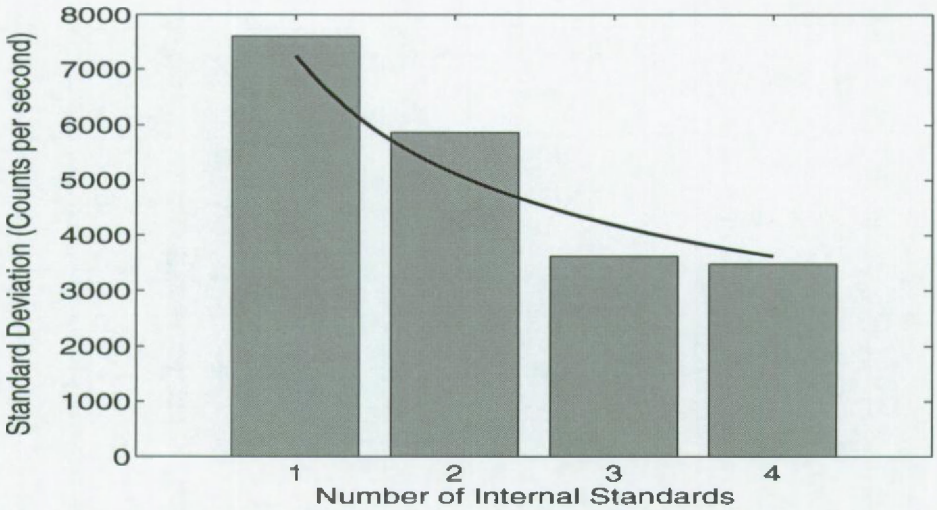

Figure 7: These four figures show the normalized profiles of Cu as a function of the subsequent measurements. In descending order: first only Ba is used as internal standard, then respectively Ca, Sr and Mg are added to the group of internal standards. The gray bounds visualize the 68% uncertainty bounds on the normalized data.

where $\sigma_p^2(n)$ is the variance on the pattern at position n and M is the number of internal standards. This relation is proven in appendix 2.C and illustrated in Figure 7: the standard deviation over the 75 subsequent measurements is plotted versus the number of internal standards in a bar diagram. The curve shows the square root of the number of internal standards. It can be seen that the variability decreases proportionally to this square root, because the uncertainty due to the drift pattern is much larger than the measurement uncertainties of signal of Cu. This means that two internal standards can decrease the variability by nearly 30%, three with 42%, four up to 50%, etc.... However, two remarks can be made:

1. the contribution of model errors, compared to the total error, will become more important, when the precision increases. So, after a normalization with an infinite number of internal standards, these signals will still have some variation, despite this proof. This could already have been noticed from the lower bound of precision that can be reached with this framework, as is discussed in paragraph 2.4.1.

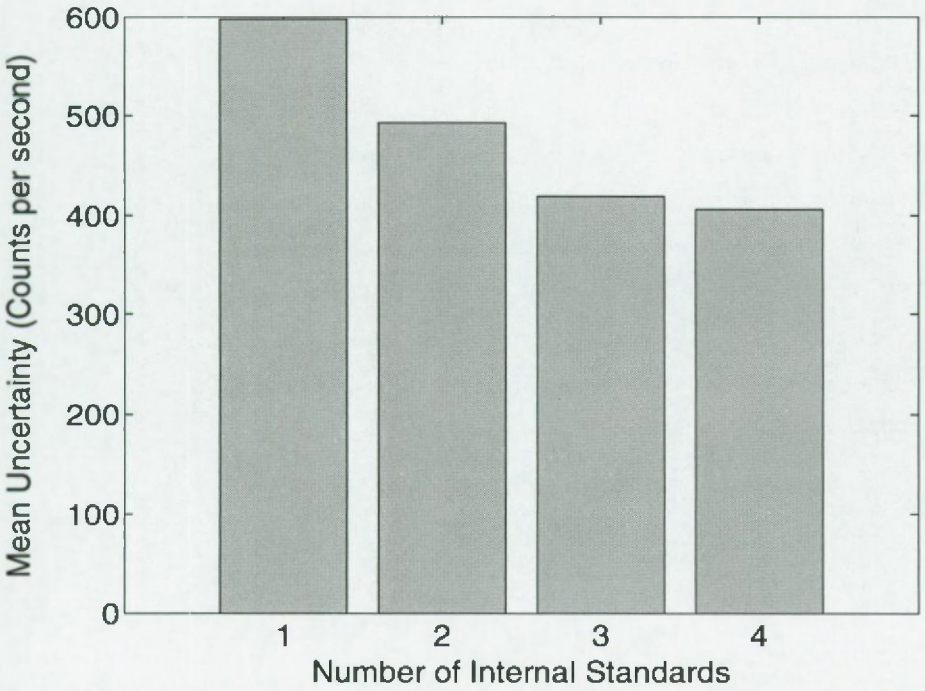

Figure 8: Mean 68% uncertainty bound over 75 subsequent measurements as a function of the number of internal standards.

2. the stochastic noise on the element, to be normalized, will remain. So the variability of elements, which can hardly be measured, will have a smaller profit from this multiple internal standard approach. The second remark can be solved by stabilizing the ablation signal; i.e. artefacts like fractionation must be tackled.

The same remarks can be made on the uncertainty bounds. These are shown as grey bounds on Figure 6. The average uncertainty also decreases with an increasing number of internal standards, but this uncertainty decreases more slowly as illustrated in Figure 8. An explanation for this is found in the first term of equation (2-21). Again, the uncertainty on the raw data, which are to be normalized, does not change with an increasing number of internal standards. As a result the uncertainty will not converge to zero.

2.6 Remarks

2.6.1 A special case: The classical normalization

To find the estimated pattern in this approach one must find the arguments that minimize equation (2-12), under the constraint that the first value of the pattern remains one. Let us now look at the case of only one internal standard. The first term in the cost function will be zero, if the scale $\alpha(1)$ is equal to the first measurement $S(1, 1)$, explicitly $S(1, 1) = \alpha(1) \cdot 1$. All subsequent terms will be zero if $p(n) = S(n, 1)/S(1, 1)$, where $p(n)$ is the pattern at position n and $S(n, 1)$ is the respective measurement of the internal standard. Explicitly written:

$$\begin{aligned} K_{\text{nwls}} &= \sum_{n=1}^{N} [S(n,1) - \alpha(1)p(n)]^2 C^{-1}(n,n) \\ &= \sum_{n=1}^{N} [S(n,1) - S(1,1) \cdot S(n,1)/S(1,1)]^2 C^{-1}(n,n) \\ &= 0 \end{aligned} \tag{2-23}$$

Again, the normalization factor is given by the inverse of the pattern, which is the classical single element normalization. Accordingly, in this special case the weighting matrix C^{-1} has no influence on the result. To conclude, the classical normalization is a special case of this multiple internal standard normalization under the condition that only one internal standard is used.

2.6.2 Internal quality control

A problem, which could occur, is that the measurement conditions suddenly change during the course of an experiment. For example: the signal of some elements drop from one measurement to another, while the signal for the others remain steady. How can such artifacts be detected? If one examines the patterns of several internal standards, he/she can detect differences, which could be due to these artifacts and redraw the measurements. An extra advantage of the cost function approach of the normalization is that the value of this cost

function is already a good measure for artifacts. If the pattern of some internal standards changes suddenly, the calculated value of the cost function will be much higher than expected. Comparison of the actual value of the cost function ($K_{\text{nwls}}(\hat{p}, \hat{\alpha})$) with its expected value (K_{noise} (see equation (2-14)) in the absence of model errors allows to detect artifacts in the measurements.

2.7 Example

So far, we used an artificial silicate standard to illustrate how this normalization works and what its properties are. Let us now turn to a real sample: a sclerosponge. The purpose of this measurement is to find relations between ecological parameters like salinity, sea water temperature, pollutions, etc.... on the one hand and variations in trace elements on the other hand [Lazareth *et al.*, 2000; Swart *et al.*, 2002]. We will illustrate the importance of a good and reliable normalization on the profile of ^{26}Mg. This sample consists of biological aragonite, so usually Ca is employed as internal standard, that is to say ^{43}Ca. The most evident internal standards for this approach are ^{43}Ca, ^{46}Ca and ^{48}Ca. 44 Measurements, with a crater diameter of 40 μm, were preformed under the experimental conditions as described in 2.3.2. The distance between the centers of two successive craters is 30 μm. Knowing that a sclerosponge grows approximately 200 μm/year [Swart *et al.*, 2002; Willenz and Hartman, 1985; Willenz and Hartman, 1999] which means that we have a sample frequency of about 7 per years. Data were acquired during 40 s in time resolved analysis. Instead of measuring the blanks after each ablation, these were measured six times in-between the measurements. These results were interpolated to correct for the blank. The first 20 s were neglected when calculating the mean signal and its standard deviation. With these data three single internal standard and one triple internal standard normalization were performed. Figure 9 shows the pattern of Mg after these three single element normalizations. From these data it is not possible to predict if the Mg concentration is increasing, decreasing or if it remains constant over time (years). In addition, it is not possible to tell either the inter-annual peaks that appear are significant variations or disturbing noise. When the multiple internal standard normalization is performed, the value of the cost function (K_{nwls}) was 46.6 (no units). Using equation (2-15) we may conclude that no artifacts occurred in the measurement of the internal standards, particularly K_{nwls} is in-between 29.9 and 56.1 (no units). The mean relative standard deviation over the 44 measurements of the three single internal standard normalizations are respectively 7.21%, 8.96% and 7.16% (standard deviation is calculated with equation (2-21) and the signal is calculated with equation (2-19)). The mean relative standard deviation of the triple internal standard normalization is 5.97%. This decrease of

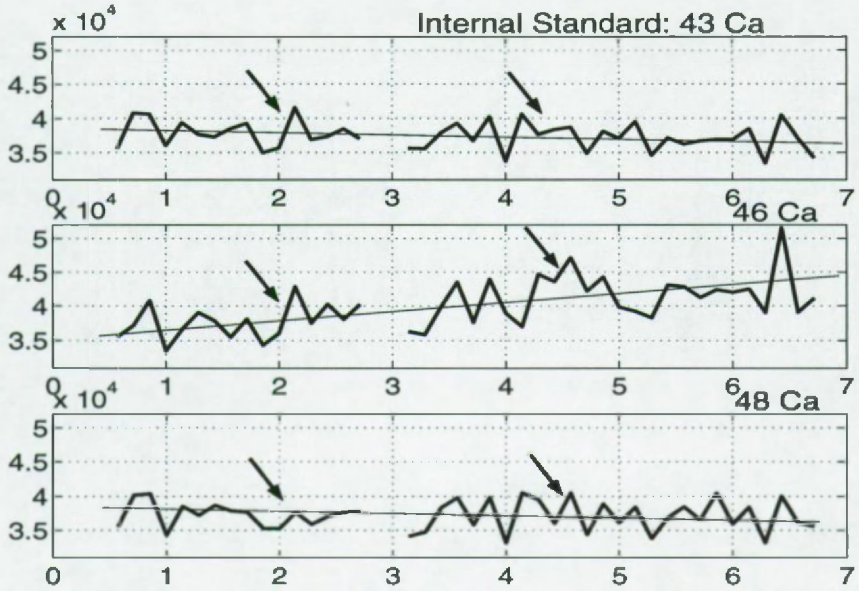

Figure 9: The intensity for 44 measurements of ^{26}Mg in biological calcite, normalized with ^{43}Ca, ^{46}Ca, ^{48}Ca. Note the different trends in the intensity (solid line) and the appearance and disappearance of peaks (arrows). Two measurements at year 3 are missing, due to a problem with the data extraction from the operating system.

about 23% is a significant improvement. Figure 10 shows the profile after a normalization with the three isotopes of Ca. Above all, this demonstrates that no trend is present in the Mg concentration and that the peaks that arise are annual and bi-annual fluctuations in the Mg-profile, which may be a consequence of the annual and bi-annual rainy season in the tropical region where this sclerosponge has lived. Anyway, a detailed discussion about the interpretation of this profile is beyond the scope of this chapter.

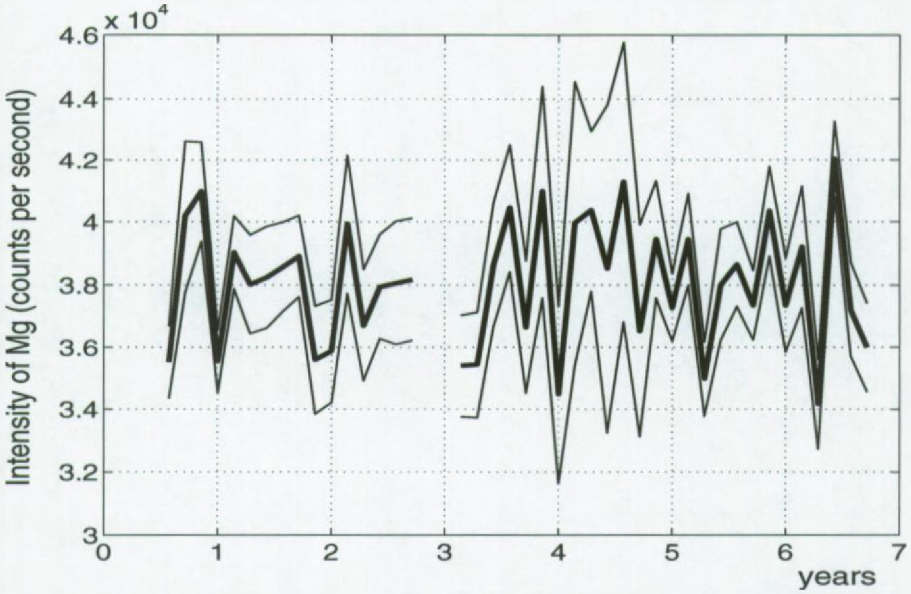

Figure 10: The intensity of ^{26}Mg, normalized with ^{43}Ca, ^{46}Ca, ^{48}Ca simultaneously. Note the absence of a trend, the annual and bi-annual variations and the large fluctuations in uncertainty, e.g. the period between the years 4 and 5 compared to that between years 5 and 6. Two measurements at year 3 are missing, due to a problem with the data extraction from the operating system.

2.8 Conclusion

In this chapter, chemometric techniques (parameter estimation and system identification) are implemented into the field of trace element analysis: a generalization of the single internal standard normalization for LA-ICP-MS measurements is reported, based on the use of multiple internal standards from which a drift pattern is extracted.

First of all, a good model to represent the measurements of the internal standards was chosen, employing a singular value decomposition. From these results we can conclude that the model error is about 2.5%. This result is independently confirmed by an a posteriori verification. Secondly, to estimate the drift pattern a cost function approach is employed, which takes for each measured internal standard not only the number of counts into account, but also its individual uncertainty. This emphasizes the more precise measurements and eliminates the stochastic errors of the different internal standards. Thirdly, it is proven that the standard deviation on the drift pattern is inversely proportional to the square root of the number of internal standards. Fourthly, uncertainty bounds on the normalized data are generated, which enables us to estimate the precision on each measured element. Finally, an extra advantage is that an internal quality check can be performed, using the value of the cost function.

A first attempt on an artificial NIST 610 sample confirms the predicted improvements on the precision. This leads to a decrease in the variability of the signal, expressed as an RSD, of 23% for two internal standards, 49% for three and 51% for four. A second attempt on a real sample (sclerosponge) explains by example how the multiple internal standards can be chosen, how the internal quality check can be performed and that the calculated uncertainty on the measurements decreases with an increasing number of internal standards.

Appendix 2.A: an estimation of the stochastic and model errors from a singular value spectrum

A brief introduction to the singular value decomposition can be found in the [Golub and Van Loan, 1990; Jennings, 1977]. It will be shown that the smallest singular values of the

singular value spectrum Σ of the noisy intensity matrix I (counts per second), defined in equations (2-1) and (2-2) and scaled by its uncertainty (counts per second) will converge to one. From this expression formulas for the model and stochastic error will be derived. First, the intensity scaled by the uncertainty (expressed as a square root of the column-covariance matrix ${C_e}^{1/2}$, with $C_e = E\{e^T e\} = {C_e}^{T/2}{C_e}^{1/2}$) is

$$I_{\text{scaled}} = I{C_e}^{-1/2} \tag{2-24}$$

The scaled intensities I_{scaled} (no units) can be modeled as

$$I_{\text{scaled}} = I_o + e{C_e}^{-1/2} \tag{2-25}$$

where I_o are the undisturbed scaled intensities (no units) and e is stochastic noise. Suppose that I_o is of rank R (with $R < M$). Then I_o can be decomposed as

$$I_o = U\Sigma_o V^T \tag{2-26}$$

where U and V are orthogonal matrices and Σ_o is a diagonal matrix $\Sigma_o = \text{diag}(\sigma_1, \ldots, \sigma_R, 0, \ldots, 0)$ describing the singular value spectrum of the model, ordered from large to small ($\sigma_1 \geq \ldots \geq \sigma_R$).

Next, the expected value of the square of I_{scaled} is calculated, assuming that the noise e has a zero mean ($E\{e\} = o$), so its expectation value is zero (equations (2-24), (2-25) and (2-26) are used)

$$\begin{aligned} E\left\{{I^T}_{\text{scaled}} I_{\text{scaled}}\right\} &= I_o^T I_o + C_e^{-T/2} C_e C_e^{1/2} \\ &= V\Sigma_o^2 V^T + I_M \\ &= V(\Sigma_o^2 + I_M)V^T \end{aligned} \tag{2-27}$$

When the number of measurements N is large enough, it can be shown that ${I^T}_{\text{scaled}} I_{\text{scaled}}$ is close to $E\left\{{I^T}_{\text{scaled}} I_{\text{scaled}}\right\}$ [Pintelon and Schoukens, 2001]. The calculated singular value spectrum is given by

$$\Sigma^2 = \Sigma_o^2 + I_M \tag{2-28}$$

So, the spectrum Σ_o can be calculated as

$$\Sigma_o = \sqrt{\Sigma - I_M} \tag{2-29}$$

It is assumed that the measurements are modeled by a bilinear model, consequently all other singular values ($\sigma_2, \ldots, \sigma_R$) will contribute to model errors. So, the total model error is given by equation (2-7). Further, the singular values of Σ_o are scaled to their uncertainty, so the noise error of this bilinear model is given by equation (2-8).

Appendix 2.B: the non-linear least-squares algorithm

In general, the analytical solution of the non-linear least-squares problem (equation (15)) is not known. The Newton-Gauss method is very well suited to deal with the least-squares minimization problem because it makes use of the structure of the cost function [Pintelon and Schoukens, 2001]. The calculations follow the lines of reference [Pintelon and Schoukens, 1996].

Minimizing equation (2-12) with respect to p gives

$$\hat{p} = [H(\alpha)^T C^{-1} H(\alpha)]^{-1} H(\alpha)^T C^{-1} S \tag{2-30}$$

where $\hat{p}$ is the estimated pattern. Substituting of equation (2-30) into (2-12) gives the final non-linear least squares cost function in the parameter α

$$K_{\text{nwls}}(\alpha) = \frac{1}{2}[P(\alpha)S_w]^T[P(\alpha)S_w] \tag{2-31}$$

where $P(\alpha) = I_{NM} - C^{-1/2}H(\alpha)[H(\alpha)C^{-1}H(\alpha)]^{-1}H(\alpha)^T C^{-T/2}$, with I_{NM} a $NM \times NM$ unity matrix and $S_w = C^{-1/2}S$. Minimization of equation (2-31) with respect to α gives the non-linear least-squares estimate of the scale factors:

$$\hat{\alpha} = \arg\min_{\alpha} K_{\text{nwls}}(\alpha) \tag{2-32}$$

This argument $\hat{\alpha}$ is found by the iterative Newton-Gauss procedure. The *k*th iteration step has the form

$$\begin{aligned} \alpha^{k+1} &= \alpha^{k} + \Delta\alpha \\ \Delta\alpha &= -J_{\alpha}^{+}\,[P(\alpha)S_w] \\ &= -\left(P\frac{\partial H}{\partial \alpha}H^{+} + \left(P\frac{\partial H}{\partial \alpha}H^{+} \right)^{T} \right) S_w \end{aligned} \tag{2-33}$$

where the superscript + denotes the Moore-Penrose pseudo-inverse $A^{+} = (A^{T}A)^{-1}A^{T}$. Good starting values for the scale factor, α^{o}, can be found using a singular value decomposition of the raw data as is explained in paragraph 2.4.1. Once the scale factor is estimated, the pattern can be calculated employing equation (2-30).

Appendix 2.C: convergence properties of the pattern

Equation (2-22) can be proved as follows: the cost function is given by equation (2-12) and can be rewritten as a function of the pattern, p:

$$K_{\text{nwls}}(p) = \frac{1}{2}[Q(p)S_w]^{T}[Q(p)S_w] \tag{2-34}$$

where $Q(p) = I_{NM} - C^{-1/2}G(p)[G(p)C^{-1}G(p)]^{-1}G(p)^{T}C^{-T/2}$ with $G(p)$ given by

$$G(p) = \begin{bmatrix} p & 0 & \dots & 0 \\ 0 & p & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & p \end{bmatrix} \tag{2-35}$$

which is a $NM \times NM$-matrix. Following the guidelines of reference [Pintelon and Schoukens, 1996], it follows that

$$\mathrm{cov}(\hat{p}) = \left[\left(\frac{\partial Q(p)}{\partial p} S_w\right)^T \left(\frac{\partial Q(p)}{\partial p} S_w\right)\right]^{-1} \tag{2-36}$$

where $\mathrm{cov}(\hat{p})$ is the covariance matrix of the pattern. After some algebra, the diagonal elements of the Fisher information matrix can be explicitly written as

$$F(n,n) = \sum_{m=1}^{M} \sum_{n=1}^{N} Y^2, \text{ with}$$

$$Y = \frac{p(n)}{\sigma_I(n,m)}\left(\frac{I(n,m)}{D\sigma_I^2(n,m)} - \frac{Z}{D}\left(\frac{2p(n)}{\sigma_I^2(n,m)} + \frac{1}{\sigma_I(n,m)}\right)\right), \text{ with} \tag{2-37}$$

$$D = \sum_{n=1}^{N} \frac{p^2(n)}{\sigma_I^2(n,m)} \text{ and } Z = \sum_{n=1}^{N} \frac{p(n)I(n,m)}{\sigma_I^2(n,m)}$$

$$F(n,n) \geq Mk \tag{2-38}$$

where M is the number of internal standards, N the number of measurements, $I(n,m)$ the nth measurement of internal standard m, $\sigma_I(n,m)$ the respective standard deviation and $p(n)$ the pattern at measurement n. Each term is a square and thus always positive, the smallest term is called k and as a consequence the right hand summation must always be greater or equal to M times k. The covariance matrix is the inverse of this matrix, so $\mathrm{cov}(p(n)) = \sigma_p^2(n) \leq 1/Mk$. This is true for all n, so equation (2-22) is proved.

References

[1] de Brauwere A., De Ridder F., Pintelon R., Elskens M., Schoukens J. and Baeyens W., 2005. Model selection through a statistical analysis of the minimum of a Weighted Least Squares cost function, Chemometrics and Intelligent Laboratory Systems, in press.

[2] Cheatham M.M., Sangrey W.F. and White W.M., 1993. Sources of error in external calibration ICP-MS analysis of geological samples and an improved non-linear drift correction procedure, Spectrochemica Acta, 48B, E487-E506.

[3] Chen Z., Doherty W. and Grégoire D.C., 1997. Application of laser sampling microprobe inductively coupled plasma mass spectrometry to the in situ trace element analysis of selected geological materials, J. Anal. At. Spectrom., 12, 653-659.

[4] Chen Z., Canil D. and Longerich H.P., 2000. Automated in situ trace element analysis of silicate materials by laser ablation inductively coupled plasma mass spectrometry, Fresenius J. Anal. Chem., 368, 73-78.

[5] Fryer B.J., Jackson S.E. and Longerich H.P., 1995. The design, operation and role of the laser-ablation microprobe coupled with an inductively coupled plasma - Mass spectrometer (LAM-ICP-MS) in the earth sciences, The Canadian Mineralogist, 33, 303-312.

[6] Golub G.H. and Van Loan C.F., 1990. Matrix Computations. Johns Hopkins Univ. Press, Baltimore MD.

[7] Günther D., Horn I. and Hattendorf B., 2000. Recent trends and developments in laser ablation-ICP-mass spectrometry, Fresenius J. Anal. Chem., 368, 4-14.

[8] Gray A.L., 1985. Solid Sample introduction by laser ablation for inductively coupled plasma source mass spectrometry, Analyst, 110, 551-556.

[9] Jackson S.E., Longerich H.P., Dunning G.R. and Freyer B.J., 1992. The application of laser-ablation microprobe - inductively coupled plasma - mass spectrometry (LAM-ICP-MS) to in situ trace-element determination in minerals, Canadian Mineralogist, 30, 1049-1064.

[10]Jeffries T.E., Jackson S.E. and Longerich H.P., 1998. Application of a frequency quintupled Nd: YAG source (lambda = 213 nm) for laser ablation inductively coupled plasma mass spectrometric analysis of minerals, J. Anal. At. Spectrom., 13, 935-940.

[11]Jennings A., 1977. Matrix Computation for Engineers and Scientists, John Wiley & Sons, London.

[12]Lazareth C.E., Willenz Ph., Navez J., Keppens E., Dehairs F. and André L., 2000. Sclerosponges as a new potential recorder of environmental changes: Lead in *Ceratoporella nicholsoni*, Geology, 28, 515-518.

[13]Pintelon R. and Schoukens J., 1996. An improved sine-wave fitting procedure for characterizing data acquisition channels, IEEE trans. Instrum. Meas., 5, 588-593.

[14]Pintelon R. and Schoukens J, System Identification A Frequency Domain Approach, IEEE PRESS, New York, 2001.

[15]Rolain Y., Schoukens J. and Pintelon R., 1997. Order estimation for linear time-invariant systems using frequency domain identification methods, IEEE Trans. Autom. Contr., 42,1408-1417.

[16]Swart P.K., Thorrold S., Rubenstone J., Rosenheim B., Harrison G.G.A., Grammer M. and Latkoczy C., 2002. Intra-annual variation in the stable oxygen and carbon and trace element composition of sclerosponges, paleoceanography, 17 (3), Art. No. 1045.

[17]Vanhaecke F., de Wannemacker G., Moens L., Dams R., Latkoczy C., Prohaska T. and Stingeder G., 1998. Dependence of detector dead time on analyte mass number in inductively coupled plasma mass spectromety, J. Anal. At. Spectrom., 13, 567-571.

[18]VG Elemental (Fisons Instruments), Plasma Quad Software Manual, Appendix E.

[19]Willenz Ph. and Hartman W.D., 1985, Proceedings of the Fifth Coral Reef Congress, Tahiti. 5. (Antenne Museum-EPHE: Moorea, French Polynesia): i-xvi, 113-118.

[20]Willenz Ph. and Hartman W.D., 1999, Memoirs of the Queensland Museum, 44: 675-685.

CHAPTER 3

DECODING NON-LINEAR GROWTH RATES IN BIOGENIC ENVIRONMENTAL ARCHIVES

Abstract - The record of an environmental proxy along a growth axis in biogenic carbonates can reflect changing environmental conditions experienced during the lifetime of the organism. When a chronology of the growth axis is not available, a method based on anchor points is commonly used to introduce a time grid. Between these anchor points a constant growth rate is generally assumed, despite the fact that growth rates change during life in most organisms. Here, we present a method which refines the constant growth rate assumption, at least in situations where the environmental proxy has a periodic component in its signal. A non-linear growth rate can then be estimated, enabling the construction of a more realistic time base. The anchor point method and the method proposed here are first compared on a synthetic record. It is shown that the choice of the anchor points, which is always a subjective one, is largely influenced by the stochastic noise on the data record. This results in biased growth rate profiles and biased time grids. Next we apply the method on the Mg-record in the bivalve *Isognomon ephippium*, on three samples of the clam *Saxidomus giganteus* from Washington state and on the Vanuatu coral stable isotope record. The method developed is applicable to a wide range of proxy records, which are measured along an indirect time grid and are complicated by non-linear accretion rates. Furthermore, this method which takes variable growth rates into account opens perspectives on quantitative modelling of the relationship between environmental proxies recorded by accretive archives and the instrumental record, like temperature or salinity.

3.1 Context

The remaining of this manuscript is devoted to signal processing. A major problem occurring in almost all proxy records is the dating of the observations, which is necessary to produce a time series. It is only the latter which can be compared with other proxies, with instrumentally measured environmental parameters or with models.

Note, before the discussion is started, that we would like to stress the difference between the time scale, which is related to the age of the record and is expressed in e.g. days, kyear,... and the time base, which is dimensionless and dates each individual observation w.r.t. the other observations. We will mainly focus on the latter.

In order to construct a time series, a signal model has to be proposed for the signal profile. This signal model is function of time and by matching the observations on this signal model, a time base is usually constructed. The calibration curves used in carbon dating is an example of such signal model. However, for the particular proxy records we will examine, carbon dates are not present. Therefore, other signal models have been used, i.e. periodic signal models.

The central idea presented in this work is the time base distortion, which is a measure for variations in accretion or growth rate. For the particular case of periodic signals this time base distortion can easily be identified. How this is done, is explained in detail in this chapter.

The ideas were developed by examining the vessel density of mangrove trees. Herein, we could see annual variations, which could not so easily be interpreted in the Fourier spectrum. In order to clarify the spectrum of the vessel density the anchor point method has been used. This is a simple conceptual approach, were the number of samples within one year is kept constant by resampling, i.e. a linear interpolation. After this operation, all non-annual vessel density peaks had disappeared, allowing us to conclude that these 'ghost' peaks around the annual periodicity were hiding somehow the information about the growth rate. The similarity of the disappearing side peaks, directed us to the work of Jan Verspecht [1994], who used a phase demodulation to identify a time base distortion in high frequency scopes.

Herein, more than one million observations are gathered in one second, while we applied it on records of eight observations per year for periods which are sometimes smaller than ten years, gathered in roughly one week work. However, the abstract problem was identical, just as the treatment.

To summarize, the signal model is coupled to the time base. Variations in this time base are isolated in a term, called the time base distortion. For periodic signals, this time base distortion can be identified.

This work was performed on data, gathered by Claire Lazareth *et al.* [2003], David Gillikin *et al.* [submitted] and Quinn *et al.* [1996].

3.2 Introduction

Earth's climate system varies on a range of time scales due to natural or anthropogenic forcings. Predictions of future climate require that global climate models are tested and validated against instrumental measurements of meteorological and environmental data, and also against long-term records of paleoclimate based on trace element and isotopic records preserved in layered substrates. Such climate and environmental information can be traced back from proxy records in many solid substrates characterized by a wide range of accretion rates, e.g. sclerosponges [Lazareth *et al.*, 2000], speleothems [Verheyden *et al.*, 2000; Finch *et al.*, 2001], corals [Sinclair *et al.*, 1998; Fallon *et al.*, 1999; Wei *et al.*, 2000; Kuhnert *et al.*, 2002; Marchall *et al.*, 2002], bivalves [Vander Putten *et al.*, 1999; Lazareth *et al.*, 2003], sediments [Weedon, 1989; Herbert, 1994] and ice cores [Petit *et al.*, 1999].

In such data-series, the environmental proxy is not measured directly as a function of time, but along a growth or accretion axis. Since we are mostly interested in the time series this distance profile has to be transformed into a time profile (Figure 1), which requires knowledge about growth or accretion rate. In general, it is not possible to estimate the growth or accretion rate and the time profile from the distance grid. So, without extra information or extra assumptions, one has to date at least two observations and assume a constant growth rate to construct a time series. This is illustrated by the dotted line in Figure 1. An example of such extra information are growth bands, present in some corals, bivalves or trees, which can be used to date the record, employing e.g. the anchor point method [Jones, 1983; Paillard *et al*, 1996]. However, this can induce non-negligible bias, as will be shown in a simulation (paragraph 3.4.2). Furthermore, even when growth lines are available, they usually only provide information about variations on annual and multi-annual scales and not on an sub-annual scale.

In the present chapter, we propose a method to transform the distance series into a time series, based on the assumption that the record is periodic. Fluctuations in growth rate are reconstructed by decoding the side peaks in the Fourier spectrum by a phase demodulation,

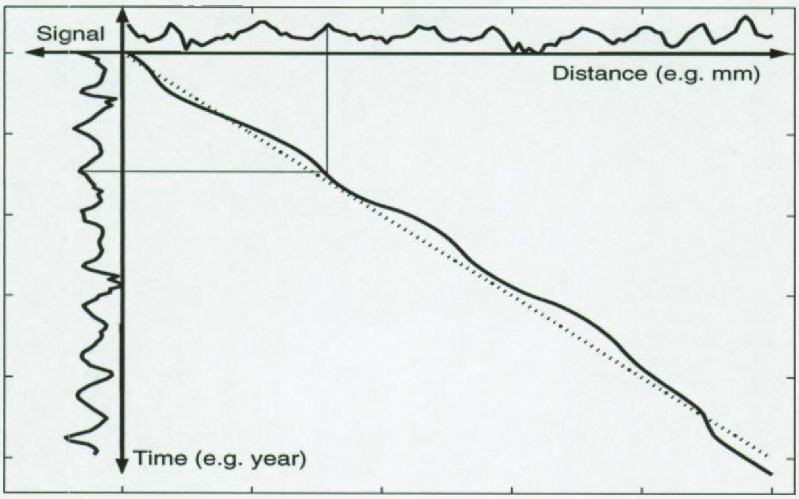

Figure 1: Conceptual graph showing the transformation from a record along a distance grid to time series. The dotted line represents a constant growth rate, while the curved line along the diagonal represents the real growth rate (data of Clam 2 are used, which is discussed in paragraph 3.6).

which are sometimes mis-interpreted as noise. The variations in growth rate are modelled as a distortion of the time base. Once the distortion is known, a more accurate time base can be reconstructed eventually resolving inter-annual variation. In the context of this work, it is important to make the distinction between the time base, which gives the relative date of each observation with respect to the first and last observation and the time scale, which measures the time interval of the record.

First, we compared the underlying assumptions of the method with those of the anchor point method [Jones, 1983; Paillard, 1996]. Next, both methods are tested and validated using simulation data. Finally, some applications are discussed:

1. the Mg-record in the bivalve *Isognomon epphippium* (Tudor Creek, Mombasa, Kenya);
2. the $\delta^{18}O$-records of three specimen of the bivalves *Saxidomus giganteus* (Pusset Sound, Washington, U.S.A.); and
3. the Vanuatu coral stable isotope data. The method developed is applicable to a large range of

accretive substrates which are measured along a distance grid and have to be transformed into time series.

3.3 Conceptual approach of the time base distortion method

The most commonly used method for converting a spatial scale to a time scale assumes a step-wise constant growth rate and is called the anchor point method. In the following sections, we first describe how variations in growth rate can be expressed as a distortion of linear growth rate, followed by a method to estimate this distortion of the time base.

3.3.1 A model for the time base distortion

The growth rate is not estimated directly. On the one hand we know the distances between subsequent observations and on the other hand we would like to know the time instances between these observations. The ratio of increments of both is the growth rate. To estimate the time instances, we start from an average constant growth rate, where the time instances between the observations are constant. So the time instance at which the nth observation was formed is given by

$$t_n = nT_s \tag{3-1}$$

In equation (3-1) T_s is the sample period (dimension is time), which is a measure for the average time instance between two subsequent observations. This constant growth rate model would correspond to the dotted line on the diagonal of Figure 1. Equation (3-1) will not be valid when the time instances are disturbed, due to variations in growth or accretion rate. This distortion of the time base can be modelled by a distortion term, $g(n)$, which is (i) zero when the growth is equal to the average growth, (ii) negative when the growth is slower than the average and (iii) positive when the growth is greater than the average. So, an improved estimate of the time instant at which observation n was formed is given by

$$t_n = nT_s + g(n)T_s \tag{3-2}$$

where the first term expresses the constant time step and $g(n)$ will be called the time base distortion (TBD) at observation position n (scalar quantity). The TBD is the difference between the dotted and the full line on the diagonal of Figure 1. Note that at this stage no

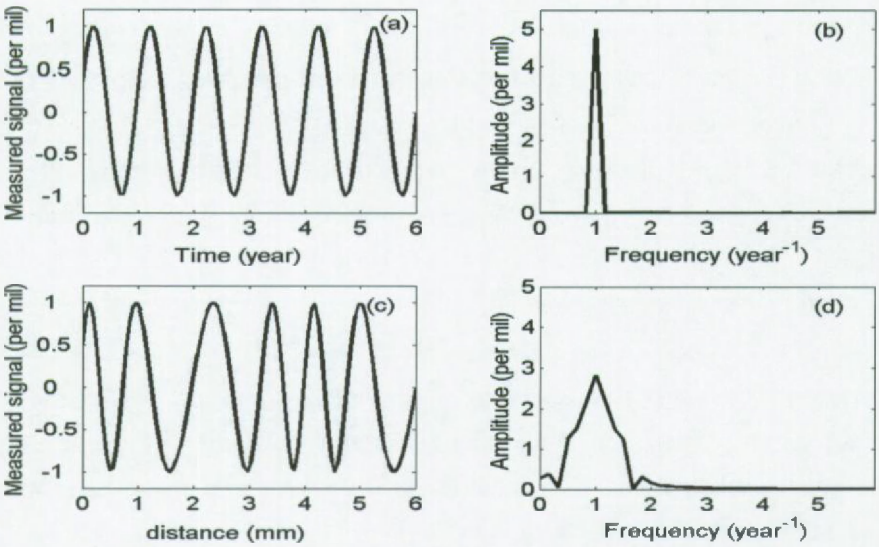

Figure 2: Deformation of the Fourier spectrum, due to variations in accretion rate: (a) sinusoidal signal as function of time, (b) its spectrum, (c) the measured signal as function of the distance (d) its spectrum assuming an constant accretion rate of one (mm/year). Notice that the TBD information is found around the harmonic and in its side peaks.

assumptions are made about the TBD: any possible variation in the growth can be expressed by equation (3-2). In the next paragraph we explain how the TBD can be estimated.

3.3.2 Identification of the time base distortion

In order to estimate the TBD, $g(n)$, we have to add an additional assumption since the data alone carry not enough information to determine the TBD uniquely. Many possibilities exist. In this chapter we assume that the sampled signal is periodic, i.e. recurring or reappearing from time to time. Periodic records can most easily be interpreted in a Fourier spectrum. Figure 2a shows a signal as function of time. The corresponding Fourier spectrum is shown in Figure 2b: one peak appears at a frequency of 1 year^{-1}. If a TBD is introduced, e.g. due to variations in growth rate, the measured record could look like Figure 2c. If a constant growth rate of one mm/year is used to construct a time base, the spectral interpretation becomes much more complicated (Figure 2d): small side peaks appear symmetrically around a broadened central peak. How the time base distortion can be reconstructed from the

broadened peaks and side peaks is elaborated in the present chapter. The mathematical details of the procedure are given in appendix 3.A and also in [Verspecht, 1994; Schoukens *et al.*, 1997]. The method has some important limitations that need to be detailed before we can interpret the results: the time base is separated by isolating a window around the first harmonic in the spectrum. The larger this window is, the more detailed the TBD will be. However, with an increasing window width the influence of the stochastic noise on the TBD will increase as well. If the width of the selected window is too large, such that side peaks of the second harmonic are included, these will be misinterpreted as a distortion of the first harmonic.

3.4 Comparison with the anchor point method

The anchor point method [Jones, 1983; Paillard *et al.*, 1996], assumes that the date of some observations is known. Next, the time instances of the observations in between these observations can be estimated by assuming a step wise constant accretion rate. This anchor point method is based on three assumptions which are compared with the assumptions made by the TBD (see also Table 1). Next a synthetic record is simulated in order to test and compare both methods.

Table 1: Comparison of the underlying assumptions made by the TBD and by the anchor point method.

	TBD	Anchor Point
1	record has a periodic component	the number of anchor points is known and each one is dated
2	two arbitrary observations are dated	first and last observation are dated
3	TBD is bandwidth limited	within two anchor points the growth rate is constant

3.4.1 Comparison of the assumptions

First assumption: in case of non-periodic records, the TBD method as proposed here, cannot be used. Therefore, we will focus on detailed data-records with sub-annual variations. In this situation, the anchor-point method makes one additional assumption for each additional year, i.e. the user has to define the timing of a (periodic) event in the record, e.g. mid-summer. Selecting these events is much more sensitive to noise than assuming a periodic record (see paragraph 3.4.2). The user has not only to decide on the timing of each event, but also the number of events. In noisy records, events can easily be overlooked, or a particular year can be mis-represented as two years. The latter problem becomes particularly acute when the record consists of multiple harmonics, which is often the case in tropical regions because of possible effects due to bimodal temperature and rainfall patterns.

Second assumption: the TBD method cannot assess the time span of a period. To translate this information into a time scale at least two events have to be dated. The anchor-point

method dates all events in between the first and last anchor point. Observations which are not in between the first and last anchor point cannot be dated.

Third assumption: due to the size of the window, used to separate the TBD, this TBD will be bandwidth limited. This means that high frequency variations in the TBD fall outside this window and will cause an error in the predicted growth rate. Another disadvantage is that the width of the window and thus the bandwidth of the TBD is chosen by the user and thus subjective. On the other hand, the anchor point method assumes a constant growth rate within two anchor points. Consequently, hiatuses or any other variation in growth rate between the selected anchor points will induce an error.

3.4.2 Comparison of both methods based on a simulation

The anchor point and TBD methods are compared using a synthetic data record consisting of 100 observations of a sinus with amplitude one and frequency one year^{-1}, the sample period is $T_s = 0.07$ year, so that exactly seven years are recorded. The time base distortion is chosen so that the corresponding growth rate slows down exponentially. The record is disturbed by Gaussian noise, such that the signal-to-noise ratio[1] equals ten. We have selected the maximum in each period as an annual anchor point. The window width, used to isolate the time base distortion, covered the spectral content in the region from $1 - 0.28 = 0.72$ year^{-1} to $1 + 0.28 = 1.28$ year^{-1}. The growth rate results obtained by both methods are compared with the actual synthetic growth rate (Figure 3). The first and last half year are excluded because the anchor point method cannot give growth rate information in these regions and the growth rates predicted by the TBD method are unreliable because of ringing [Schoukens *et al.*, 1997]. Both reconstructed growth rates more or less follow an exponential decrease, although the TBD method provides a better match to the simulated growth rate. The root mean square values of the difference between the calculated and simulated growth rates are 70 μ m year^{-1} for the anchor point method and 20 μ m year^{-1} for the TBD method. We have chosen this simulation to illustrate that even under optimal conditions (signal-to-noise-

1. The signal-to-noise-ratio is the proportion of the root-mean-square value of the noise-free signal over the root-mean-square value of the particular noise realization.

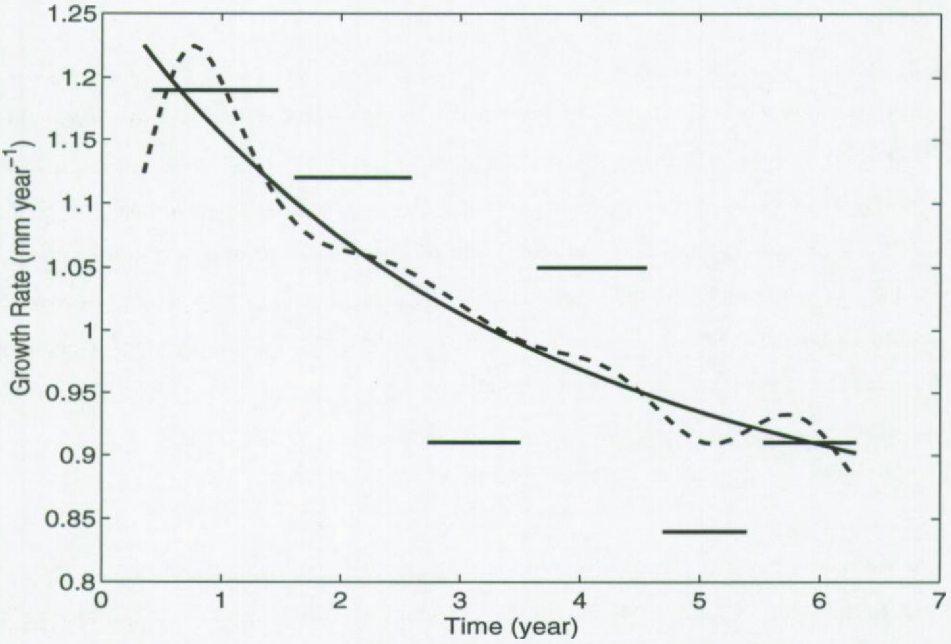

Figure 3: Simulated (full line) and reconstructed growth rates (dashed line: TBD method and horizontal line: anchor point method).

ratio of ten) the anchor point method can produce unreliable results. The reason for this is that the user has to decide which events are selected as anchor points. The noise on the measurements has a large influence on the selection of a particular event (e.g. a maximum or minimum in the record). Selecting the wrong event as anchor point will largely influence the growth rate and will consequently bias the time grid in the preceding and following period. The TBD method on the contrary bears on all observations, and hence the noise on a particular event will have a much smaller influence on the result.

3.5 Case study 1: the bivalve *Isognomon ephippium*

In this first example the TBD method is discussed in detail, to clarify the range of applications, benefits and shortcomings. First, the record is analyzed while assuming a constant growth rate. This way, the largest extrema in the spectrum can already be interpreted and the spectral contents, which carry information about growth rate can be shown. Next, results obtained via the TBD method and the anchor point method are compared. It is well known from literature that sea surface temperature is a major external control on Mg incorporation in calcite [Klein *et al*, 1996; Lazareth *et al*., 2003], and we will use the correlation between SST and the Mg-signal as a validation of the deduced time base: high Mg values correspond to high SST and vise versa.

3.5.1 Data record

The *I. ephippium* specimen was collected in August 1998 from Tudor Creek, nearby the city of Mombasa, Kenya. A more detailed description of the measurements and a discussion of other trace elements, as well as the analysis of other specimens of this species can be found in [Lazareth *et al*., 2003]. Monthly averaged values of sea surface temperature (SST) were obtained from the web site http://ingrid.ldgo.columbia.edu/SOURCES/.IGOSS/.nmc/.monthly/.sst as blended from ship, buoy and bias-corrected satellite data.

High-resolution spatial sampling of the calcite layer was performed with a Laser Ablation Inductively Coupled Plasma-Mass Spectrometer system from the umbo towards the growing tip (one laser shot approximately every 230 μm with a diameter of 60 μm). In this way, successively formed layers were sampled. Ca was used as internal standard and the Mg-signals were normalized (subtraction of the mean, and divided by the standard deviation).

3.5.2 A constant growth rate

Assuming the minima and maxima of the Mg-signal in *I. ephippium* reflect seasonal temperature cyclicity [Lazareth *et al*., 2003], an average growth rate of 4.8 mm year^{-1} (seven maxima over a distance of 33.5 mm) was calculated. The Mg-signal and the temperature

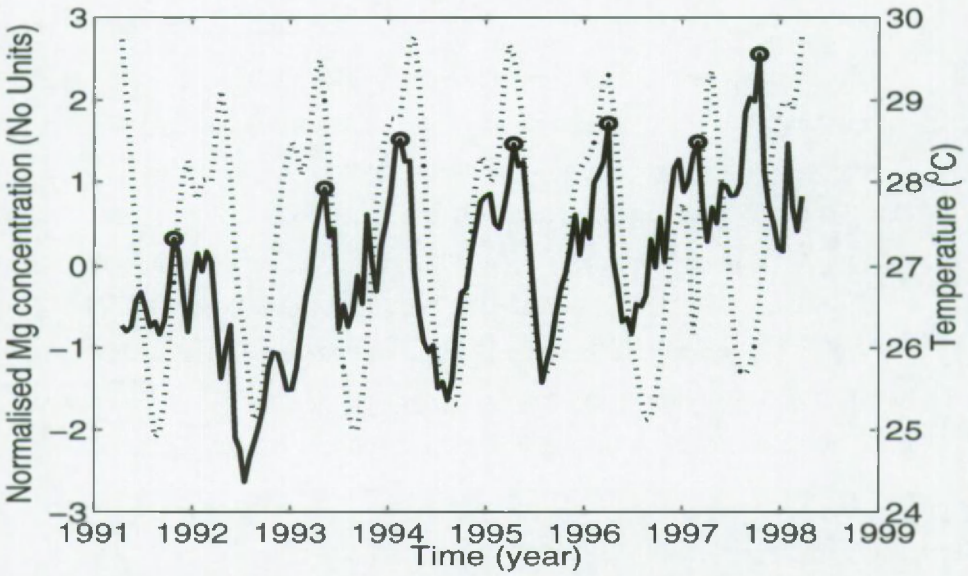

Figure 4: Mg-signal measured along the growth axis in an *Isognomon ephippium* (full line) versus instrumental SST (dotted line), assuming a constant growth rate (correlation coefficient R = 0.37). The 'o' indicates annual maxima that we have chosen to test the anchor point method.

record are shown in Figure 4. However, we noticed the occurrence of a second smaller Mg-peak just before the annual Mg maxima (Figure 4), which can also be noticed in the SST profiles. Despite these similarities, three peculiarities must be pointed out:

1. Although the smaller peaks appear too often at the same place to reflect stochastic noise, the peak correspondence between Mg-signal and SST is rather poor (Figure 4);
2. The Mg-concentration increases along the growth axis, suggesting some physiological control on Mg incorporation (e.g. growth rate);
3. From the observed deviation the distance between successive maxima in the Mg profile with those in the SST profile, it is clear that bivalve growth differed from year to year.

Assuming a constant growth rate, the correlation coefficient between Mg and SST is rather weak (R = 0.37 (n = 145, significance level $p < 0.01$); some peaks coincide, but others are shifted along the time axis). During the cooler seasons of 1993 and of 1997 the Mg-signal decreases, as it does in the other years, but the minimum does not reach as low. A possible

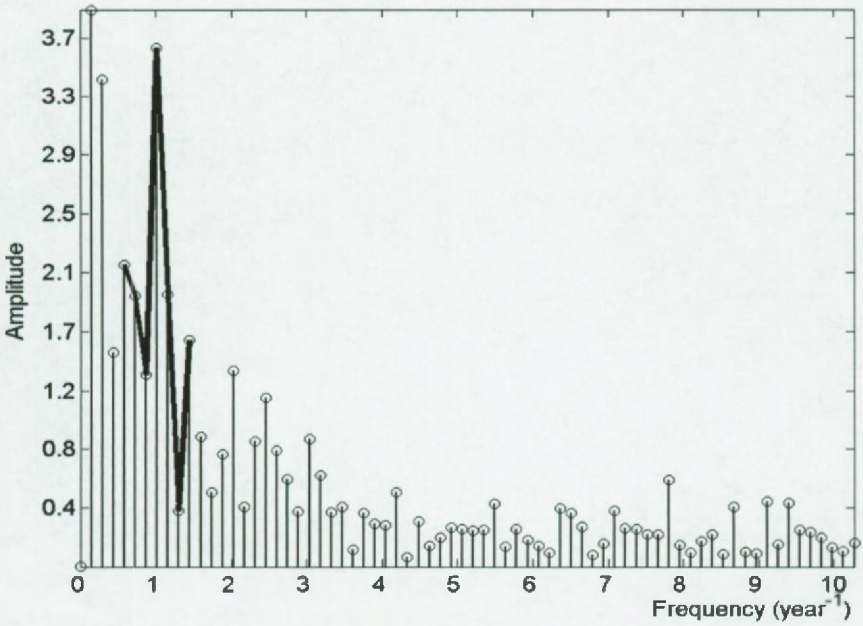

Figure 5: Discrete Fourier spectrum of the Mg-signal in *I. ephippium*. The window is represented by the full line, which covers the spectral contents from 0.57 to 1.43 year^{-1}.

explanation is that the specimen stopped growing during particular events and corresponding sequences of the proxy record are missing. Such offsets could result in for the poor correlation observed.

To summarize, the assumption of a constant growth rate is inadequate to predict relations on an inter-annual scale (even annual variations suffer from variability in growth rate).

3.5.3 Non-linear growth

The largest peak of the Mg-signal in Figure 5 reflects trends in the proxy record. The second largest peak of the Fourier spectrum is found at a frequency of one year^{-1} (assuming a constant growth rate of 4.8 mm year^{-1}). Several other peaks can be identified, viz. a peak at a frequency of approximately two year^{-1}, corresponding to variations with a six month period

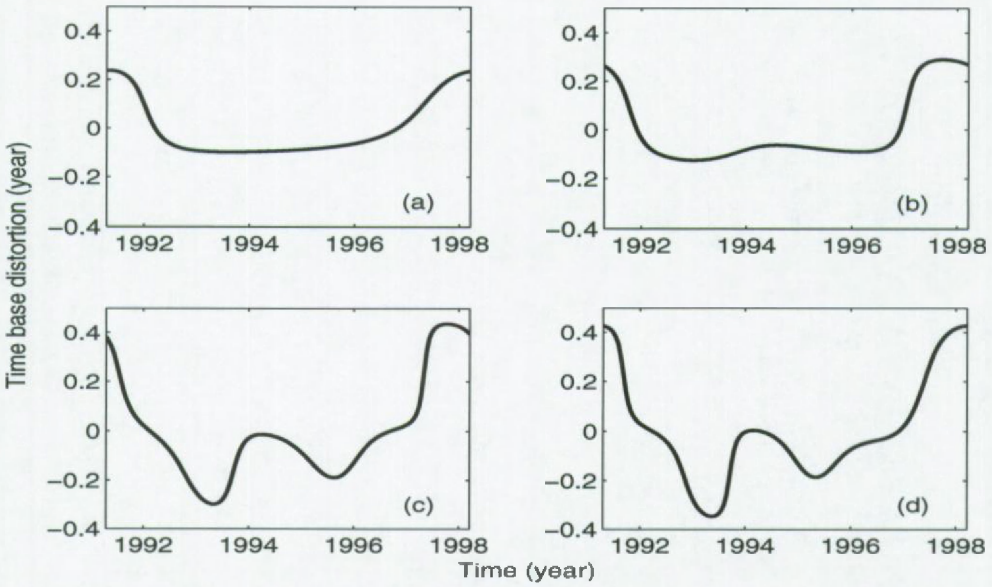

Figure 6: The TBD changes with increasing width of the window: (a) from 0.857 to 1.143 year^{-1}, (b) from 0.714 to 1.286 year^{-1}, (c) 0.571 to 1.429 year^{-1} and (d) 0.429 to 1.571 year^{-1}. Notice that the beginning and end match. This bias is due to the Gibbs phenomenon.

in the Mg-signal are present; even a third harmonic can be noticed. However, many other peaks in the spectrum remain unidentified and cause the weak correlation with SST.

The TBD method: the TBD is isolated by applying the windowing and phase demodulation technique. The window covered the spectral contents from 0.57 year^{-1} to 1.43 year^{-1}, i.e. the region covered by the thick solid line in Figure 5. Besides the data-record itself, the choice of this width is actually the only input demanded from the user. Figure 6 illustrates the influence of this choice: four plots of the TBD are shown with an increasing width of the window. Figures 6a and 6b do not show much variation in contrast to 6c and 6d. When selecting the window width, the user should keep in mind that

1. by increasing the window width to right, the possibility increases that the TBD around the second harmonic is erroneously included in the estimated TBD around the first harmonic;
2. a long term trend, if present, appears at the left side of the first harmonic. To cope with long term trends, we have de-trended the record before estimating the TBD.

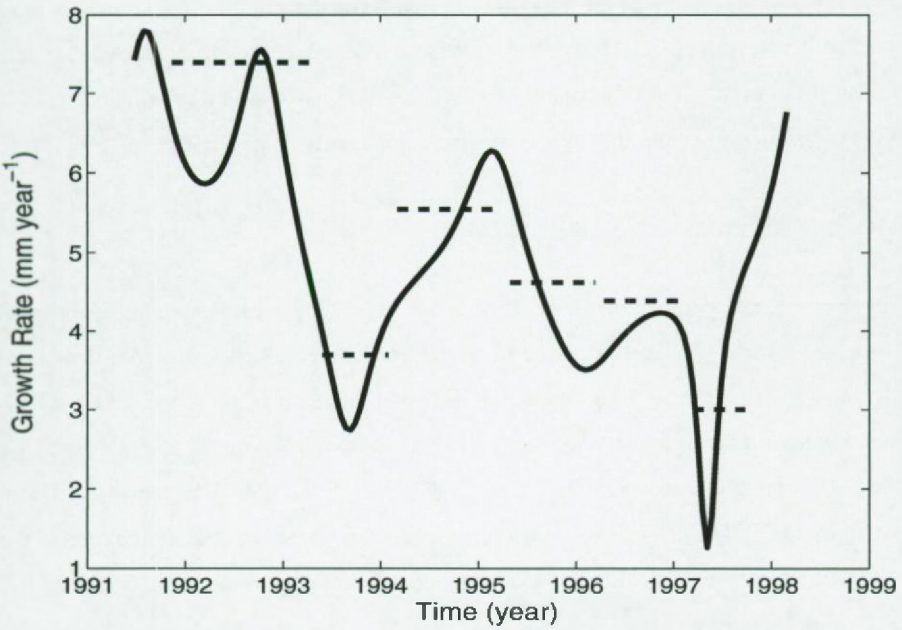

Figure 7: Growth rate of *I. ephippium* as function of the same time base as used in Figure 4: TBD (solid line) and anchor point method (dotted line).

Anchor point method: it is assumed that the peaks in the Mg-signal are annual and coincide with the peaks in the SST record. The peaks that we have selected are shown on Figure 4. Between two subsequent peaks a linear growth rate is assumed and the data before the first peak and after the last one are lost.

3.5.3.1 The reconstructed growth rate

Using equation (3-2) an improved growth rate can be calculated. Figure 7 shows the growth rates calculated by both methods. Although the TBD method changes the time grid by only 5%[1], the influence on the position of one single peak can be large. The TBD method provides a much more detailed profile of the variations in growth rate, but on the average,

1. Root mean square value of $g(n)T_s$ divided by the root mean square value of t, with g and t defined in equation (3-2).

both follow the same trend, i.e. a decrease in growth rate, disturbed at two occasions. The most remarkable feature of the TBD growth rate is the detection of two large dips in growth rate (cooler seasons of 1993 and 1997). It is possible that growth rate stopped completely during these events, but this cannot be discerned because of the limited band width of the window, which filters out the high frequency components in the TBD.

3.5.3.2 The correlation with SST

The discrete data points along the growth axis are no longer equidistant after correcting for growth rate variation. Therefore, the measurements were re-sampled before comparing this Mg record with the SST. Employing the TBD method, a correlation coefficient R = 0.60 is found (between the normalized Mg and SST record), while the anchor point method yields a correlation coefficient of R = 0.57 (n = 145, both $p < 0.01$). Although the difference is small, for the TBD method it was not a priori necessary to assume that all maxima of the Mg-record coincide with SST maxima. The anchor point method, on the contrary, a priori links in this example Mg-maxima to SST-maxima.

Figure 8 compares the re-sampled Mg-data on the re-constructed time grid using the TBD method with the SST record. Note that nearly all yearly maxima and minima in both records correspond one to one. This is highlighted when taking a closer look at the Fourier spectra of both the Mg and the SST record (Figure 9):

1. the annual frequency in the Mg remains, as expected, but the amplitude is larger (compare the amplitude in Figure 5 with that in Figure 9;
2. the peaks due to the TBD have disappeared and this not only around the first harmonic (which we used to calculate the TBD), but also around the second and third harmonic, although we have not used this information directly; this shows that the TBD acts identically on all harmonics as is expected from theory [Verspecht, 1994; Schoukens *et al.*, 1997]. This validates the initial assumptions that were made;
3. the second harmonic (bi-annual variations) stands out higher relative to the surrounding frequencies;
4. even a third harmonic can be distinguished from the noise;

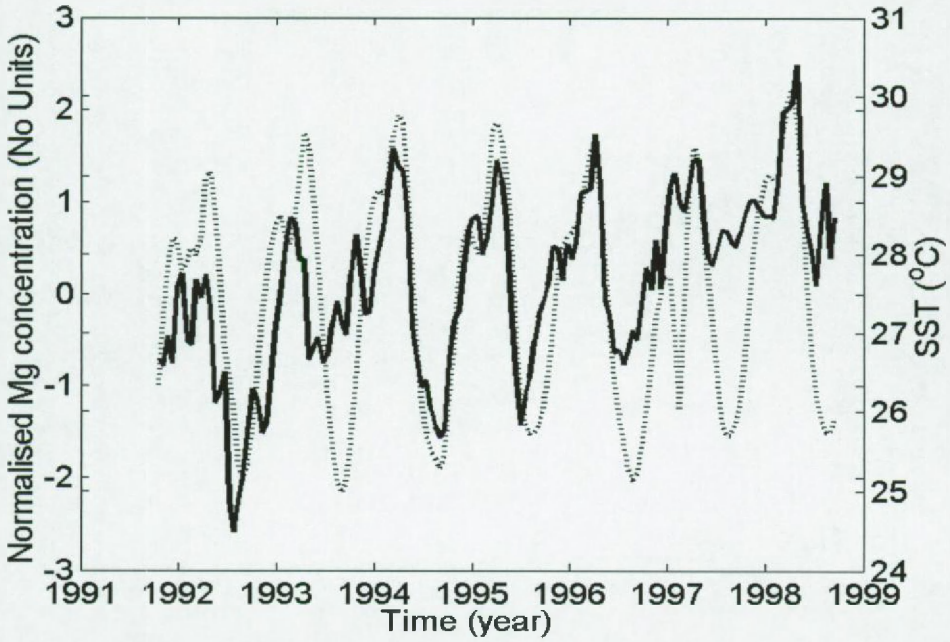

Figure 8: Mg-signal (solid line) on the improved time grid compared with instrumental SST (dotted line).

5. these three harmonics are also found in the spectrum of the SST;
6. the relative magnitude of the harmonics in the Mg-signal and in the SST correspond as well, which indicates a static linear relation between both [Pintelon and Schoukens, 2001a];
7. the Mg-record consists of a periodic signal, a trend and noise. The latter is now equally distributed over all frequencies, which means that the record is disturbed with white noise, probably due to measurement uncertainties and natural (short time) variations.

3.5.3.3 Trends in the record

There is no trend for SST but Mg clearly has one (Figure 8). This trend is isolated from the harmonics in the Fourier spectrum, and is expressed by the broad utmost left peak. A possible explanation is that an unknown process, additive to the SST, influences Mg (e.g. aging). When de-trending the Mg-record, the correlation with SST increases from $R = 0.60$ to $R = 0.76$. Thus most variation of the Mg-proxy can be explained by SST and a long term trend.

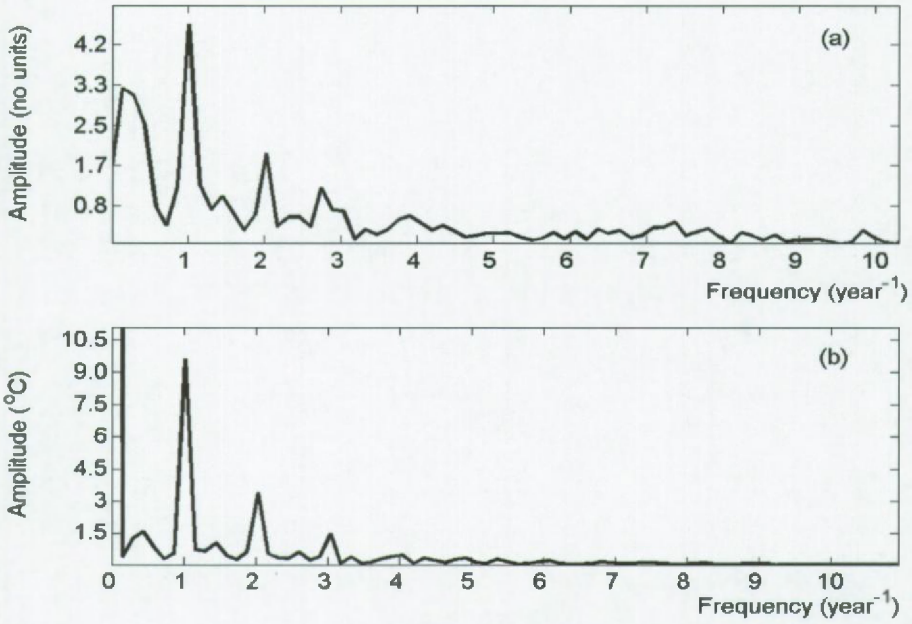

Figure 9: (a) Fourier spectrum of the Mg-signal on the improved time grid; (b) spectrum of the instrumental SST.

3.6 Case study 2: *Saxidomus giganteus*

In a second example the TBD method is illustrated on the stable oxygen isotope records ($\delta^{18}O$) measured in the aragonite shells of *S. giganteus* from Puget Sound (Washington State, U.S.A.) in September 2001. For a detailed description of the experimental setup and interpretation of the results, we refer the interested reader to [Gillikin *et al.*, subm.]. This paper discusses the problems related with building transfer functions between stable oxygen isotopes and temperature in the particular case where the salinity is hardly known. This site generally experiences a SST range of 8°C (8°C to 16°C) and has a salinity of 27.7±1.06 ‰, with occasional (less than yearly) periods of freshwater input reducing salinity to ca. 20 ‰. A thick-section of the aragonite shell was continuously sampled using a computer-controlled microdrill from growing tip to half way to the umbo. The carbonate powder (±100 μg) was processed using an automated carbonate device (Kiel III) coupled to a Finnigan Delta+XL. Data were corrected using an internal laboratory standard and are reported relative to V-PDB in conventional notation. Precision is generally better than 0.08 ‰.

Three different specimen were studied and the raw data are shown in Figure 10a to 12a (see [Gillikin *et al.*, subm.] for more details). Since the specimens come from the same sampling site, we can expect that they have similarly recorded environmental conditions. The correlation between the records can thus be used to compare the anchor point method with the TBD method presented here. The three records consist of 190, 123 and 55 observations (for clam 1, 2 and 3 respectively; clam 3 was not fully sampled) and covers periods of approximately five to nine years (assuming that periodic variations are annual). The dotted lines in Figure 10B to 12B show the power spectra of the records, assuming constant growth rate. The full line shows the spectrum after correction for TBD, using a window covering the spectral contents from 0.762 to 1.238 $year^{-1}$(Clam 1), 0.628 to 1.372 $year^{-1}$ (Clam 2) and 0.730 to 1.270 $year^{-1}$ (Clam 3). Clearly, TBD correction leads to a simplification of the Fourier spectra: ghost peaks, hiding the growth rate information have now disappeared, while annual periodicity is highlighted. No other harmonics appear in the spectra. The three $\delta^{18}O$-records on the TBD corrected time bases are shown in Figure 13. Similar features can

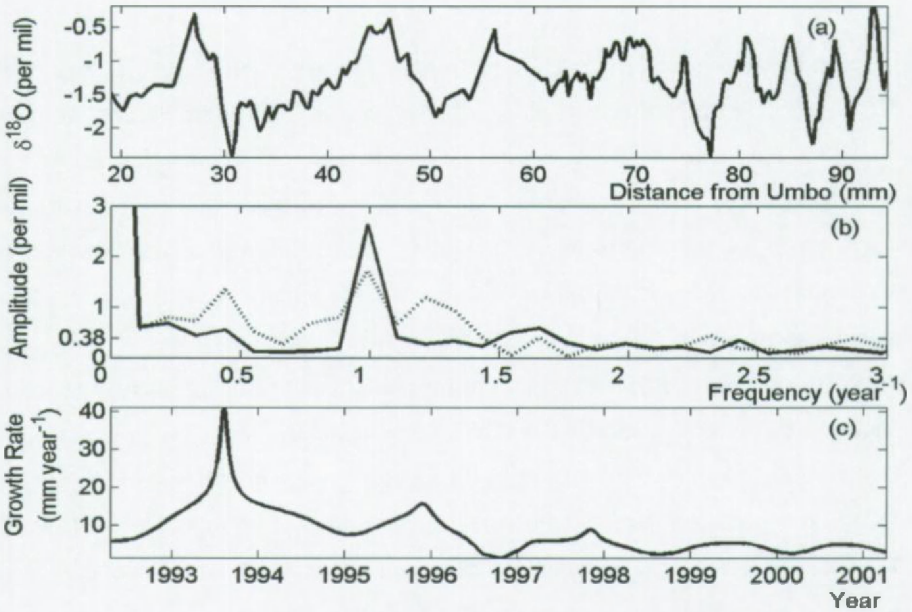

Figure 10: First *Saxidomus giganteus* from Washington state: (a) raw data, (b) Fourier spectra before (dotted line) and after the TBD method (full line), and (c) growth rate.

be found in all records, such as the broad relatively high minimum during summer 1998 and the much more negative values in summer 1997. These similarities are reflected by the high correlation coefficients between the three series (Table 2). In order to calculate these correlation coefficients the observations were interpolated on the time axis of clam 2. Using the correlation between the records as criterion, the time base estimated with the TBD method is in all cases slightly better then the one calculated with the anchor point method. The reason for this is that the anchor point method fixes the annual minima, while the phase

Table 2: The correlation coefficients, R, between the three *Saxidomus giganteus* stable isotope records sampled at the same site in Washington state.

Correlation, R	1 vs. 2	1 vs. 3	2 vs. 3
TBD	0.91	0.90	0.88
Anchor Point	0.81	0.79	0.85

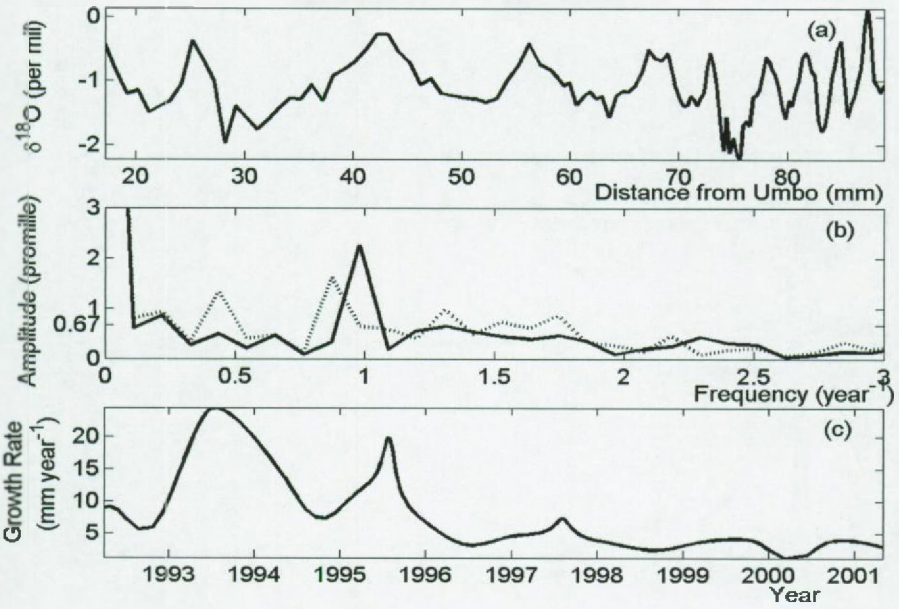

Figure 11: Second *Saxidomus giganteus* from Washington state: (a) raw data, (b) Fourier spectra before (dotted line) and after the TBD method (full line), and (c) growth rate.

demodulation restores the periodic component at all observations. Clearly, the TBD method outperforms the anchor point method.

In Figure 10C to 12C, the growth rates of the samples are shown. The TBD method can give a more refined growth rate profiles. Similar trends can be detected in the growth rate profiles in all three specimens: on average, growth rate decreases with time, while in between 1997 and 1998 growth rate was high for the three clams, while it was low in late 1996 and early 2000 for clam 1 and 2. Again we emphasize that the width of the window limits the variations in growth that can be detected by this method. This means that detailed variation on sub-annual scale is out of reach in this example. To solve this problem, parametric growth rate models can be used [Schoukens *et al*, 1997].

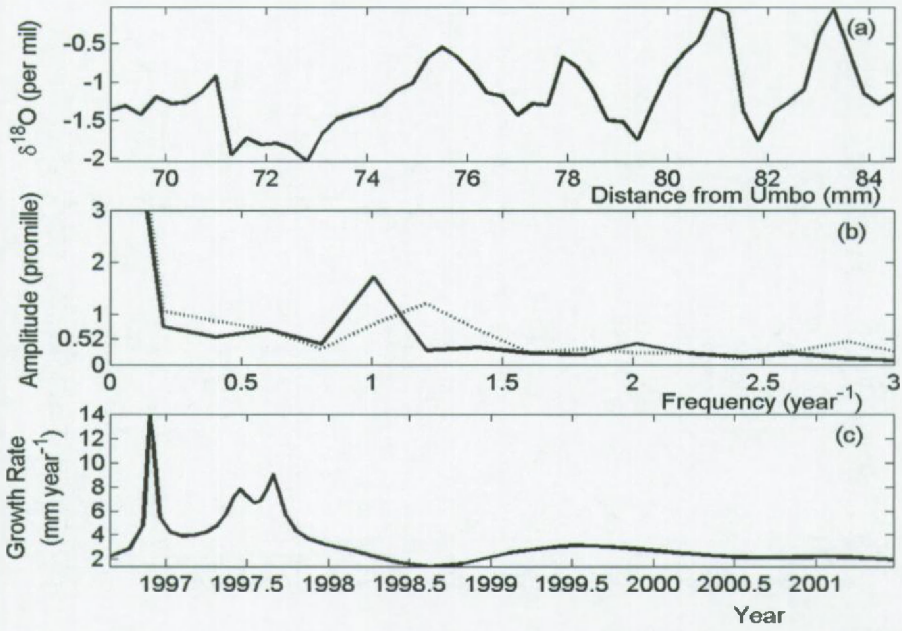

Figure 12: Third *Saxidomus giganteus* from Washington state: (a) raw data, (b) Fourier spectra before (dotted line) and after the TBD method (full line), and (c) growth rate.

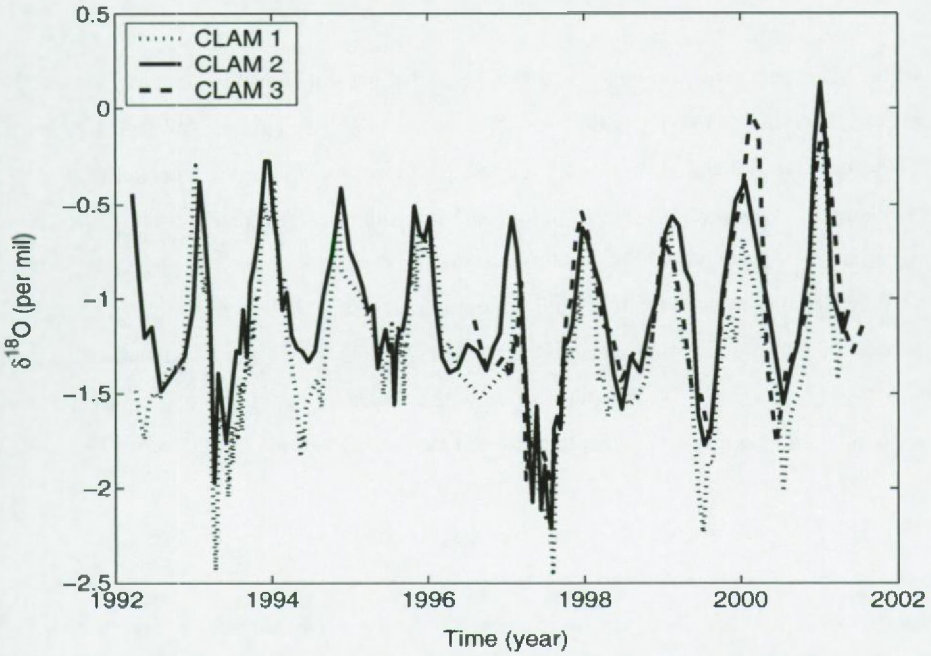

Figure 13: The three $\delta^{18}O$-signals as function of the time, which was estimated by the phase demodulation method.

3.7 Case study 3: Vanuatu coral stable isotope record

In this paragraph we illustrate how the TBD method can be used as a refinement of an existing time base and how an unknown frequency on a distorted time base can be identified. Therefore, we processed the Vanuatu coral $\delta^{18}O$-record, covering a period of 173 years, published by Quinn *et al.*, [1996] (additional information about the instrumentation and the data can be downloaded from http://www.ngdc.noaa.gov/paleo/). Ages were originally derived from annual density banding, combined with the anchor point method. This procedure may induce non-negligible errors in the time base, as we showed earlier for the simulation (paragraph 3.4.2). Assuming now that the record is periodic, the TBD, g, can be calculated and the timing of each observation can be refined.

The data-record is shown in Figure 14a and its spectrum in Figure 14b. Notice that the spectral information is spread along the frequencies 0.05 to 0.10 $year^{-1}$, which makes the identification of the underlying 14-15 year periodicity a difficult task. In this case study, we do not know in advance the distorted frequency around which the TBD can be estimated. Figure 14b shows several neighboring peaks and it is unclear which are caused by variations in accretion rate or by the hidden periodic variation in the record. To overcome this problem, the TBD was initially estimated around the largest peak in the region of interest (0.075 $year^{-1}$) using a window covering the spectral contents from 0.046 $year^{-1}$ to 0.104 $year^{-1}$. As shown in appendix 2, an incorrect initial estimate of the frequency will cause a linear trend in the TBD. The slope of this linear trend is proportional to the mismatch between the initially estimated frequency and the 'true' frequency. An iterative procedure is used to refine the frequency estimate until the optimal frequency is found (corresponding to a slope of zero). The Fourier spectrum corresponding to the improved time base is shown in Figure 14c: the record can be decomposed in a trend (the broad peak left in the spectrum) and one single peak with a frequency of 0.0693 $year^{-1}$ (i.e. a periodicity of 14.42 years). To validate the improvements made, the original and the TBD time base were first detrended (sixth order polynomial). Next a periodic signal was matched to the remaining signal [Pintelon and Schoukens, 1996; Pintelon and Schoukens, 2001a] and the uncertainty was estimated from the mismatch of these signals (assuming circular normally distributed white noise). This is visualized by the 95% uncertainty bound on the spectra [Pintelon and Schoukens, 2001b]. Refining the time base increases the signal-to-noise-ratio on the harmonic from 4.2 to 8.2.

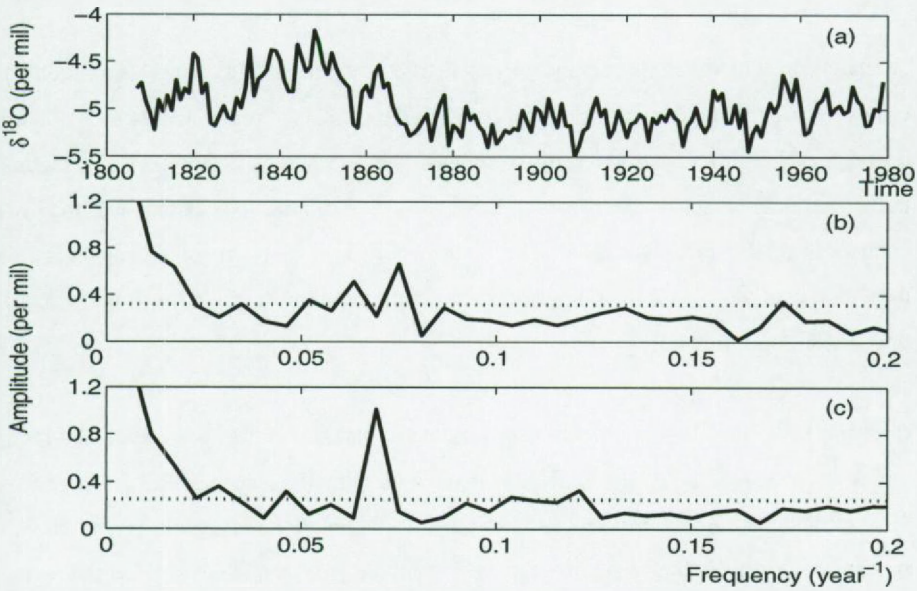

Figure 14: (a) The measured $\delta^{18}O$-signal of the coral in the time domain (before TBD correction); (b) the corresponding Fourier spectrum; (c) Fourier spectrum of the time base corrected growth rate. The dotted lines visualize the 95% uncertainty bounds.

To conclude, assuming that a periodic component with a frequency in the region 0.05 to 0.10 $year^{-1}$ is present in the record, the TBD is useful to refine the initially estimated time base. This leads to the identification of a single harmonic with a frequency of 0.0693 $year^{-1}$ (i.e. occurring at a periodicity of 14.42 years), which simplifies the spectral interpretation quite a lot.

3.8 Conclusion

We have presented a new method to reconstruct changes in the growth/accretion rate of proxy environmental records, based on the periodic content of proxy series. This approach is useful for a better reconstruction of the time series. The two weak points of this method are the constraints, that (i) the record has to have a periodic component and that (ii) the bandwidth of the TBD is limited. The main benefit is that it can reconstruct variations in growth or accretion rates, including short growth or accretion stops, in much more detail than the anchor point method.

Both methods were first compared using synthetic data. Next they were applied to the Mg-record in a bivalve, *I. ephippium,* from Kenya, where the Mg-record could be correlated with SST. In this example the benefits and limitations of both methods are illustrated. Next, three specimens of the clam *Saxidomus giganteus* sampled from the same site are processed and discussed. In this case the correlation between the three records along the estimated time bases is used to illustrate the better performance of the TBD method compared to the anchor point method. Finally, the TBD approach applied on the Vanuatu coral stable isotope record dated, using growth bands and the anchor point method. In this last case study, we have illustrated that the TBD method can also be applied to refine the time base estimated with the anchor point method. Our TBD approach performed well on a simulation, as well as on five real world examples. Interpretation of other proxy records could also benefit form this approach that enables to reconstruct a more detailed time base.

Appendix 3.A: estimating the time base distortion

The mathematical framework was first proposed in [Verspecht, 1994] and worked out in more detail in [Schoukens *et al.*, 1997]. The Fourier spectrum, S, of the record, s, can be calculated using a discrete Fourier transformation algorithm (DFT):

$$S(\omega) = \mathrm{DFT}[s(t)] \tag{3-3}$$

The TBD is isolated by windowing the spectrum, $S(\omega)$, with a rectangular window, $c(\omega)$, around the first harmonic. This cuts out an estimate, $\hat{G}(\omega)$, of the spectrum (for example, an estimate of the shifted spectrum $\hat{G}$ could be the triangle in Figure 2b).

$$G(\omega) = c(\omega)S(\omega) \tag{3-4}$$

with $c(\omega) = 1$ if ω is in the frequency range of interest (for example $\omega_0 - b < \omega < \omega_0 + b$, with b half the width of the window) and $c(\omega) = 0$ elsewhere. The calculation of an estimate of the TBD, $\hat{g}(n)$, is completed by employing the inverse discrete Fourier transformation (IDFT)

$$\hat{g}(n) = \frac{\varphi[\exp(-j(\omega_0/\omega_s)2\pi n)\mathrm{IDFT}(\hat{G})]}{\omega_0} \tag{3-5}$$

with $\varphi(u)$ the phase of u, $j = \sqrt{-1}$, ω_0 the fundamental radial frequency (2π/age (year^{-1})) and ω_s the radial sampling frequency $2\pi/T_s$, with T_s the sample period (average time gap between two subsequent observations). The exponential term shifts the frequency by $-\omega_0$.

Appendix 3.B: trends in the time base distortion

The argument of the Fourier series of the record is given by

$$\omega(nT_s + g(n)T_s) \tag{3-6}$$

where ω is the fundamental frequency, n is the observation number, T_s is the sample period and $g(n)$ is the TBD at observation n. If the fundamental frequency, around which the TBD manifests, is initially incorrectly estimated, this estimated frequency, $\hat{\omega}$, can be written as a scaled actual frequency, ω,

$$\hat{\omega} = (1 + \alpha)\omega \tag{3-7}$$

where α is the relative frequency error ($\alpha < 0$ means an underestimation of the fundamental frequency and thus an underestimation of the actual age of the specimen; $\alpha = 0$ means a

good estimation and $\alpha > 0$ means an overestimation and thus an overestimation of the actual age). Equation (3-7) is now inserted into equation (3-6) in order to express the TBD as a linear trend additive to another TBD

$$\begin{aligned} \hat{\omega}(nT_s + h(n)T_s) &= \omega(1+\alpha)(nT_s + h(n)T_s) \\ &= \omega(nT_s + \alpha nT_s + (1+\alpha)h(n)T_s) \\ &= \omega(nT_s + g(n)T_s) \end{aligned} \tag{3-8}$$

from which it follows that

$$g(n) = (1+\alpha)h(n) + \alpha n \tag{3-9}$$

where $h(n)$ is related to $g(n)$ as equation (3-9) shows that a miss-identification of the fundamental frequency ($\hat{\omega}$ instead of ω) leads to a linear trend (αn) in the TBD, $g(n)$. In the latter case the value of α can be used to improve the estimation of the fundamental frequency.

REFERENCES

[1] Fallon S.J., McCulloch M.T., van Woesik R. and Sinclair D.J., 1999. Corals at their latitudinal limits: laser ablation trace element systematics in porites from Shirigai Bay, Japan. Earth and Planetary Science Letters 172, 221-238.

[2] Finch A.A., Shaw P.A., Weedom G.P. and Holmgren K., 2001. Trace element variation in speleothem aragonite: potential for palaeoenvironmental reconstruction. Earth and Planetary Science Letters 186, 255-267.

[3] Gillikin D.P., Ulens H., De Ridder F., Elskens M., Keppens E., Baeyens W. and Dehairs F., submitted. Environmental and biological controls on oxygen and carbon isotopes in the aragonitic bivalve Saxidomus giganteus: implications for palaeoclimate studies. Submitted to Palaeogeography Palaeoclimatology Palaeoecology.

[4] Herbert T.D., 1994. Reading orbital signals distorted by sedimentation: models and examples. Spec. Publ. -Int. ass. Sediment. 19, 483-507.

[5] James J.F., A Student's Guide to Fourier Transforms with Applications in Physics and Engineering, Cambridge University Press, 2002.

[6] Jones D.S., 1983. Sclerochronology: reading the record of the molluscan shell. American

Scientist 71, 384-390.

[7] Kuhnert H., Pätzold J., Schnetger B. and Wefer G., 2002. Sea-surface temperature variability in the 16th century at Bermuda inferred from coral records. Palaeogeography, Palaeoclimatology, Palaeoecology 179, 159-171.

[8] Lazareth C. E., Willenz, P., Navez J., Keppens E., Dehairs F. and André L., 2000. Sclerosponges as a new potential recorder of environmental changes: Lead in *Ceratoporella nicholsoni*. Geology 28, 515-518.

[9] Lazareth C.E., Vander Putten E., André L. and Dehairs F., 2003. High-resolution trace element profiles in shells of the mangrove bivalve *Isognomon ephippium*: a record of environmental spatio-temporal variations? Estuarine, Coastal and Shelf Science 56, 1-12.

[10]Marshall J.F. and McCulloch M.T., 2002. An assessment of the Sr/Ca ratio in shallow water hermatypic corals as a proxy for sea surface temperature. Geochimica et Cosmochimica Acta 66, 3263-3280.

[11]Martinson D.G., Pisias N.G., Hays J.D., Imbrie J., Moore T.C. and Shackleton N.J., 1987. Age dating and the Orbital Theory of the Ice Age: Development of a High Resolution 0 to 300,000-year Chronostratigraphy. Quaternary Research 27, 1-29.

[12]Paillard D., Labeyrie L. and Yiou P., 1996. Macintosh program performs time-series analysis, Eos Transactions AGU, 77(39), 379.

[13]Petit J., Jouzel J., Raynaud D., Barkov N., Basile I., Bender M., Chapellaz J., Davis J., Delaygue G., Demotte M., Kotlyakov V., Legrand M., Lipenkov V., Lorius C., Pépin L., Ritz C., Saltzman E. and Stievenard M., 1999. 420,000 years of climate and atmospheric history revealed by the Vostok deep Antarctic ice core, Nature, 399, 429-436.

[14]Klein R.T., Lohmann K.C., and Thayer C.W., 1996. Bivalve skeletons record sea-surface temperature and delta O-18 via Mg/Ca and O-18/O-16 ratios. Geology 24[5], 415-418.

[15]Pintelon R. and Schoukens J., 1996. An improved Sine-Wave Fitting Procedure for Characterizing Data Acquisition Channels. IEEE transactions on instrumentation and measurement, 45(2), 588-593.

[16]Pintelon R., and Schoukens J., 2001a. System Identification: A Frequency Domain Approach, IEEE Press.

[17]Pintelon R. and Schoukens J., 2001b. Measurement of Frequency Response Functions Using Periodic Excitations, Corrupted by correlated Input/Output Errors. IEEE transactions on instrumentation and measurement, 50(6), 1753-1760.

[18]Quinn T.M., Crowley T.J. and Taylor F.W., 1996. New stable isotope results from a 173-year

coral record from Espiritu Santo, Vanuatu. Geophysical Research Letters, 23[23], 3413-3416.

[19]Schöne B.R., 2002. A 'clam-ring' master-chronology constructed from a short-lived bivalve mollusk from the northern Gulf of California, USA. The Holocene 13, 39-49.

[20]Schoukens J., Pintelon R. and Vandersteen G., 1997. A Sinewave Fitting procedure for Characterizing Data Acquisition Channels in the Presence of Time Base Distortions and Time Jitter. IEEE Transactions on Instrumentation and Measurement 40 (4), 1005-1010.

[21]Sinclair D.J., Kinsley L.P.J. and McCulloch M.T., 1998. High resolution analysis of trace elements in corals by laser ablation ICP-MS. Geochimica et Cosmochimica Acta 62 (11), 1889-1901.

[22]Vander Putten E., Dehairs F., André L. and Baeyens W., 1999. Quantitative in situ microanalysis of minor and trace elements in biogenic calcite using infrared laser ablation – inductively coupled plasma mass spectrometry: a critical evaluation. Analytica Chimica Acta 378, 261-272.

[23]Vander Putten E., Dehairs F., Keppens E. and Baeyens W., 2000. High resolution distribution of trace elements in the calcite shell layer of modern *Mytilus edulis*: environmental and biological controls. Geochimica et Cosmochimica Acta 64 (6), 997-1011.

[24]Verheyden S., Keppens E., Fairchild I.J., McDermott F. and Weis M., 2000. Mg, Sr and Sr isotope geochemistry of a Belgian Holocene speleothem: implications for palaeoclimatic reconstructions. Chemical Geology 169: 161-144.

[25]Verspecht J., 1994. Accurate Spectral Estimation Based on Measurements with a Distorted-Timebase Digitizer. IEEE Transactions on instrumentation and measurement 43 (2), 210-215.

[26]Weedon G.P., 1989. The detection and illustration of regular sedimentary cycles using Walsh power spectra and filtering, with examples from the Lias of Switzerland. Journal of the Geological Society, London 146, 133-144.

[27]Wei G., Sun M., Li X. and Nie B., 2000. Mg/Ca, Sr/Ca and U/Ca ratios of a porities coral from Sanya Bay, Hainan Island, South China Sea and their relationship to sea surface temperature. Palaeogeography, Palaeoclimatology, Palaeoecology 162, 59-74.

CHAPTER 4

REDUCTION OF THE GIBBS PHENOMENON APPLIED TO NON-HARMONIC TIME BASE DISTORTIONS

Abstract - A sine wave fitting procedure for characterizing measurements of a periodic signal in the presence of additive noise and an unknown time base distortion is presented. If the time base distortion is modelled by a Fourier series, it suffers from the Gibbs phenomenon (ringing) at the borders of the data record. Usually, this is solved by ignoring data samples at the borders. The latter is unacceptable for very short data records where measuring a sample is (very) expensive and/or (very) time consuming. This chapter presents a solution that suppresses the ringing in the estimated time base distortion without ignoring data samples at the borders. The theory is illustrated on simulations and on real vessel density in the wood of a mangrove tree from Kenya (*Rhizophora mucronata*).

4.1 Context

In the previous chapter, the ideas and the time base distortion concept have been presented. However, two problems, which are already mentioned have to be treated: the Gibbs phenomenon or ringing, which causes a bias in the time base distortion near the borders in the time window and the choice of the window width in the frequency domain.

The method, as presented so far, is not able to solve these problems. Therefore, a parametric representation will be introduced. In such a parametric representation, the time base distortion will be expressed in some bases, like a polynomial or a Fourier series. It is not possible to predict in advance which basis will perform best. Dependent on the time base, it can be possible that specific bases will work much better than others. We have limited ourselves to orthogonal polynomials and Fourier bases.

In a Fourier basis, the Gibbs phenomenon is still present. The time base distortion is discontinue at the borders, which means that the left and right side in the time domain do not match. The basis functions in the Fourier basis are all continuous at the borders. The Gibbs phenomenon is consequently the error one makes by fitting continuous functions on a discontinuous profile. As long as the number of basis functions is lower than the number of observations, an error will be made. In engineering applications the samples near the borders are simply ignored to minimize this problem. Because of the shortness of the data-records used in climate studies and the costs of gathering samples, ignoring samples is not an option. Therefore, an alternative solution is proposed. This alternative is compared with the polynomial basis.

This work was performed on data gathered by Anouk Verheyden *et al.* [submitted].

4.2 Introduction and outline of the problem

The estimation of a harmonic signal in the presence of additive noise and an unknown time base distortion (T.B.D.) has been studied intensively in the literature [Verspecht, 1994; Schoukens *et al.*, 1997; Vandersteen et al, 2001; Stenbakken and Deyst, 1998a; Stenbakken and Deyst, 1998b]. Two different approaches can be distinguished. The first starts from a single observation (data record) of the harmonic signal and uses a parametric time base distortion model [Verspecht, 1994; Schoukens *et al.*, 1997; Stenbakken and Deyst, 1998b]. The second uses several well-chosen observations of the harmonic signal and uses a non-parametric time base distortion representation [Vandersteen et al, 2001; Stenbakken and Deyst, 1998a]. In general, the first approach suffers from ringing at the borders of the data record, which is not the case for the second approach. In [Schoukens *et al.*, 1997] the problem is solved by deleting samples at the borders. This chapter studies the estimation of an harmonic signal in the presence of additive noise and a time base distortion, starting from a (very) short data record, where measuring one sample is (very) expensive and/or (very) time consuming.

Such measurements can, for example, be encountered in tropical dendochronology. Tropical tree species often lack annual growth rings (see [Detienne, 1989]). Furthermore, even in tree species that display growth rings, ring width data, a measure for the productivity of the tree, may not always provide environmental information [Mushove *et al.*, 1995; February and Stock, 1998; Verheyden, 2004]. However, detailed anatomical measurements of wood have the potential to give information about the past environment in which the tree grew. For instance, changes in vessel density and diameters (vessels are tube-like cells responsible for the water transport from the roots to the leaves of a plant) may reflect changes in environmental conditions (e.g., [Verheyden, 2004; Sass and Eckstein, 1995; Astrade and Begin, 1997]). Detecting these changes is of major importance to understand forest dynamics and to ensure a sustainable management of the forests.

In this study, the vessel density in a mangrove tree was measured manually by counting the vessels employing a microscope, aided with image analysis software (Analysis 3.0). Counting the density along a radial axis from bark to pith takes about a full day of manual work (depending on the size of the tree) [Verheyden, 2004]. Besides the cost and time, the

data record can be fairly short, therefore ignoring data in order to reduce ringing is clearly not satisfactory. Other data-records suffering from the same problems, e.g. time base distortions due to changes in growth or accretion rate, include sclerosponges [Lazareth *et al.*, 2000; Rosenheim *et al.*, 2004], speleothems [Verheyden *et al.*, 2000; Finch *et al.*, 2001], corals [Sinclair *et al.*, 1998; Fallon et al, 1999; Wei et al, 2000; Kuhnert et al, 2002; Marshall and McCulloch, 2002], bivalves [Vander Putten *et al.*, 1999; Lazareth *et al.*, 2003], sediments [Weedon, 1989; Herbert, 1994] and ice cores [Petit *et al.*, 1999].

This chapter proposes a sine wave fitting procedure that suppresses ringing of the estimated parametric time base distortion model without data removal at the borders.

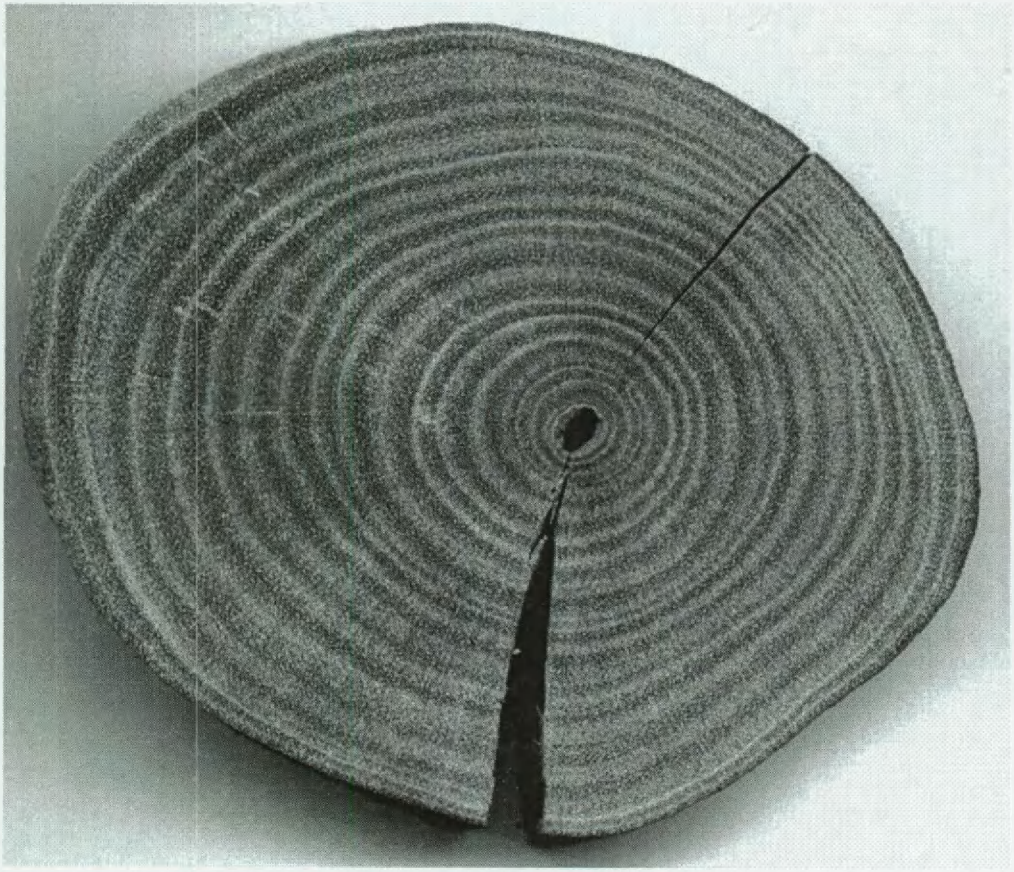

Figure 1: Cross section of a mangrove tree. The dark rings corresponds to low vessel density, while the light rings corresponds to high densities.

4.3 Sampling and methodology

A stem disc of the mangrove, R. mucronata, was collected in November 1999 from Gazi bay, Kenya (39.5°E, 4.4°S), located 40 km south of Mombasa. The sample is now part of the xylarium of the Royal museum for Central Africa (RMCA), Tervuren, Belgium (Tervuren wood collection, accession numbers: Tw55891).

The rainfall along the Kenyan coast shows a bimodal distribution, which is locally expressed in terms of the long rains (from April to July) and the short rains (from October to

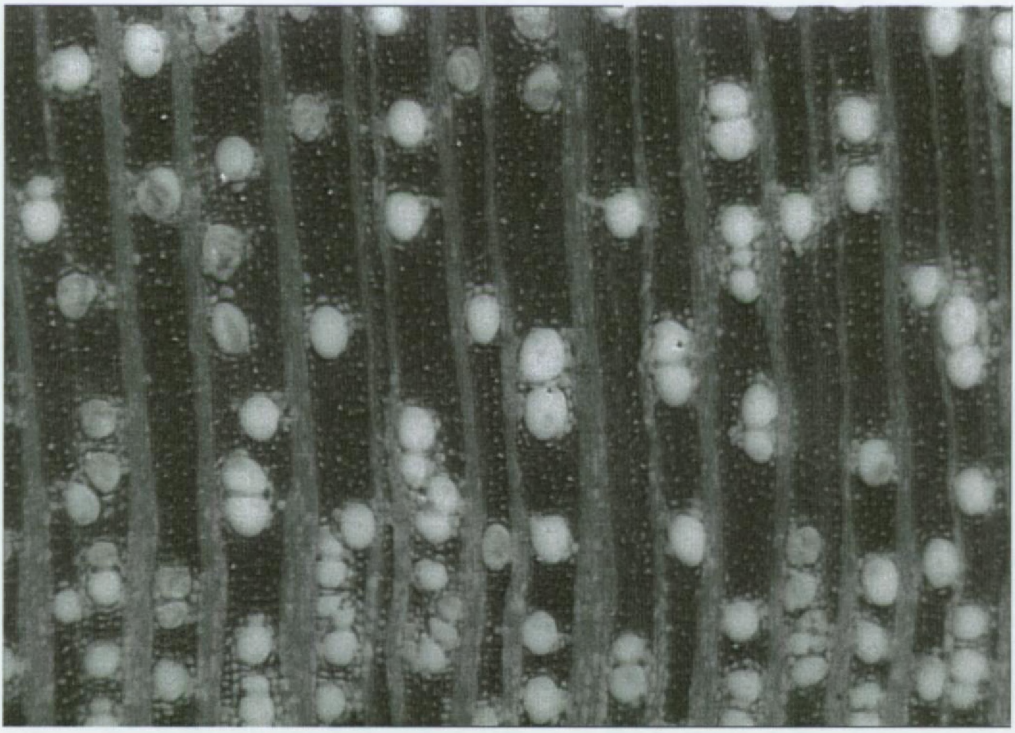

Figure 2: Microscopic view of the vessels.

November), with a mean annual precipitation of 1144 mm (1890-1985) [Lieth *et al.*, 1999]. The temperature ranges from 23.3 to 29.9 °C with a mean annual temperature of 26.4 °C (1931-1990) [Leith *et al.*, 1999].

The stem disc, Figure 1, was air dried and its transversal section was sanded (grain 100 to 1200). Prior to measurements, the sample was treated with white wasco-crayons, in order to enhance the delimitation of vessel elements. The vessel density was measured directly on the polished stem disc, along a radial transect from bark to pith in adjacent windows (Figure 2). Window size was set to 300 µm height and 2100 µm width. The number of vessels in each area were counted at an optical magnification of 12x using image analysis software (AnalySIS 3.0) and recalculated to the number of vessels per square millimeter.

4.4 The model

A sine wave fitting procedure [Pintelon and Schoukens, 1996] is combined with the T.B.D. identification technique of [Verspecht, 1994] in order to identify the harmonic content of the data. The signal model must describe the physical signal as closely as possible. This reduces the modeling errors and therefore increases the precision of the results. It is assumed that the continuous-time signal model, $s(t)$, for the measurements is given by

$$s(t) = A_0 + \sum_{k=1}^{h} A_k \cos(k\omega t) + A_{k+h} \sin(k\omega t) \tag{4-1}$$

where t is the time variable, A_0 is the offset, A_k and A_{k+h} are the unknown amplitudes of the kth harmonic ($k \in \{1, \ldots, h\}$) and ω is the unknown fundamental angular frequency. The analog signal, $s(t)$, is sampled at the time instances (see equation (3-2) on page 45)

$$t_n = (n + g(n))T_s \tag{4-2}$$

where $n \in \{0, \ldots, N-1\}$, with N the number of samples, $T_s = f_s^{-1}$ the mean sample period, f_s the sampling frequency and $g(n)$ the deterministic, unknown T.B.D. The time instance, t_n, is the time the specimen took to grow the sampling distance. Variations in growth rate will consequently cause variations in the sample moments, described by the value of the T.B.D., $g(n)$. In order to characterize this T.B.D., parametric models are used. In general the T.B.D. can be expanded in any set of (orthogonal) basis functions, $\phi_l(n)$,

$$g(n) = \sum_{l=1}^{b} B_l \phi_l(n) \tag{4-3}$$

where B_l are the unknown coefficients of the T.B.D. ($l \in \{1, \ldots, b\}$). A good choice of basis functions will limit the number of coefficients needed to describe the T.B.D. Note that this model is linear in the parameters, B_l, which simplifies the estimation considerably. Two sets of basis functions are discussed in this article, i.e. orthonormal polynomials and a Fourier series.

If the time base distortion is reconstructed for a record where an exponentially decreasing accretion rate (without much variation) is assumed, a polynomial model would probably perform better. On the other hand, if the accretion rate is influenced by climatological conditions, which vary mainly with a yearly periodicity, a Fourier basis would perform probably better. Considering that we do not know in advance which effect will overrule the other, both the polynomial and Fourier basis set are used. Hiatuses (stops in growth) are unlikely to occur in mangrove trees [Verheyden, 2004], but no additional constraints such as smoothness, continuity, etc.... have been implemented in order to keep the number of applications as wide as possible. In other occasions, when accretion hiatuses are present (e.g. [Vander Putten 2000]), B-splines or other sets can be more useful than the basis sets used in this chapter.

4.4.1 Orthonormal polynomials

The orthonormal polynomials are actually Legendre polynomials [Abramowitz and Segun, 1968], generated using the Gram-Schmidt orthogonalisation procedure starting from the polynomials $y = x^2, \ldots, x^{b+1}$ in the domain $x \in [0, N-1]T_s$. We have chosen orthogonal polynomials in order to avoid numerical problems in the estimation of the parameters. Note that the offset and the linear term are not used. The reason for this is that the corresponding coefficients cannot be identified uniquely:

1. the offset in the T.B.D. can be rewritten as a linear phase shift (time delay) in the signal model; and
2. the coefficient of the linear term acts as a shift of the fundamental frequency.

To show this, suppose an offset α and a linear trend $\beta n T_s$ in the T.B.D., $g(n) = \alpha + \beta n T_s + h(n)$, with $h(n)$ the remaining part of the T.B.D. The argument in equation (4-1) can be rewritten so that the offset and linear trend are replaced by a change in the angular frequency and a linear phase shift

$$\begin{aligned} k\omega[nT_s + g(n)T_s] &= k\omega[nT_s + \alpha T_s + \beta n T_s + h(n)T_s] \\ &= k\omega[1+\beta]\left[nT_s + \frac{h(n)}{1+\beta}T_s\right] + k\omega(\alpha T_s) \\ &= k\omega_1[nT_s + h_1(n)T_s] + k\omega_1\tau \end{aligned} \tag{4-4}$$

where $\omega_1 = (1+\beta)\omega$ is the shifted angular frequency, $\tau = \alpha T_s/(1+\beta)$ the delay, and $h_1(n)$ is the remaining T.B.D.

4.4.2 Fourier series and the Gibbs phenomenon

The parametric representation of the T.B.D. as a Fourier series is given by

$$g(n) = \sum_{l=1}^{b/2} B_l \cos(2\pi n l T_s) + B_{l+b/2} \sin(2\pi n l T_s) \qquad (4\text{-}5)$$

In the examples used in this chapter, no prior knowledge of the shape of the T.B.D. is available, so the polynomial model may be intuitively a more obvious set to describe the unknown T.B.D. Moreover, there is no reason why the T.B.D. model should be periodic. So, ringing will occur when equation (4-5) is used to approximate a non-periodic T.B.D.

In [Schoukens *et al.*, 1997] it is proposed to use only a part of the complete measurement period: some points near the borders are ignored during the fit. This reduces ringing considerably. As was pointed out, each measurement point is of high value, so dropping measurements is not ideal. For that reason an alternative is designed to reduce ringing in the T.B.D. model: extra samples are added at the borders. These extra samples will be ignored in the final estimation of the model parameters, but allow the procedure to reduce ringing inside the measurement window. Figure 3A shows a T.B.D. model, which is estimated in the initial measurement window. Notice the Gibbs phenomenon, which mainly acts near the borders. Figure 3B shows the same T.B.D., but this time some extra space is created outside the measurement window. Because ringing occurs mainly at the borders, it is exported mostly outside the measurement window. After the model is matched, these extra borders and most of ringing are dropped.

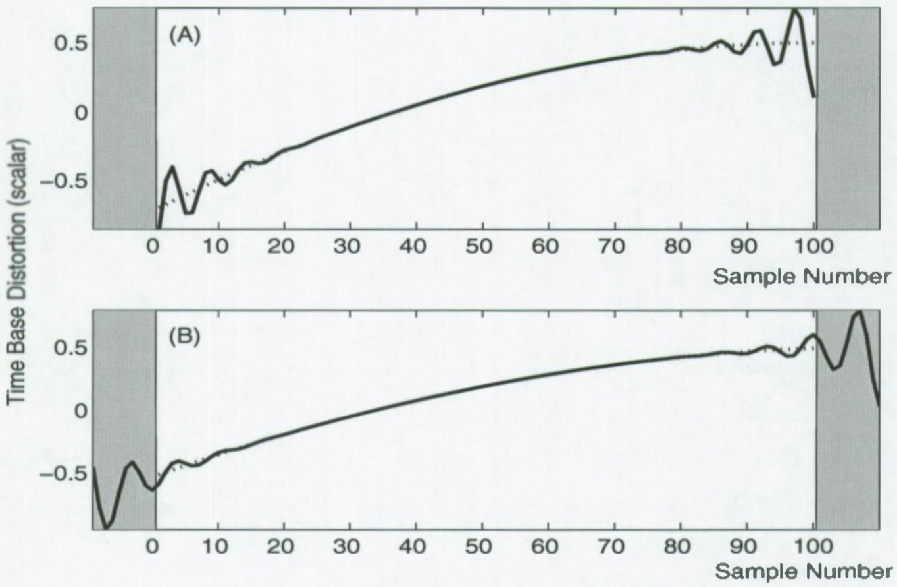

Figure 3: Conceptual illustration of the Gibbs phenomenon (ringing). Measurement window (white rectangular), enlarged borders (grey rectangulars), modelled T.B.D. (full line) and unknown T.B.D. (dotted line).

4.5 Estimation of the model parameters A_k, B_l and ω

4.5.1 Definition of the estimators

The parameters of the signal model and of the T.B.D. model are estimated simultaneously by minimizing a cost-function with respect to an extended parameter vector

$$\vartheta = [\omega, A_0, \ldots, A_{2h}, B_1, \ldots, B_b]^T \tag{4-6}$$

including not only the signal parameters, but also the T.B.D. model parameters. The optimal set of parameters in least square sense is given by

$$\hat{\theta} = \arg\min_{\theta} K_{\text{nls}}(\theta) \tag{4-7}$$

The explicit expression of the cost-function is

$$K_{\text{nls}} = \frac{1}{2}\sum_{n=1}^{\tilde{N}}\left[s(n) - \left(A_0 + \sum_{k=1}^{h} A_k \sin(k\tilde{\omega}t_n) + A_{k+h}\cos(k\tilde{\omega}t_n)\right)\right]^2 \tag{4-8}$$

where $s(n)$ is the nth sample, t_n is defined by the equations (4-2) and (4-3), A_k and A_{k+h} are defined in equation (4-1). For the optimization of (4-8) a Levenberg-Marquardt algorithm is preferred in order to improve the convergence. Further, each row in the Jacobian was scaled by its standard deviation in order to improve the numerical conditioning. When the polynomial or Fourier T.B.D. model with unchanged borders is used, $\tilde{N} = N$, with N the number of samples, and $\tilde{\omega} = \omega$, with ω defined in equation (4-1); when samples are dropped near each border $\tilde{N} = N_s$, with $N_s = N - 2q$ the remaining number of samples, and $\tilde{\omega} = \omega$; when the borders are enlarged with q samples at each border, $\tilde{N} = N$ and $\tilde{\omega} = \omega(1 + (2q)/N)$, with $N_L = N + 2q$. We will refer to the model where samples are dropped as the *reduced model* and to the model with enlarged borders as the *enlarged model*. Notice that if the fundamental frequency in the reduced model would have been rescaled, as is done in the enlarged model, ringing would be introduced again. Under the latter circumstances, no space is created for the Gibbs phenomenon to occur.

4.5.2 Starting value problem

The identification consists of a sequence of four steps, which are only summarized in this chapter. First we will discuss how the set of parameters is estimated when the measurement window is not altered (model with $\tilde{N} = N$ and $\tilde{\omega} = \omega$), i.e. we will focus on the polynomial and the Fourier T.B.D. model, where the borders are not altered. The estimation of the initial values is done in precisely the same way as in chapter 3:

1. a non-parametric estimation of the T.B.D. is performed: its influence acts as a phase modulation, characterized in the spectrum by the lines appearing around the harmonics. The T.B.D. can be isolated by employing a frequency window around the first harmonic in order to cut out an estimate of its spectrum, $\hat{G}$. The calculation of the estimate can be done by shifting the spectrum to D.C. and then by calculating the inverse Fourier transform.
2. The initial values of the coefficients are generated by matching the proposed model to this non-parametric estimate of the T.B.D.
3. Also, the sampling instances are calculated using equation (4-3).
4. An initial estimation of the amplitudes, A_k, and fundamental frequency, ω, of the signal is done by minimizing a least squares cost function, in the same way as is done in [Pintelon and Schoukens, 1996], but on the newly calculated sampling moments. Mostly, an initial estimation of the fundamental frequency can be done visually by counting the number of peaks in the signal.

For the reduced model, with $\tilde{N} = N_s$ and $\tilde{\omega} = \omega$, the estimation of initial values remains unchanged. In order to gather starting values for the enlarged model, with $\tilde{N} = N$ and $\tilde{\omega} = \omega \cdot N_L/N$, a larger measurement window is created, consisting of the original one, neighbored at the left by the q last observations and at the right by the q first observations. The procedure will still induce ringing and will thus not estimate the parameters correctly. Regardless, the initial values seem to be appropriate to make the procedure converge toward a good set of parameters that minimizes the cost function (equation (4-7)). For both the reduced and enlarged model, steps (2) to (4) remain unaffected.

4.5.3 Advantages and disadvantages of the different approaches

The main advantage of the polynomial model is that no ringing occurs, but the numerical conditioning is worse than for the other Fourier base based models. Therefore, when a detailed description of the T.B.D. is desired and a lot of basis functions are used, the Fourier models may be preferred. An extra feature of the Fourier models is that they can absorb an inaccurate estimation of the initial angular frequency, because this would cause a linear trend in the T.B.D. model (see, e.g., Figure 2 in chapter 8). The polynomial model cannot absorb this, because the linear coefficient is not estimated. The linear coefficient, β as defined in equation (4-4), can be removed when using a Fourier bases by changing the fundamental angular frequency ω to ω_1 :

$$\omega_1 = \omega(1 + \beta) \tag{4-9}$$

4.6 Results of study based on simulation

4.6.1 Comparison of the different models

In this section a harmonic signal, distorted by a non-periodic T.B.D. is simulated. No extra noise was added in order to concentrate on the systematic errors caused by ringing. First the polynomial model for the T.B.D. is used. Next the Fourier model is employed, without arrangements to reduce ringing. Finally the reduced and enlarged model are used. These different estimates for the T.B.D. are compared, employing the residual cost function as criterion. The harmonic signal is constructed by a sinus with frequency 8 Hz on a grid of 100 samples ($[1, \ldots, N]/N$, with $N = 100$). Because certain biota have an exponentially decreasing growth rate [Schöne et al, 2002; Gillikin *et al.*, submitted], an exponentially decreasing T.B.D. was chosen in the simulation. The offset and the linear trend were removed from the T.B.D.

One harmonic is used to reconstruct the signal, while the number of parameters used in the T.B.D. model was 14, regardless of the model (Fourier or polynomial). Figure 4 shows the difference between the estimated and true T.B.D. Comparison revealed that:

1. the polynomial model (order 15) was able to reconstruct the T.B.D. with hardly any error. A residual cost function of $4\ 10^{-16}$ was found.
2. the Fourier model, without any arrangements against the leakage converged towards a residual cost function of $1.2\ 10^{-2}$.
3. the reduced Fourier model, where 10% of the measurements were not used to match the model, had a residual cost function of $4\ 10^{-4}$. This value is already corrected for the reduction of samples: because now only 90 samples were used to match the model, the cost function is evidently lower [Pintelon and Schoukens, 2001]:

$$K_{\text{corr}} = K_{\text{nls}} \frac{N - n_\theta}{N_s - n_\theta} \tag{4-10}$$

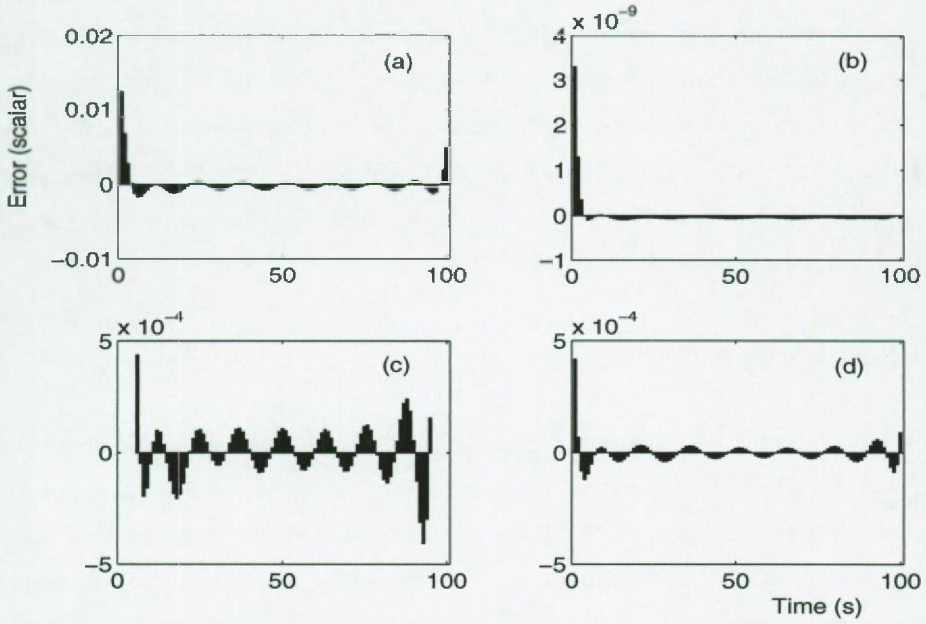

Figure 4: The absolute error between the estimated T.B.D. and the T.B.D. used in the simulation for (a) the Fourier model, (b) the polynomial model, (c) the reduced Fourier model and (d) the enlarged Fourier model.

where K_{nls} is the residual cost function, N the total number of samples (here: 100), N_s is the number of samples on which the model is matched (here: 90) and n_θ is the number of model parameters (here: 14+2+1). We can conclude that the ringing contribution to the residual cost function is reduced by a factor 30, but at the cost of removing 10 samples and hence lack of knowledge of the T.B.D. at these points.

4. when each border was enlarged with 10% of the total sample number (enlarged model), a residual cost function of 2.8 10^{-5} is found, which is 14 times smaller than in (3) and over 400 times smaller than (2). The benefit is that no samples were lost at the borders, but because the total sample window is enlarged, the frequency resolution increases, while the number of T.B.D. parameters, b, remains equal. Consequently, the bandwidth of the T.B.D. model becomes smaller (only the first lines were used).

To conclude, the best model seems to use the polynomial base instead of a Fourier series to expand to the T.B.D. In the polynomial model, no systematic errors were present and no

arrangements against ringing were necessary. When the T.B.D. is modelled by a Fourier series, without any arrangements, a significant misfit is found due to ringing. The samples near the borders are discarded, which reduces the residual cost function by a factor 30, but 10% of the observations are lost. The alternative is to enlarge the borders, while the model is still matched on the original time window. The latter has a slightly better performance: no samples are lost, the cost function is reduced by a factor 400, but the bandwidth of the T.B.D. model decreases proportionally to the enlargement.

4.6.2 Optimal size of the extra borders

If more lines are added at the borders, the ringing will become smaller. On the other hand the bandwidth of the T.B.D. model will decrease and this will cause a larger systematic error. So, what is the optimal number of the extra samples at the borders? A strategy is presented to find a rough estimate of the optimal size [Fletcher, 1991]. To avoid long calculation times, the residual cost function is calculated for three possible enlargements of the borders. A parabolic function is fitted on these three values and the minimum of this parabolic function is chosen as an optimal value (for example $p = \{1, 10, 20\}$). This procedure is illustrated on the simulation presented in the previous paragraph and seems to work well as long as the optimal value is between the initial values. Figure 5 shows the cost function as function of the border size, q. The optimal value for this case study is an enlargement of 26%, i.e. $q = 13$ (the number of extra samples at each border). The predicted values are distributed around this value with a mean of 13.6 samples and a standard deviation of 1.7 samples, depending on the chosen values of the enlargement. The actual value depends slightly on the choice of the initial values, but the value of the cost function hardly differs within the interval defined by twice the standard deviation, as can be seen in Figure 5. A good choice of the enlargement parameter, p, can reduce the cost function substantially: 60 dB in this simulation. In real examples, the systematic error caused by the limited bandwidth of the T.B.D. model can be more or less important. So for each case the enlargement parameter, q, has to be estimated again. When the bandwidth of the T.B.D. model is increased by increasing the parameter b in equation (4-3), the optimal value of q has to be re-estimated. Estimating both parameters can be done employing a model selection criterion (see chapters 5 and 6).

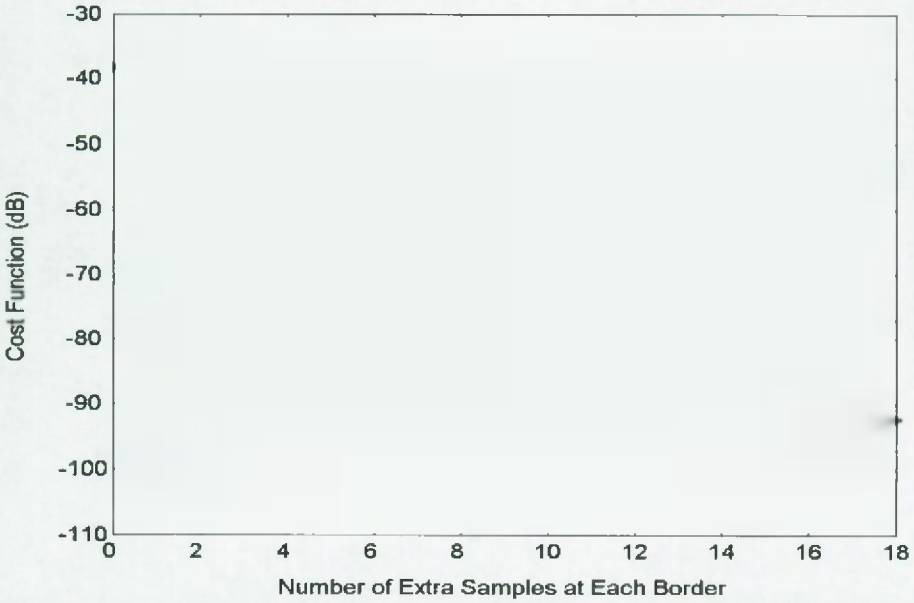

Figure 5: The cost function as function of the number of extra samples at the border. an optimum ('o') is reached when the decrease, caused by reducing the ringing, is compensated by an increase of the cost function, caused by the decreasing bandwidth.

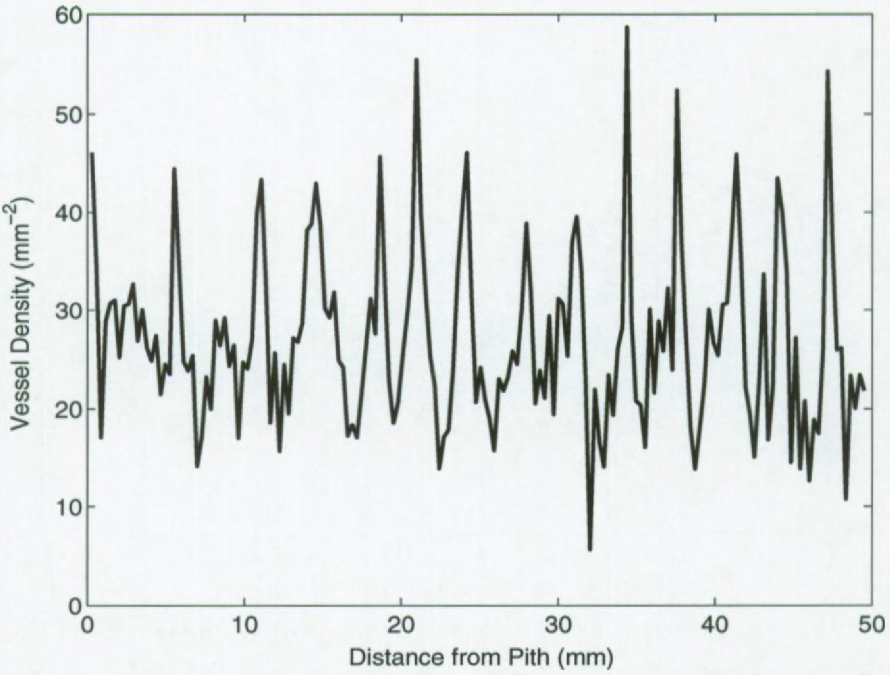

Figure 6: The vessel density of a mangrove tree as function of the distance from the pith.

4.7 Real world application

We now apply this procedure to a record of the vessel density measured along a radius from pith to bark of a mangrove tree from Kenya (*R. mucronata*). Vessel density may be related to environmental parameters such as temperature and/or precipitation[1] (see [February, 1998; Verheyden, 2004; Verheyden *et al.*, submitted]). We can therefore expect periodicity in the signal, which is shown in Figure 6. The data record is 170 samples long and covers approximately 15 years. Its spectrum is shown in Figure 7, where we have assumed a constant growth rate of 3.3 mm/year. The annual variation can be seen and has been

1. The beginning and ending of each season differs from year to year. These variations can be described by time jitter [Schoukens *et al.*, 1997], but this is not handled in this chapter.

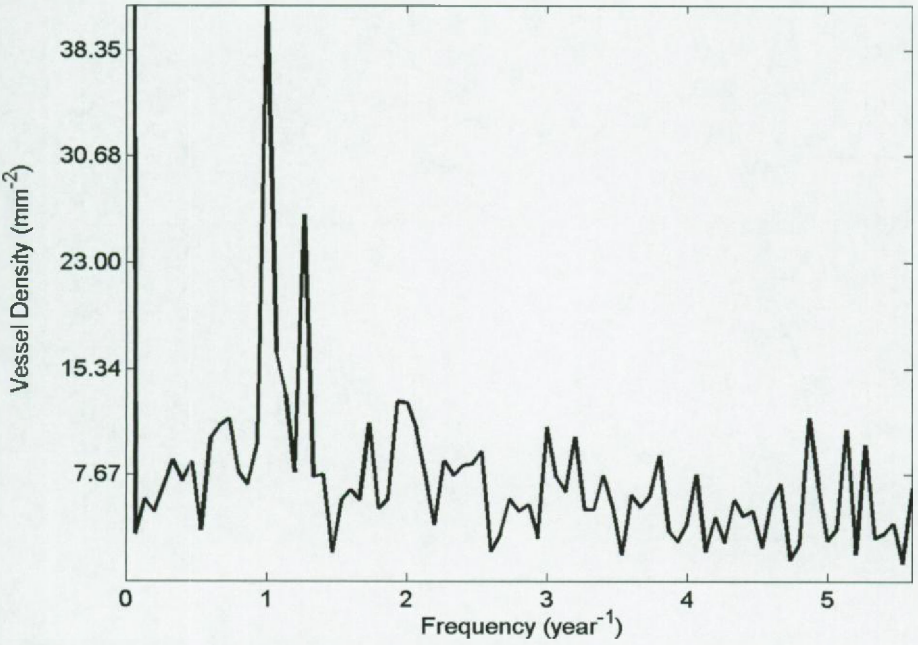

Figure 7: The spectrum of the vessel density, assuming a constant accretion rate.

confirmed [Verheyden, 2004], but one can wonder if the bi-annual climatic fluctuations present along the Kenyan coast (see methodology) are reflected in this signal. Growth rate variations make an interpretation of the spectrum hard. For instance, one can wonder what the second peak at 1.25 year^{-1} could mean. In order to give a correct interpretation of the spectrum, we have used the T.B.D. methodology.

A signal model consisting of three harmonics, A_k with $k \in \{1, 2, 3\}$, with a initial fundamental frequency of $\omega = 2\pi$ year^{-1}, is used; furthermore, the time base model employed 14 parameters, B_l with $l \in \{1, \ldots, 14\}$. This model complexity was ad hoc confirmed following the lines of chapter 6.

The residual cost function obtained from the Fourier model for the T.B.D. on the original data set is 2844 (no removal or addition of samples). The polynomial model with the same number of coefficients gives a slightly better result, viz. 2613. Employing the reduced model, where 5% of the measurements were dropped near each border ($q = 5$), decreases

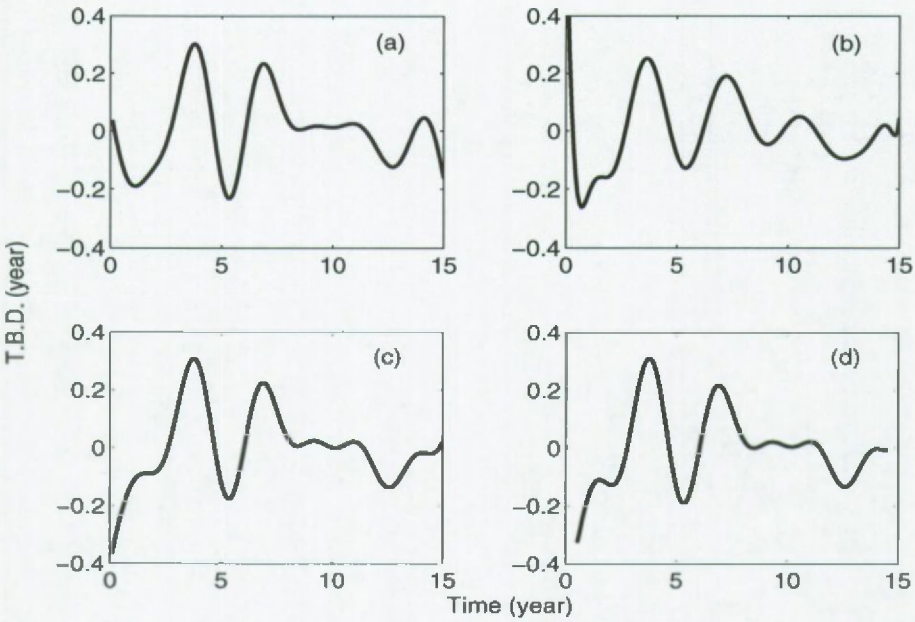

Figure 8: The different T.B.D.s: (a) a Fourier model with ringing, (b) a polynomial model, (c) Fourier model with enlarged borders and (d) a Fourier model with reduced borders.

the residual cost functions to 2387 (corrected by equation (4-10) for the loss of samples). The enlarged model, with the optimal value $q = 6$ achieves a similar result, viz. 2343. To conclude, ringing seems responsible for at least 17% of the residual cost function. When ringing is treated, the Fourier model seems to perform better than the polynomial model.

Comparing the different Fourier models (Figure 8), (a), (c) and (d), shows that ringing causes a severe difference in the estimation of the T.B.D., especially at the borders, e.g. at the left (c) curves down and (a) curves up, whereas the opposite appears at the right side. The general the shape of (c) and (d) (Fourier model) compared with (b) (polynomial model) is the similar, although local differences are present. Because of a lower cost function the Fourier models are preferred. (c) and (d), both Fourier models without ringing, are very similar, except that (c) also estimates the T.B.D. at the borders.

The objective of the construction of these time bases is the comparison with environmental parameters (see [Verheyden, 2004; Verheyden *et al.*, submitted] for more details). In order to

illustrate the method, the vessel density signal is compared with the maximum monthly temperature (^{o}C), monthly precipitation (mm) and mean monthly relative humidity (Table

Table 1: The correlation, R, between several environmental parameters and the vessel density, constructed on different time bases (the offset was removed before the correlation coefficients were calculated).

Correlation, R (%)	Maximum Temperature	Precipitation	Humidity
Constant growth rate	18	7	-28
Polynomial T.B.D. model	62	10	-56
Fourier T.B.D. model with ringing	66	12	-61
Fourier T.B.D. model with reduced borders	69	9	-61
Fourier T.B.D. model with enlarged borders	69	10	-62

1). Environmental data were obtained from the Kenya Meteorological Department in Mombasa. The raw vessel density data were interpolated on the newly reconstructed time frame. When no time base correction was performed the correlation with environmental parameters is weak. The correlation with the maximum temperature increases from 18% to typically 69% and from minus 28% to typically minus 62% with humidity when a time base correction is used. Without the time base correction, the misfit is largely due to misfits in the peak's position; with a time base correction the remaining misfit is mainly due to differences in the amplitude of the peaks (see Figure 9 and 10). Comparison of the different T.B.D. models in Table 1, shows that the differences in correlation are too small to distinguish between the different time base distortion models based on the correlation with environmental parameters. When the different environmental parameters are compared with the vessel density, it follows that both temperature and humidity seem to correlate reasonably well with the vessel density. The precipitation signal is much less regular, which is reflected by the weak correlation (see [Verheyden, 2004; Verheyden *et al.*, submitted]).

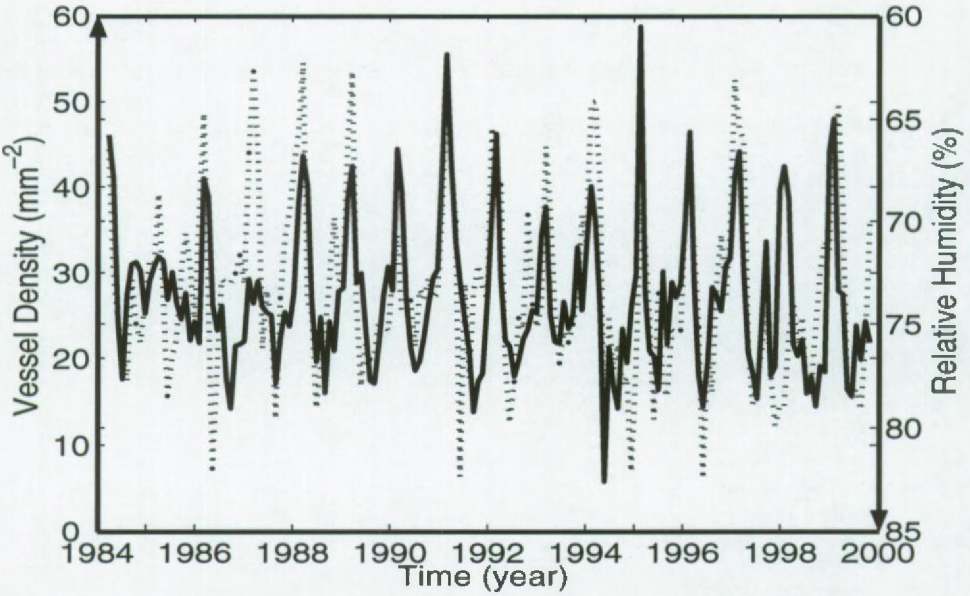

Figure 9: The vessel density on a corrected time grid (full line) versus the relative humidity (dotted line). To stress the anti-correlation between both the vessel density increases from the bottom to the top, while the relative humidity decreases.

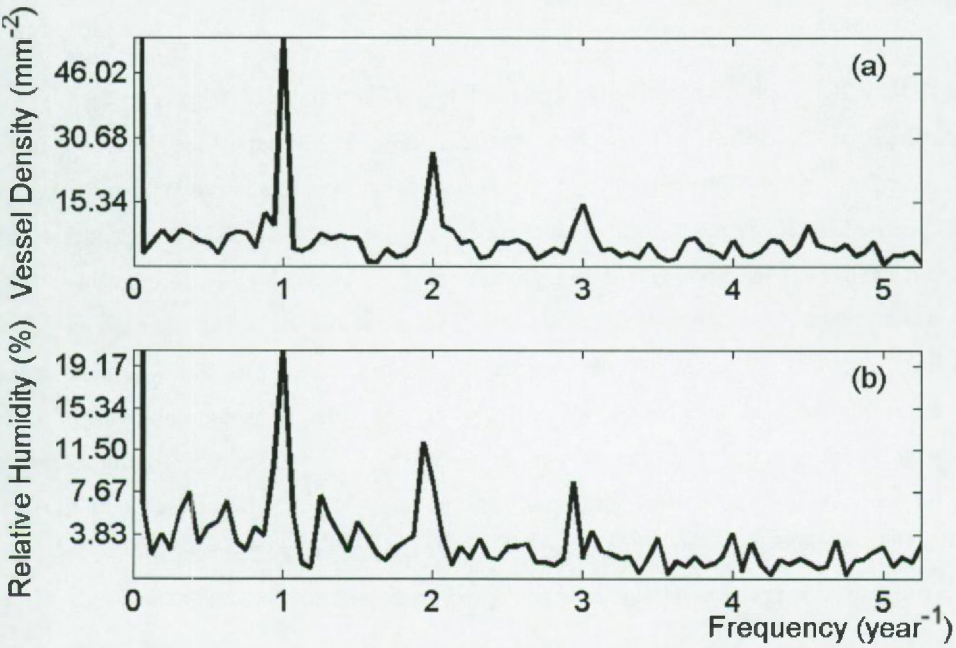

Figure 10: (a) The spectrum of the vessel density, corrected for variations in growth rate; (b) the spectrum of the relative humidity over the last 16 years.

4.8 Conclusion

A time base distortion approach was used to reconstruct time series, e.g. density in a mangrove tree. When a Fourier base was used to model the time base distortion, it was biased due to the Gibbs phenomenon. In order to circumvent this bias two adaptations were proposed. First, we used a polynomial model for the time base distortion. Under certain circumstances, like slowly varying growth rates, this model outperforms. However, a disadvantage of the polynomial model is that it becomes ill-conditioned when the number of basis functions increases. On the contrary, the Fourier series was still well conditioned. Secondly, we have adapted the model based on a Fourier basis by employing a model window that is larger than the measurement window. In these additional borders, ringing occurs. Next, we delete these borders and thus most of the bias. The advantage of enlarging the borders is that no samples are lost. The disadvantage is that the bandwidth of the time base distortion model is smaller. Employing more parameters can remedy this disadvantage.

In the simulation, an exponentially decreasing growth rate was chosen and the Fourier basis with ringing, the polynomial basis and two Fourier models remedied for ringing were compared. Therefore, we dropped some samples near the borders, as advised in [Schoukens *et al.*, 1997], and we enlarged the model window, as described in this chapter. Both solutions reduced the ringing considerably, but in this simulation, the polynomial model outperformed. Which basis matches better, depends upon the particular growth rate, which is unknown. In our botanical example, the Fourier model, remedied for ringing, matched better, resulting in a lower cost function.

The procedures discussed in this chapter are applicable to a wide array of short periodic data records, from both temperate and tropical regions, where all data points are critical. The work of validating and or calibrating environmental proxies often entails correlating short instrumental records (e.g., 2-10 years) with accretionary records (e.g., speleothems, corals, sclerosponges, molluscs) [Marschall and McCulloch, 2002Schone et al, 2002; Lazareth *et al.*, 2003; Rosenheim *et al.*, 2004; Gillikin *et al.*, submitted]. Such work should greatly benefit from the proposed procedure.

REFERENCES

[1] Abramowitz M. and Segun I.A., 1968. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover Publications, inc., New York.

[2] Astrade L. and Bégin Y., 1997. Tree-ring response of *Populus tremula L.* and *Quercus robur L.* to recent spring floods of the Saône River, France. Ecoscience 4(2): 232-239.

[3] Détienne, P., 1989. Appearance and periodicity of growth rings in some tropical woods. IAWA Bulletin 10(2): 123-132.

[4] Fallon S.J., McCulloch M.T., van Woesik R. and Sinclair D.J., 1999. Corals at their latitudinal limits: laser ablation trace element systematics in porites from Shirigai Bay, Japan. Earth and Planetary Science Letters 172, 221-238.

[5] February, E.C. and Stock W.D., 1998. An assessment of the dendrochronological potential of two *Podocarpus* species. Holocene 8(6): 747-750.

[6] Finch A.A., Shaw P.A., Weedom G.P. and Holmgren K., 2001. Trace element variation in speleothem aragonite: potential for palaeoenvironmental reconstruction. Earth and Planetary Science Letters 186, 255-267.

[7] Fletcher R., 1991. Practical methods of optimization, John Wiley and Sons Ltd., Chichester.

[8] Gillikin, D.P., Ulens H., De Ridder F., Elskens M., Keppens E., Baeyens W. and Dehairs F. Environmental and biological controls on oxygen and carbon isotopes in the aragonitic bivalve *Saxidomus giganteus*: implications for palaeoclimate studies. Submitted to Palaeogeography Palaeoclimatology Palaeoecology.

[9] Herbert T.D., 1994. Reading orbital signals distorted by sedimentation: models and examples. Spec. Publ. -Int. ass. Sediment. 19, 483-507.

[10]Kuhnert H., Pätzold J., Schnetger B. and Wefer G., 2002. Sea-surface temperature variability in the 16th century at Bermuda inferred from coral records. Palaeogeography, Palaeoclimatology, Palaeoecology 179, 159-171.

[11]Lazareth C.E., Willenz, P., Navez J., Keppens E., Dehairs F. and André L., 2000. Sclerosponges as a new potential recorder of environmental changes: Lead in *Ceratoporella nicholsoni*. Geology 28, 515-518.

[12]Lazareth C.E., Vander Putten E., André L. and Dehairs F., 2003. High-resolution trace element profiles in shells of the mangrove bivalve *Isognomon ephippium*: a record of environmental spatio-temporal variations? Estuarine, Coastal and Shelf Science 56, 1-12.

[13]Lieth H., Berlekamp J., Fuest S. & Riediger S., 1999. Climate diagrams of the world. CD-

Series: Climate and Biosphere (eds H. Lieth, J. Berlekamp, S. Fuest & S. Riediger) Blackhuys Publishers, Leiden, The Netherlands.

[14]Marshall J.F. and McCulloch M.T., 2002. An assessment of the Sr/Ca ratio in shallow water hermatypic corals as a proxy for sea surface temperature. Geochimica et Cosmochimica Acta 66, 3263-3280.

[15]Mushove P.T., Prior J.A.B., Gumbie C. and Cutler D.F., 1995. The effects of different environments on diameter growth increments of *Colophospermum mopane* and *Combretum apiculatum.* Forest Ecology and Management 72: 287-292.

[16]Petit J., Jouzel J., Raynaud D., Barkov N., Basile I., Bender M., Chapellaz J., Davis J., Delaygue G., Demotte M., Kotlyakov V., Legrand M., Lipenkov V., Lorius C., Pépin L., Ritz C., Saltzman E. and Stievenard M., 1999. 420,000 years of climate and atmospheric history revealed by the Vostok deep Antarctic ice core, Nature, 399, 429-436.

[17]Pintelon R. and Schoukens J., 1996. An improved sine wave fitting procedure for characterizing data acquisition channels, IEEE Trans. Instrum. Meas., 44, 588-593.

[18]Pintelon R. and Schoukens J., 2001. System Identification: a Frequency Domain Approach, IEEE PRESS, New York.

[19]Rosenheim, B.E., Swart P.K., Thorrold S.R., Willenz P., Berry L. and Latkoczy C., 2004. High-resolution Sr/Ca records in sclerosponges calibrated to temperature in situ. Geology 32(2): 145-148.

[20]Sass U. and Eckstein D., 1995. The variability of vessel size in beech (*Fagus sylvatica*) and its ecophysiological interpretation. Trees-Structure and Function 9 (5): 247-252.

[21]Schöne B., Lega J., Flessa K. and Goodwin D., Dettman D., 2002. Reconstructing daily temperatures form growth rates of the intertidal bivalve mollusk *Chione cortezi* (northern Gulf of California, Mexico), Palaeogeography, Palaeoclimatology, Palaeoecology, 184, 131-1462.

[22]Schoukens J., Pintelon R. and Vandersteen G., 1997. A Sinewave Fitting procedure for Characterizing Data Acquisition Channels in the Presence of Time Base Distortion and Time Jitter, IEEE Trans. Instrum. Meas., 46, 1005-1010.

[23]Sinclair D. J., Kinsley L. P. J. and McCulloch M. T., 1998. High resolution analysis of trace elements in corals by laser ablation ICP-MS. Geochimica et Cosmochimica Acta 62 (11), 1889-1901.

[24]Stenbakken G. and Deyst J., 1998a. Time-base nonlinearity determination using iterated sine-fit analysis, IEEE Trans. Instrum. Meas., 47 (5), 1056-1061.

[25]Stenbakken G. and Deyst J., 1998b. Comparison of time base nonlinearity measurement

techniques, IEEE Trans. Instrum. Meas., 47 (1), 34-39.

[26]Verheyden, A., De Ridder F, Schmitz N., Beeckman H. and Koedam N. Linking vessel density to climate variability in Rhizophora mucronata. Submitted to Trees-Structure and Function.

[27]Verheyden, A., 2004. Rhizophora mucronata as a proxy for changes in environmental conditions, a study of the wood anatomy, stable isotope chemistry and inorganic composition of a Kenyan mangrove species. Ph.D. thesis Vrije Universiteit Brussel, Brussels, Belgium, 228 p.

[28]Verheyden S., Keppens E., Fairchild I.J., McDermott F. and Weis M., 2000. Mg, Sr and Sr isotope geochemistry of a Belgian Holocene speleothem: implications for palaeoclimatic reconstructions. Chemical Geology 169: 161-144.

[29]Verspecht J., 1994. Accurate Spectral Estimation Based on Measurements with a Distorted-Timebase Digitizer, IEEE Trans. Instrum. Meas., 43 (2), 210-215.

[30]Vander Putten E., Dehairs F., André L. and Baeyens W., 1999. Quantitative in situ microanalysis of minor and trace elements in biogenic calcite using infrared laser ablation – inductively coupled plasma mass spectrometry: a critical evaluation. Analytica Chimica Acta 378, 261-272.

[31]Vander Putten E., Dehairs F., Keppens E., Baeyens W., 2000. High resolution distribution of trace elements in the calcite shell layer of modern Mytilus edulis: Environmental and biological controls. Geochimica et Cosmochimica Acta 64, 997-1011.

[32]Vandersteen G., Rolain Y. and Schoukens J., 2001. An Identification Technique for Data Acquisition Characterization in the Presence of Nonlinear Distortions and Time Base Distortions, IEEE Trans. Instrum. Meas., 50, 1355-1363.

[33]Weedon G.P., 1989. The detection and illustration of regular sedimentary cycles using Walsh power spectra and filtering, with examples from the Lias of Switzerland. Journal of the Geological Society, London 146, 133-144.

[34]Wei G., Sun M., Li X. and Nie B., 2000. Mg/Ca, Sr/Ca and U/Ca ratios of a porities coral from Sanya Bay, Hainan Island, South China Sea and their relationship to sea surface temperature. Palaeogeography, Palaeoclimatology, Palaeoecology 162, 59-74.

CHAPTER 5

MODIFIED AIC AND MDL MODEL SELECTION CRITERIA FOR SHORT DATA RECORDS

Abstract - The classical model selection rules such as Akaike information criterion (AIC) and minimum description length (MDL) have been derived assuming that the number of samples (measurements) is much larger than the number of estimated model parameters. For short data records AIC and MDL have the tendency to select too complex models. This paper proposes modified AIC and MDL rules with improved finite sample behavior. They are useful in those measurement applications where gathering a sample is very time consuming and/or expensive.

5.1 Context

In the previous chapter, a parametric representation of the time base distortion approach was discussed. The choice of the window size was replaced by the choice of the number of parameters in the time base distortion model. This choice is still arbitrary and thus subjective. In order to objectify this choice, model selection criteria are used. This should lead to the desirable situation that different people will draw the same conclusion from the same data. Therefore, a the concept of complexity is introduced. In this work, the complexity is defined by the number of parameters in the model. The more complex a signal or time base model is, the more parameters are needed to describe it. Within this context, the central question is: "What is the desired complexity?"

An additional complexity, besides that of the time base, is the number of harmonics in the signal model. In the frequency domain, the time base distortion manifests around each harmonic. If more harmonics are present, the time base distortion can be estimated around each of these. This has the same advantage as the averaging in the instrument calibration (Chapter 1): at the right side of the first harmonic the time base distortion is coded, perturbed by stochastic noise; at the left side, the complex conjugated is found, again perturbed by noise. The same pattern can be found around the first overtone, if present, and eventually around all others. So, the complicated task is to estimate the complexity of the time base distortion (what is significant and what is noise?), the number of harmonics and to optimize this model in order to filter of the stochastic noise.

A very important aspect of identification theory, is that we do not have to invent for each new case study new procedures to solve them. Therefore, we try to propose generally applicable rules, which guide us to the best solution. However, one of the most powerful rules, i.e. the model selection, has not been adapted for the special case, where the data-record is short, compared to the complexity of the model. Therefore, these rules have been refined. In order to remain in this general framework, and to illustrate the power of such rules, several completely different examples are solved.

5.2 Problem statement

An identification procedure typically consists in estimating the parameters of different models, and next selecting the optimal model complexity within that set. Increasing the model complexity will decrease the systematic errors, however, at the same time the model variability increases [Ljung, 1999; Pintelon and Schoukens, 2001]. Hence, it is not a good idea to select the model with the smallest cost function within the set because it will continue to decrease when more parameters are added. At a certain complexity the additional parameters no longer reduce the systematic errors but are used to follow the actual noise realization on the data. As the noise varies from measurement to measurement, the additional parameters increase only the model variability. To avoid this unwanted behavior the cost function is extended with a model complexity term that compensates for the increasing model variability. Summarized the model selection criterion should be able to detect undermodeling (= too simple model) as well as overmodeling (= too complex model). Undermodeling occurs when the true model does not belong to the considered model set. For example, unmodelled dynamics and/or nonlinear distortions in linear system identification, or too small a number of sinewaves and/or nonperiodic deterministic disturbances in signal modeling. Overmodeling occurs when the considered model includes the true model and is described by too many parameters.

Two popular model selection criteria are the Akaike information criterion, AIC [Akaike, 1974], and the minimum description length, MDL [Rissanen, 1978; Schwarz, 1978]. Under the hypothesis of Gaussian disturbances they take the form

$$\frac{K(\hat{\theta}(Z), Z)}{N} e^{p(n_\theta, N)} \quad \text{with penalty} \quad p(n_\theta, N) = \begin{cases} \dfrac{2n_\theta}{N} & \text{AIC} \\ \dfrac{\ln(N) n_\theta}{N} & \text{MDL} \end{cases} \tag{5-1}$$

when the noise variance is unknown [Ljung, 1999], and

$$\frac{K(\hat{\theta}(Z), Z)}{N}(1 + p(n_\theta, N)) \tag{5-2}$$

with $p(n_\theta N)$ defined in (5-1) for known or prior estimated noise variance(s) [Schoukens *et al.*, 2002]. In (5-1) and (5-2), $\hat{\theta}(Z)$ are the estimated model parameters

$$\hat{\theta}(Z) = \arg\min_{\theta} K(\theta, Z), \tag{5-3}$$

$n_\theta = \dim(\theta)$ the number of free parameters in the model, $K(\theta, Z)$ the quadratic-like cost function corresponding to the Gaussian maximum likelihood estimator

$$K(\theta, Z) = \frac{1}{2} e^T(\theta, Z) e(\theta, Z) \tag{5-4}$$

with $e(\theta, Z)$ the N by 1 vector of the (weighted) residuals of the model equation (= difference between measurements and model), Z the measurements, and $N = \dim(e)$ the number of data points. According to the AIC or MDL[1] criteria (5-1) or (5-2), the optimal model complexity is obtained by minimizing (5-1) or (5-2) over the given set of models. Note that (5-1) reduces to (5-2) if $p(n_\theta N) \ll 1$, which is mostly the case in system identification and signal modeling problems. For these cases the minimizers of (5-1) and (5-2) coincide.

The MDL rule has a much better reputation than the AIC rule: the MDL criterion gives strongly consistent estimates ($N \to \infty$) of the order of autoregressive moving average (ARMA) noise processes [Hannan, 1980], while the AIC criterion has a strong tendency to select too complex models (see [Pintelon and Schoukens, 2001], and Section 5.4). Both criteria are nevertheless inappropriate for short data records where the number of data samples N is not much larger than the number of parameters n_θ. Indeed, selecting a model with $n_\theta = N$ gives $K(\hat{\theta}(Z), Z) = 0$, and both AIC and MDL are exactly zero. Hence, the global minimum of (5-1) and (5-2) is attainted by the most complex models, which is of course totally undesirable. It explains why AIC and MDL select far too complex models for

1. In case of errors-in-variables (EIV) problems $\ln(N)$ in the MDL penalty (5-1) is replaced by $\ln(\text{rank}(C_n))$, where C_n is the covariance matrix of the noise on measurements Z (see [Pintelon and Schoukens, 2001], Chapter 17). For example, for single input single output systems within an EIV framework, $\text{rank}(C_n) = 2N$. Note that for multivariable EIV problems $\dim(Z) > \dim(e)$.

short data records (see Section 5.4). The same phenomenon has been observed in (non)linear regression problems with unknown noise variance [Sugiura, 1978; Hurvich and Tsai, 1989; McQuarrie *et al.*, 1997] and autoregressive time series modeling [Hurvich and Tsai, 1989; Broersen, 2000]. In [Broersen, 2000] tailor made finite sample AIC criteria for autoregressive-order selection are proposed, while in [Sugiura, 1978; Hurvich and Tsai, 1989] a corrected AIC rule for general (non)linear regression problems is given. The corrected AIC rule [Sugiura, 1978; Hurvich and Tsai, 1989] for problems with unknown noise variance has the form

$$\text{AIC}_\text{c}\text{: } \frac{K(\hat{\theta}(Z), Z)}{N} e^{p_c(n_\theta, N)} \text{ with penalty } p_c(n_\theta, N) = \frac{2(n_\theta + 1)}{N - n_\theta - 2} \tag{5-5}$$

Similarly to the classical AIC criterion (5-1) AIC_c has also the tendency to select too complex models (see [McQuarrie *et al.*, 1997] and Section 5.4). Therefore, the following improvement of (5-5) has been proposed in [McQuarrie *et al.*, 1997]

$$\text{AIC}_\text{u}\text{: } \frac{K(\hat{\theta}(Z), Z)}{N - n_\theta} e^{p_c(n_\theta, N)} \tag{5-6}$$

with $p_c(n_\theta, N)$ defined in (5-5). According to the sample size, AIC_c (very small) or AIC_u (moderate to large) is preferred [McQuarrie *et al.*, 1997]. Which model selection criterion to use also depends on the intended application: physical interpretation or prediction. In case of physical interpretation the criterion that selects most the true model should be preferred, while in case of prediction the criterion that minimizes the prediction error on validation data should be chosen. In the latter case the optimal model order may strongly depend on the sample size [Broersen, 2000]. In this paper we focus the attention to the first problem. Note that no finite sample criteria are available for estimation problems with known or prior estimated noise variance(s).

This paper proposes (i) the MDL equivalent of (5-5) for estimation problems with unknown noise variance, and (ii) small sample AIC and MDL rules for estimation problems with known noise variance(s) (see Section 5.3). The performance of the new rules is compared with the existing criteria on three simulation examples (Section 5.4) and one real measurement example (Section 5.5).

5.3 Modified AIC and MDL criteria

5.3.1 The new model selection rules

When the noise variance is unknown, the new MDL rule has the form

$$\text{MDL}_c\text{:}\quad \frac{K(\hat{\theta}(Z), Z)}{N} e^{p_c(n_\theta, N)} \quad \text{with penalty } p_c(n_\theta, N) = \frac{\ln(N)(n_\theta + 1)}{N - n_\theta - 2} \tag{5-7}$$

while for known or prior estimated noise variance(s) we propose

$$\frac{K(\hat{\theta}(Z), Z)}{N - n_\theta}(1 + p_s(n_\theta, N)) \quad \text{with penalty } p_s(n_\theta, N) = \begin{cases} \dfrac{2n_\theta}{N - n_\theta} & \text{AIC}_s \\ n_\theta \dfrac{\ln(N)}{N - n_\theta} & \text{MDL}_s \end{cases} \tag{5-8}$$

The rationale for deriving (5-7) and (5-8) is given in the next section.

5.3.2 Rationale

The new MDL_c rule (5-7) is derived from the corrected AIC rule (5-5), by analogy of AIC and MDL in (5-1). One could do exactly the same with the AIC_u rule (5-6) giving an MDL_u rule that has the tendency to select too simple models. Therefore, the choice (5-7) has been retained.

The starting point for deriving (5-8) are the AIC and MDL rules assuming (i) that the variance(s) of the disturbing noise source(s) is (are) known, and (ii) that no model errors are present

$$\begin{aligned} &\text{AIC:} && \frac{1}{N}(K(\hat{\theta}(Z), Z) + n_\theta) \\ &\text{MDL:} && \frac{1}{N}\left(K(\hat{\theta}(Z), Z) + \frac{n_\theta}{2}\ln(N)\right) \end{aligned} \tag{5-9}$$

(see [Ljung, 1999; Pintelon and Schoukens, 2001; Schoukens *et al.*, 2002]). In the absence of model errors the expected value of the global minimum of the maximum likelihood cost function equals

$$\mathrm{E}\{K(\hat{\theta}(Z), Z)\} = (N - n_\theta)/2 \tag{5-10}$$

(see [Pintelon and Schoukens, 2001], Theorem 17.12). Hence, on average, the value between brackets of AIC (5-9) will be

$$\begin{aligned} K(\hat{\theta}(Z), Z) + n_\theta &\approx \frac{N - n_\theta}{2} + n_\theta \\ &\approx \frac{N - n_\theta}{2}\left(1 + \frac{2n_\theta}{N - n_\theta}\right) \\ &\approx K(\hat{\theta}(Z), Z)\left(1 + \frac{2n_\theta}{N - n_\theta}\right) \end{aligned} \tag{5-11}$$

Making similar calculations for MDL we get

$$\frac{K(\hat{\theta}(Z), Z)}{N}(1 + p_s(n_\theta, N)) \tag{5-12}$$

where $p_s(n_\theta, N)$ is defined in (5-8).

Next recall that the factor $K(\hat{\theta}(Z), Z)/N$ in the classical (5-1) and corrected (5-5) AIC criteria is an estimate of the variance σ^2 of the innovations (= driving white noise source of the disturbance)

$$\hat{\sigma}^2 = \frac{2}{N}K(\hat{\theta}(Z), Z) = \frac{1}{N}e^T(\hat{\theta}(Z), Z)e(\hat{\theta}(Z), Z) \tag{5-13}$$

(see [Ljung, 1999; Sugiura, 1978; Hurvich and Tsai, 1989]). In the absence of modeling errors n_θ linear dependencies exist among the residuals $e(\hat{\theta}(Z), Z)$ (see [Pintelon and Schoukens, 2001], Chapter 17) and, therefore, an improved variance estimate is obtained as

$$\hat{\sigma}^2 = \frac{2}{N - n_\theta} K(\hat{\theta}(Z), Z) \tag{5-14}$$

Replacing $K(\hat{\theta}(Z), Z)/N$ by $K(\hat{\theta}(Z), Z)/(N - n_\theta)$ in (5-12) finally gives (5-8).

Although (5-8) has been derived in the absence of model errors (overmodeling), it is also used in the presence of model errors (undermodeling). In that case $2K(\hat{\theta}(Z), Z))/(N - n_\theta)$ is the sum of the innovations variance σ^2 and the remaining modeling errors. Note that a similar reasoning has been applied in [Ljung, 1999] for deriving AIC (5-1).

5.4 Simulation examples

Three simulation examples are given: a time domain signal modeling problem, a frequency domain system identification problem and noise modeling problem.

5.4.1 Signal modeling example

The goal in this example is to estimate simultaneously the harmonic content (fundamental frequency and the complex amplitudes) and the time base distortion parameters of the following signal model $s_m(n, \theta)$

$$s_m(n, \theta) = A_0 + \sum_{k=1}^{h} A_k \cos(k\omega t_n T_s) + A_{k+h} \sin(k\omega t_n T_s) \tag{5-15}$$

for $n = 0, 1, \ldots, N-1$, where A is a vector of length $2h$ with the unknown amplitudes and ω is the unknown fundamental angular frequency and T_s is the sample period,. The time instances, t_n, are given by

$$t_n = n + \sum_{l=1}^{b/2} B_l \cos(2\pi n l T_s) + B_{l+b/2} \sin(2\pi n l T_s) \tag{5-16}$$

where B is a vector of length b with the unknown time base distortion parameters. Note that the number of time base distortion parameters, b, and the number of harmonics, h, are unknown. The signal parameters $\theta = [\omega, A^T, B^T]^T$ are estimated by minimizing

$$K(\theta, Z) = \frac{1}{2} \sum_{n=0}^{N-1} (s(n) - s_m(n, \theta))^2 \tag{5-17}$$

w.r.t. θ, where $s(n)$ is the observed signal (see chapters 3 and 4 for the details).

The chosen parameter values in the Monte-Carlo simulation are $\omega = 2\pi 5$ year^{-1}, $A = [1, 0, 0, 1]^T$ $(h = 2)$, $B = [1, 1, 0, 0, 0, 1]^T$ $(b = 6)$, and $N = 50$. Zero mean white Gaussian noise was added such that the signal-to-noise ratio of the measurements equals 2 or 1.5. Note that the variance of the disturbing noise is unknown, and hence has to be estimated. The optimal model order (h, b) is selected by minimizing AIC and MDL (5-

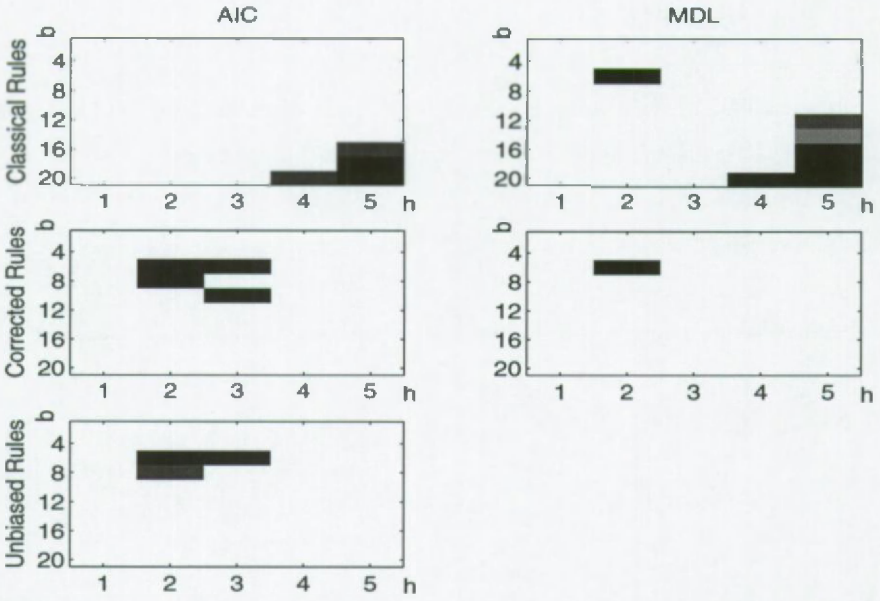

Figure 1: Graphical visualization of the selected model complexity (h, b) by the AIC (left) and MDL (right) criteria in the simulation ($SNR = 2$). The first line gives the result of the old rules (5-1), the second line of AIC_c (5-5) and the new MDL_c (5-7), and the bottom line of AIC_u (5-6). Notice the contraction of the results around the true model ($h = 2$ and $b = 6$).

1), AIC_c (5-5), AIC_u (5-6) and MDL_c (5-7) over the set $h = 1, \ldots, 5$ and $b = 2, 4, \ldots, 20$ ($n_\theta = b + 2h + 1$).

This simulation was ran one hundred times to show its average performance. Figure 1 and Table 1 show the number of times that AIC and MDL (5-1), AIC_c (5-5), AIC_u (5-6) and MDL_c (5-7) selected a certain model order. The following observations can be made:

1. The undermodeling ($h = 1$ and $b = 2, 4$) is clearly detected by all model selection criteria, which is not the case for the overmodeling ($h > 2$ and/or $b > 6$).
2. The classical AIC and MDL rules (5-1) often overestimate the model order: the average selected model orders are too high, especially regarding the time base distortion.
3. When the SNR is high (≥ 2 in this simulation) the new MDL_c rule outperforms AIC_u, while for low SNR values AIC_u performs better.

4. As could be expected for the classical rules (5-1), AIC selects higher orders than MDL. The same is valid for AIC_c (5-5) w.r.t. MDL_c (5-7).

Table 1: Results signal modeling simulation example for two signal-to-noise ratios (SNR).

SNR = 2	**classical rules (5-1)**		**AIC_c (5-5)**	**AIC_u (5-6)**	**new rule MDL_c (5-7)**
	AIC	**MDL**			
% selection true model $(h, b) = (2, 6)$	0%	10%	69%	86%	98%
mean(h, b) ± std(h, b)	(5.0, 19.4) ± (0.2, 1.4)	(4.4, 16.6) ± (1.0, 5.1)	(2.3, 6.7) ± (0.8, 1.7)	(2.1, 6.3) ± (0.4, 1.0)	(2.0, 6.0) ± (0.0, 0.3)
SNR = 1.5	**classical rules (5-1)**		**AIC_c (5-5)**	**AIC_u (5-6)**	**new rule MDL_c (5-7)**
	AIC	**MDL**			
% selection true model $(h=2, b=6)$	0%	4%	59%	83%	63%
mean(h, b) ± std(h, b)	(4.9, 19.2) ± (0.4, 1.4)	(4.6, 17.6) ± (0.8, 3.5)	(2.5, 7.3) ± (0.9, 2.5)	(2.1, 6.2) ± (0.6, 1.2)	(2.0, 5.2) ± (0.3, 1.6)

5.4.2 System identification example

The frequency domain system identification example consists in the identification of a discrete-time, finite impulse response (FIR) system $G(z^{-1})$

$$G(z^{-1}, \theta) = \sum_{r=0}^{R} g_r z^{-r} \text{ with } \theta = [g_0, g_1, \ldots, g_R]^T \tag{5-18}$$

from noisy observations of the input/output discrete Fourier transform (DFT) spectra $U(k)$, $Y(k)$

$$\begin{aligned} Y(n) &= Y_0(n) + N_Y(n) \\ U(n) &= U_0(n) + N_U(n) \end{aligned} \tag{5-19}$$

$U_0(n)$, $Y_0(n)$ stand for the true unknown input/output DFT spectra and $N_U(n)$, $N_Y(n)$ for the input/output errors. For a given order R, the model parameters are estimated by minimizing the maximum likelihood cost function [8]

$$K(\theta, Z) = \sum_{n=0}^{N-1} \frac{\left|Y(n) - G(z_n^{-1}, \theta)U(n)\right|^2}{\sigma_Y^2(n) + \sigma_U^2(n)\left|G(z_n^{-1}, \theta)\right|^2 - 2\mathrm{Re}(\sigma_{YU}^2(n)\bar{G}(z_n^{-1}, \theta))} \tag{5-20}$$

with σ_Y^2, σ_U^2 and σ_{YU}^2 the known noise (co-)variances of the disturbing noise sources N_Y and N_U; $z_k = \exp(j2\pi n/N)$; N the number of frequencies in the band $[0, 1)$ (normalized w.r.t. f_s); Re( ) the real part of; and where the overline stands for the complex conjugate. Since $z_{N-n} = \bar{z}_n$, $Y(N-n) = Y(n)$ and similarly for $U(n)$ and $\sigma_{YU}^2(n)$, (5-20) contains only N independent real residuals: $N/2-1$ complex ($n = 1, 2, \ldots, N/2-1$) and two real ($n = 0, N/2$). Hence, the AIC and MDL model selection rules (5-2) should be applied with N and $n_\theta = R+1$. Because of the errors-in-variables framework (all observations are noisy), the MDL rules should be used with $\ln(2N)$ instead of $\ln(N)$ (see footnote 1 and [8]).

The true FIR systems $G_0(z^{-1})$ consists of the twenty first samples ($r = 19$ in (5-18)) of the impulse response of a sixth order digital Chebyshev filter with a pass band ripple of 0.1 dB, and a cutoff frequency of 0.1 (normalized w.r.t. to the sampling frequency f_s). One thousand runs of a Monte-Carlo simulation are performed with $N = 80$. For each run independent (over the frequency k), circular complex normally distributed errors (see [8]) with standard deviations $\sigma_Y = \sigma_U = 0.05$ and $\sigma_{YU} = 0$ are added to the true input/output spectra $U_0(k) = 1$ and $Y_0(n) = G_0(z_n^{-1})U_0(n)$; and the model parameters θ are estimated for model orders $R = 14, 15, \ldots, 24$. The optimal model order R is selected by minimizing the classical (5-2), and the new (5-8) AIC and MDL rules over the set $R = 14, 15, \ldots, 24$

($n_\theta = R + 1$).Table 2 shows the number of times a model of order R has been selected for

Table 2: Number of times an FIR model (5-18) of order n has been selected (true model order = 19), and mean value and standard deviation of the selected model order

model order R	classical rules (5-2)		new rules (5-8)	
	AIC	MDL	AIC_s	MDL_s
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	*356*	*607*	*845*	*923*
20	125	136	78	49
21	103	85	45	18
22	112	55	16	4
23	130	55	10	4
24	174	62	6	2
mean value	21.1	20.0	19.3	19.1
standard deviation	1.9	1.5	0.79	0.51

the different model selection criteria. The following observations can be made:

1. The undermodeling (R = 14, 15, ..., 18) is clearly detected by all model selection criteria, which is not the case for the overmodeling (R = 20, 21, ..., 24).
2. The classical AIC and MDL rules (5-2) clearly overestimate the model order.
3. The new AIC_s and MDL_s rules (5-8) outperform: the number of times the correct model order is selected is higher, the mean value of the selected orders is closer to the true value, and the standard deviation of the selected orders is smaller.
4. As could be expected for the classical rules (5-2), AIC selects higher orders than MDL. The same is valid for the new rules (5-8).

It can be concluded that the new MDL_s rule (5-8) outperforms on this example.

5.4.3 Noise modeling example

Consider an autoregressive (AR) noise model of order R

$$y(t) + a_1 y(t-1) + a_2 y(t-2) + \ldots + a_R y(t-R) = e(t) \tag{5-21}$$

where $e(t)$ is a zero mean white noise source with variance σ^2. The parameters of the AR-model a_1, a_2, ..., a_R are estimated form N observations $y(0), y(1), \ldots, y(N-1)$ (no observations of $e(t)$ are available). Since the data record is short, it is necessary to estimate also the initial conditions $y(-1)$, $y(-2)$, ..., $y(-n)$ of the AR-process, and the vector of the noise model parameters equals

$$\theta = [a_1, a_2, \ldots, a_R, y(-1), y(-2), \ldots, y(-R)]^T \tag{5-22}$$

This can be done either in the time domain [6] or in the frequency domain [8] leading to exactly the same results [8]. The estimate $\hat{\theta}$ is found by minimizing

$$K(\theta, Z) = \frac{1}{2} \sum_{t=0}^{N-1} \left| (1 + \textstyle\sum_{r=1}^{R} a_r q^{-r}) y(t) \right|^2 \tag{5-23}$$

w.r.t. θ, where q^{-1} is the backward shift operator. For the AR-modeling example (5-21), the performance of the small sample AIC and MDL rules (5-5) to (5-7) can also be compared to the combined information criterion (CIC)

$$\frac{K(\hat{\theta}(Z), Z)}{N-R} e^{p_{cic}(R, N)} \tag{5-24}$$

where

$$p_{cic}(R, N) = \max\left(\prod_{i=0}^{R} \frac{1 + v(i)}{1 - v(i)} - 1, 3 \sum_{i=0}^{R} v(i) \right) \text{ and } v(i) = 1/(N + 2 - 2i) \tag{5-25}$$

for the least squares estimator (5-23) (see [2]).

The true AR-process has the following parameters: $R = 4$, $a_1 = -1.9602$, $a_2 = 2.2454$, $a_3 = -1.6450$, $a_4 = 0.5776$ (poles at $0.8e^{\pm 0.12\pi j}$ and $0.95e^{\pm 0.42\pi j}$), and $\sigma = 1$. One thousand runs of a Monte-Carlo simulation are performed with $N = 50$. For each run the model parameters θ and the noise variance σ^2 are estimated for model orders $R = 1, 2, \ldots, 10$. The optimal model order n is selected by minimizing AIC and MDL (5-1), AIC_c (5-5), AIC_u (5-6), MDL_c (5-7), and CIC (5-24) over the set $R = 1, 2, \ldots, 10$ ($n_\theta = 2R$).

Table 3 shows the number of times a model of order R has been selected for the different model selection criteria. The following observations can be made:

Table 3: Number of times an AR model (5-21) of order n has been selected (true model order = 4), and mean value and standard deviation of the selected model order.

model order R	classical rules (5-1)		AIC_c (5-5)	AIC_u (5-6)	CIC (5-24)	new rule MDL_c (5-7)
	AIC	MDL				
1	0	0	0	0	0	0
2	0	1	0	1	0	3
3	11	72	40	74	20	167
4	*645*	*881*	*880*	*896*	*763*	*824*
5	106	38	67	25	116	5
6	72	7	10	4	49	1
7	48	1	3	0	21	0
8	49	0	0	0	19	0
9	29	0	0	0	5	0
10	40	0	0	0	7	0
mean value	5.0	4.0	4.1	4.0	4.4	3.8
standard deviation	1.7	0.4	0.4	0.3	1.0	0.4

1. Contrary to the signal modeling (see Section 5.4.5.4.1) and the system identification example (see Section 5.4.5.4.2), both too simple (undermodeling) and too complex (overmodeling) models are selected.
2. The classical AIC (5-1) and the CIC (5-24) rules clearly overestimate the model order.
3. The AIC_u rule (5-6) performs better than all the other rules: the number of times the correct model order is selected is higher, the mean value of the selected orders is closer to the true value, and the standard deviation of the selected orders is smaller. For $N \geq 90$ the MDL_c rule (5-7) outperforms. Below this N-threshold the new MDL_c rule selects too often too simple models and, hence, the AIC_u rule should be used.

Notice

1. Although CIC less often selects the true model than AIC_c and AIC_u, the mean square error of the power spectra of the selected models are about the same. It shows that the variance of CIC (too complex models) is traded for the bias of AIC_c and AIC_u rules (too simple models).
2. When decreasing the number of samples N in the simulation ($N < 50$), first the AIC_c performs best ($N = 40$), and next the CIC rule ($N = 26$).

5.4.4 Conclusion

In general for physical modeling the new MDL_c (5-7) (unknown noise variance) and MDL_s (5-8) (known or prior estimated noise variances) rules outperform on small data records. Below some signal-to-noise ratio (SNR) threshold the new MDL_c (5-7) and MDL_s (5-8) rules have the tendency to select too often too simple models and, hence, should be replaced by respectively the existing AIC_u /AIC_c rules and the new AIC_s (5-8) rule. This SNR threshold is problem dependent and should be established by simulation. On real data with unknown noise variance, the SNR of the measurements can be calculated a posteriori using (5-14) as variance estimate. An alternative way for choosing between the different AIC and MDL criteria consists in applying the metaselection concept of [De Luna and Skouras, 2003]. A disadvantage of the method is that S times more models must be estimated, where S is the number of data subsets.

5.5 Real measurement example

The actual example is a data record[1] sampled from a the skeleton of a bivalve, i.e. a Saxidomus giganteus (the number of samples is $N = 63$). Herein the ratio of two isotopes of oxygen was measured. It lived at the coast of Washington State (U.S.A.). Collecting such a data record of the stable oxygen isotopes in a shell takes about 1.5 days work of manual sampling and another 1.5 days for automatically analyzing the samples. This signal is a proxy for the sea surface temperature and the fluctuations have an annual periodicity, which is disturbed by growth rate variations. The annual periodicity can be modeled by equation (5-15) and the variations in growth rate by equation (5-16). Again, we will try to estimate the number of harmonics, h, and the number of time base distortion parameters, b, necessary to model the signal properly. The most complex model consists of five harmonics and twenty time base distortion parameters. Figure 2 illustrates the results found by the different criteria, i.e. the classical, corrected and new rules of the AIC and MDL criterion for problems with unknown noise variance. The classical AIC selected the most complex model within the model set ($h = 5, b = 20$), while the classical MDL criterion was able to select already a reasonable model complexity ($h = 1, b = 8$). The AIC_c and AIC_s rules both selected a model with $h = 1, b = 8$. The new MDL_c criterion selected a model with $h = 1, b = 6$. Both models corresponded to a SNR value of three. A Monte-Carlo simulation with this SNR value and a model with $h = 1, b = 6$ was ran 50 times. This simulation showed that the model selection was performed above the MDL threshold and that the new MDL_c rule (5-7) is to be preferred. To conclude

1. A thick-section of the aragonite shell was continuously sampled using a computer-controlled micro drill from growing tip to half way to the umbo. The carbonate powder (±50 μg) was processed using an automated carbonate device (Kiel III) coupled to a Finnigan Delta+XL. Data were corrected using an internal laboratory standard and are reported relative to V-PDB in conventional notation. Precision with this instrumental set up is generally better than 0.08‰.

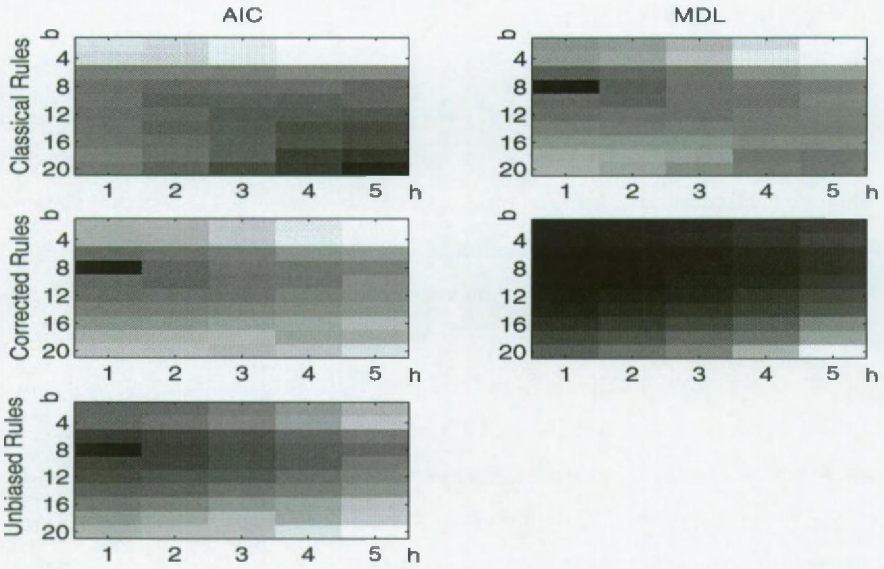

Figure 2: Graphical visualization of the AIC (left) and MDL (right) values as a function of model complexity (h, b) for the real world example: black are the lowest values and white are the highest. The first line gives the result of the old rules (5-1), the second line of AIC_c (5-5) and the new MDL_c (5-7), and the bottom line of AIC_u (5-6).

1. While the old MDL rule (5-1) selected a reasonable model, this is not the case for the old AIC rule (5-1);
2. The AIC_c (5-5) and AIC_u (5-6) rules select the same model complexity. This is not in contradiction with the simulation results of Table 1 where it can be seen that both rules select the same model with high probability;
3. As can be seen in Figure 2, the new MDL_c rule (5-7) selected a slightly less complex model than AIC_c and AIC_u. Thus independent of the rules the AIC selects more complex models than the MDL criterion.

5.6 Conclusion

Based on an intuitive reasoning, modified AIC and MDL model selection rules for short data records have been derived. Such criteria were not available for estimation problems with known or prior estimated noise model. Based on the simulation results and the real measurement data the following conclusion can be drawn. No single model selection criterion outperforms under all circumstances. For sufficiently high signal-to-noise ratio's (SNR) or sample sizes the new MDL_c (5-7) and MDL_s (5-8) rules outperform. Below some SNR (or sample size) threshold they should be replaced by respectively the existing AIC_c (5-5) or AIC_u (5-6) rules and the new AIC_s (5-8) rule. In AR-modeling the CIC (5-24) rule is also a potential candidate. The choice between the different finite sample criteria can be based on the SNR or sample size threshold (to be established by simulation) or by the meta model selection approach in [3]. The classical rules (5-1) and (5-2) are a special case of the new rules (5-7) and (5-8). With an increasing number of samples the new rules (5-7) and (5-8) converges asymptotically to the old rules (5-1) and (5-2). So, the new rules can be used for short as well as for long data records.

References

[1] Akaike H., 1974. A new look at the statistical model identification, IEEE Trans. on Automatic Control, vol. AC-19, pp. 716-723.

[2] Broersen P.M.T., 2000. Finite sample criteria for autoregressive order selection, IEEE Trans. Sign. Proc., vol. 48, no. 12, pp. 3550-3558, 2000.

[3] De Luna X. and Skouras K., 2003. Choosing a model selection strategy, Scand. J. Statist., vol. 30, pp. 113-128, 2003.

[4] Hannan E.J., 1980. The estimation of the order of an ARMA process, The Annals of Statistics, vol. 8, no 5, pp. 1071-1081, 1980.

[5] Hurvich C.M. and Tsai C.L., 1989. Regression and time series model selection in small samples, Biometrica, vol. 76, pp. 297-307, 1989.

[6] Ljung L., 1999. System Identification: Theory for the User (second edition), Prentice Hall, Upper Saddle River.

[7] McQuarrie A., Shumway R. and Tsai C.L., 1997. The model selection criterion AICu, Statistics & Probability Letters, vol. 34, no. 3, pp. 285-292, 1997.

[8] Pintelon R., and Schoukens J., 1999. Time Series Analysis in the Frequency Domain, IEEE Trans. Sign. Proc., vol. 47, no. 1, pp. 206-210.

[9] Pintelon R. and Schoukens J., 2001. System Identification: A Frequency Domain Approach, IEEE Press, New York (USA).

[10]Rissanen J., 1978. Modeling by shortest data description, Automatica, vol. 14, pp. 465-471, 1978.

[11]Schoukens J., Rolain Y. and Pintelon R., 2002. Modified AIC rule for model selection in combination with prior estimated noise models, Automatica, vol. 38, no. 5, pp. 903-906, 2002.

[12]Schwarz G., 1978. Estimating the dimension of a model, The annals of statistics, vol. 6, no. 2, 461-464, 1978.

[13]Sugiura N., 1978. Further analysis of the data by Akaike's information criterion and the finite corrections, Comm. Statist., vol. A7, pp. 13-26, 1978.

CHAPTER 6

PERIODICITY AND ACCRETION RATE: NUMERICAL ISSUES

Abstract - When an environmental proxy is measured, its time base is often unknown. Here a methodology is presented which can estimate the time base for periodic proxy-signals. First the method is proposed, next several case studies are discussed. In order to estimate the time base a parametric model is proposed. Special attention is given to the optimization strategy of this model and to the model complexity, measured by the number of parameters, which is determined employing an adapted automated model selection criterion.

6.1 Context

Immediately after the initial estimate of the time base distortion, we started working on the implementation of the parametric approach. From each problem that could be solved, another one was raised. Two of these problems are already discussed in the previous two chapters. In this chapter, several other smaller problems are handled, such as the optimization strategy used to avoid local minima, the reduction of the number of parameters and the local time inversion problem.

In this chapter the structure of the algorithm is discussed: how and where are the three previous ideas implemented? How are local minima avoided? This is necessary if one wants to implement a model selection based on the residual cost function. Sometimes too many parameters are optimized. How can redundant parameters be detected and removed? How can time inversion problems be avoided? Finally, some previously discussed examples are re-evaluated.

6.2 Introduction

In this section, first the signal and time base models are recapitulated and next solutions are given for some problems which were encountered, while implementing the algorithm, like

1. gathering initial values for the parameters;
2. Sometimes it can occur that the more complex models get stuck in local minima. Additional measures are used to minimize this risk;
3. Sometimes it can occur that too many parameters are used: the redundant parameters are identified and removed;
4. Due to the parametric representation of the time base, neighboring observations can be altered, which would mean that the time is locally inverted. This artefact in circumvented.

In order to understand the sequence of steps it is important to keep in mind that we will finally optimize all models within a given model set and compare the corresponding model selection criterion. Therefore, we start with a very simple model, optimize its parameters and repeat this with a slightly more complex model. This is repeated until a given maximum tolerable model complexity is reached.

The approach is illustrated on several proxy records, like clams, a coral and a mangrove tree.

6.2.1 The signal and time base model

In this and the next paragraph the equations used are briefly recapitulated. The signal under investigation is assumed to be periodic, sampled along a equidistant distance grid[1]. It is assumed that the discrete-time signal, $s(t_n)$, is given by

$$s(t_n) = A_0 + \sum_{k=1}^{h} A_k \cos(k\omega t_n) + A_{k+h} \sin(k\omega t_n) \qquad (6\text{-}1)$$

1. If this assumption is not full-filled, the data should first be interpolated.

where t_n is the unknown time variable at sample position $n \in \{1, ..., N\}$, A_0 the offset, A_k and A_{k+h} are the unknown amplitudes of the kth harmonic, ω is the unknown fundamental angular frequency and h the number of harmonics, yet to be identified. The time instance between two subsequent samples is not constant, because of the variations in accretion rates. These distortions of the time base are modelled by a function[1], $g(n)$, called the time base distortion (T.B.D.)

$$g(n) = \sum_{l=1}^{b/2} B_l \cos(2\pi n l T_s) + B_{l+b/2} \sin(2\pi n l T_s) \tag{6-2}$$

where B is a vector of length b, yet to be identified, with the unknown time base distortion parameters. The time instances, t_n, are given by

$$t_n = (n + g(n))T_s \tag{6-3}$$

where T_s is the sample period.

6.2.2 Cost function and Gibbs phenomenon

The parameters to be estimated within one model are grouped in a vector

$$\theta = [\omega, A, B]^T \tag{6-4}$$

(T stands for the transpose). The optimal set of parameters is given by

$$\hat{\theta} = \arg\min_{\theta} K(\theta) \tag{6-5}$$

where the explicit expression for the least squares cost function is given by

$$K(\theta) = \frac{1}{2} \sum_{n=1}^{N} [s(n) - \sum_{k=1}^{h} A_k \cos(k\tilde{\omega} t_n T_s) + A_{k+h} \sin(k\tilde{\omega} t_n T_s)]^2 \tag{6-6}$$

1. A polynomial model is tested too, but it seems to perform worse on the examples tested so far. Herein orthogonal polynomials were used, which have more troubles with describing discontinuity than a Fourier based model.

Here, the model window is enlarged, compared to the measurement window, so that the Gibbs phenomenon occurs mainly outside the original measurement window. These virtual borders are ignored in the final estimation of the model parameters, but allows the procedure to reduce the Gibbs phenomenon inside the measurement window. Briefly, this can be done by replacing ω by $\tilde{\omega}$, with $\tilde{\omega} = \omega(N+q)/N$.

6.2.3 Starting value problem and optimization strategy

Minimizing cost function (6-6) will only be successful if one can start from a reasonable set of initial values. Otherwise the local optimization method will possibly not be able to converge toward a good minimum. In order to minimize the risk of converging towards a local minimum, the optimization strategy is performed in three steps (for a structure of the algorithm, see Figure 1)

1. *Initializing the angular frequency:* as is shown in chapter 3 a linear trend occurs in the time base distortion if the initial frequency is badly chosen. The reason for this is that a structural degeneracy occurs in the model: it is not possible to estimate the fundamental frequency and a linear trend with this model, because a perturbation in one causes a proportional change in the other and the new model parameters match equally well the measurements. But luckily this degeneracy can be exploited to estimate much more accurately the angular frequency, if we forbid linear trends in the time base distortion. The non-parametric estimation of the time base distortion is used, as described in chapter 3. Suppose that the true radial frequency falls inside the spectral window, but does not equal the initial estimated frequency, a trend will appear in the time base distortion. The slope of this trend is proportional to the mismatch between the initially estimated frequency and the true one. An iterative procedure is used to tune the frequency, until no trend occurs in the time base distortion.
2. *Non-parametric estimation of the T.B.D.:* a non-parametric estimation of the T.B.D. is made. Therefore, the anchor point or phase demodulation method can be used (see chapter 3).
3. *Initialization of the T.B.D. parameters and signal parameters:* initial values for the T.B.D. parameters can be gathered by matching equation (6-2) on the non-parametric T.B.D. representation, estimated in chapter 3, which can easily be done, because equation (6-2) is linear in the parameters, B. Equation (6-3) is used to get more precise sample positions.

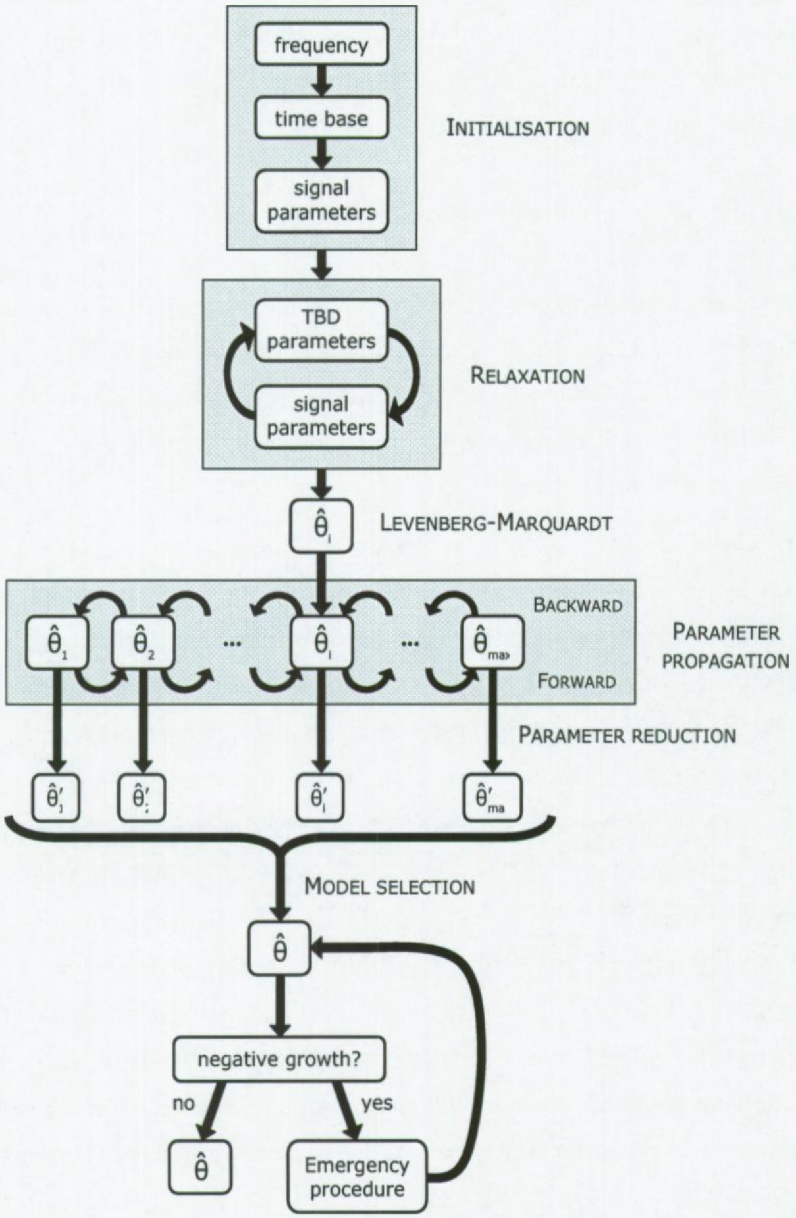

Figure 1: Structure of the algorithm.

4. *Initialization of the signal parameters, A and* ω : these are matched employing the time base estimated in equation (3) and the algorithm described in [Pintelon and Schoukens, 1996].

5. *Relaxation:* alternating, the T.B.D. parameters and the signal parameters are optimized, while the other set is remained fixed. Note that optimizing the T.B.D. parameters, while the signal parameters are constant is, in fact, the parametric orbital tuning method proposed by Martinson *et al.* [1982]. This relaxation algorithm is stopped when the largest relative variation in the parameter vector, θ, is lower than a certain numerical stop criterion (typically 10^{-3}). This step in the optimization is implemented to increase the calculation speed but it will not influence the final results.

6. *Final estimation of parameters:* all parameter values are estimated together employing a Levenberg-Marquardt algorithm (this algorithm is preferred over the Newton-Gauss algorithm, because the convergence area is larger). Further, each row in the Jacobian was scaled by its standard deviation in order to improve the numerical conditioning.

6.2.4 Backward/Forward parameter propagation

If more complex models are optimized, the risk is still present that the optimization get stuck in a local minimum. This is illustrated in Figure 2, where the residual cost functions are shown as function of the number of harmonics and the number of time base distortion parameters for clam 2 (see paragraph 6.3.1). This can easily be understood. Take for instance a model with complexity (h, b) and a more complex model with complexity (h', b'). If now, the additional parameters are set equal to zero, the residual cost function of the more complex model will equal that one of the simpler model. But, if these additional parameters are optimized with the Levenberg-Marquardt algorithm, the cost function will remain the same or will decrease. Consequently, for an increasing number of parameters, the residual cost function should monotonically decrease and, therefore, the peaks shown in Figure 2 must be due to local minima. To avoid these minima a so-called forward- backward propagations is used, where the parameters of a neighboring model complexity are tested to optimize a given model complexity.

Forward parameter propagation: to avoid these local minima, all models, except the most simple one, are optimized starting from the estimated parameters of the neighboring simpler

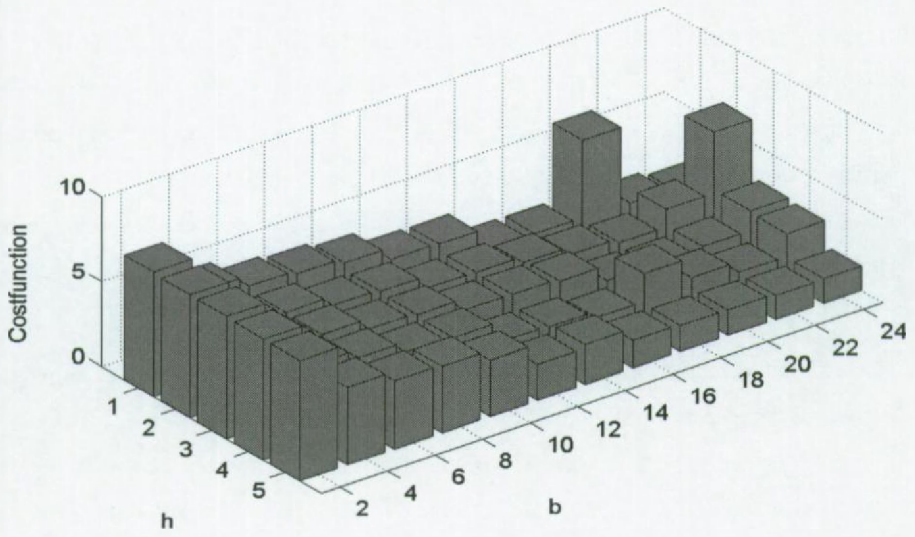

Figure 2: Residual cost function before the backward/ forward algorithm, as function of the the number of harmonics, h, and the number of time base distortion parameters, b.

model, i.e. a model with complexity (h, b) is once initialized with the parameters estimated with the model with $2(h-1)$ signal and b T.B.D. parameters and once with h signal and $b-2$ T.B.D. parameters. The new parameter values are set zero before the optimization starts. After the optimization, the residual cost functions are compared with the original one and those parameters, corresponding to the lowest residual cost function are retained.

After the optimization of all models in the forward step, the *backward parameter propagation* is performed. Analog to the previous step, it seems possible that a lower residual cost function is reached if the initial values are copied from a more complex model. A model with complexity (h, b) is once initialized with the parameters copied from the optimization of the model with $2(h+1)$ signal and b T.B.D. parameters and once with h signal and $b+2$ T.B.D. parameters. The redundant parameters are dropped before the optimization. Again that set of estimated parameters is chosen which corresponds to the lowest residual cost function.

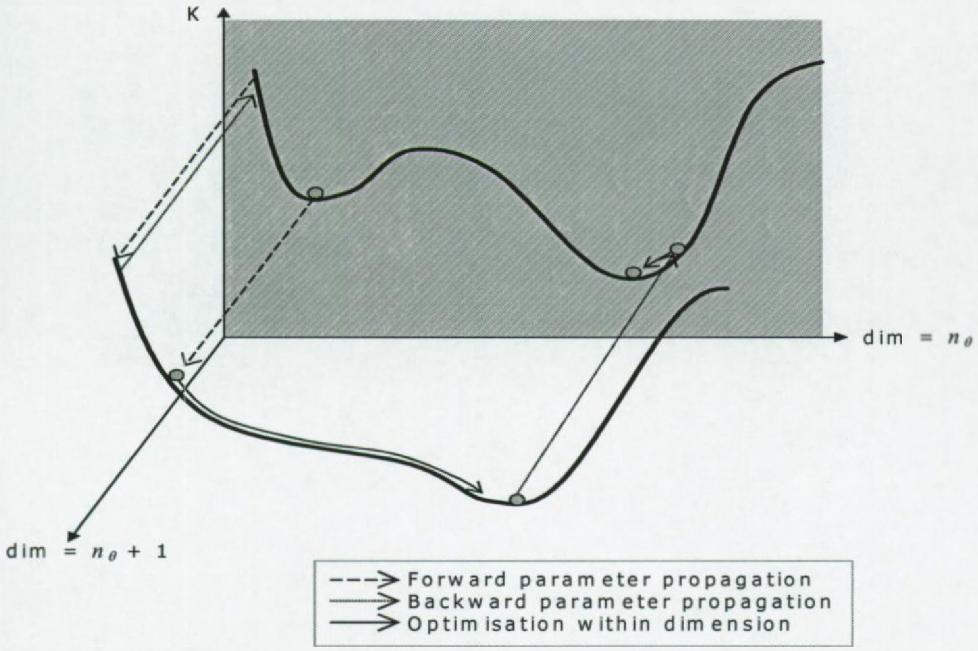

Figure 3: Illustration of how a local minimum is avoided by the backward/forward parameter propagation algorithm.

Alternating the forward and backward step until all optimizations have converged guarantees at least that all models are in the same local minimum and can thus be compared. Figure 3 illustrates how a local minimum can be avoided by iteratively using the backward/forward parameter propagation. This loop is stopped when the relative cost function decrease is less than a numerical stop criterion, e.g. 10^{-8}. The maximum number of times this forward-backward propagation loop is processed is ten, but convergence was always reached in less than five iterations. Figure 4 shows the residual cost functions after this procedure: every model has a lower residual cost function than the more simple models and some residual cost functions have still decreased, although their value was already below that of the more simple models (e.g. $(h, b) = (1, 8)$, where the cost function decreased from 4.4 to 3.4). In order to be sure of the global minimum, global optimization methods should be implemented [Jaulin *et al.*, 2001; De Brauwere, 2003], but the disadvantage of these methods, at the moment, are that the calculation time becomes too large.

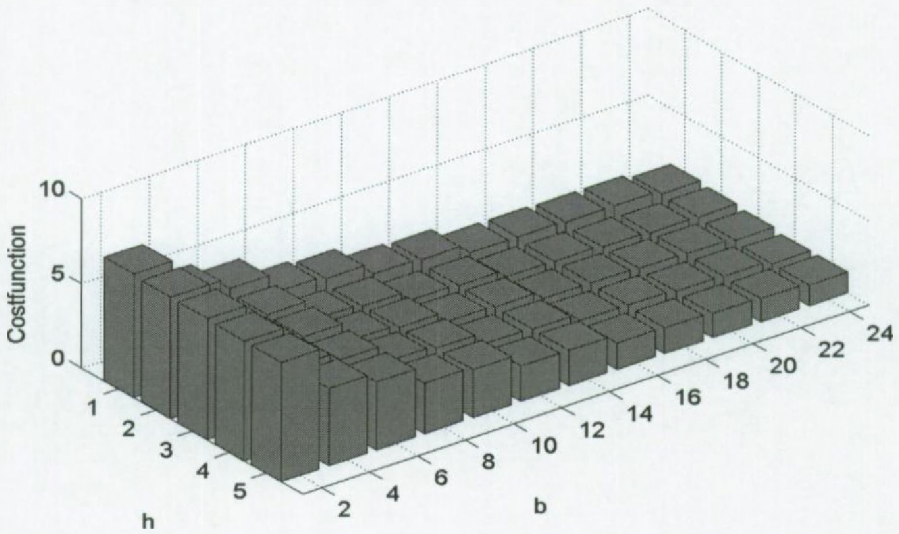

Figure 4: Residual cost function after the backward/ forward algorithm, as function of the number of harmonics, h, and the number of time base distortion parameters, b.

6.2.5 Parameter reduction

Before we can compare different models, the impact of the individual parameters within one single model has to be evaluated. Some parameters, which do not impart to the description of the data must be removed. Suppose, for instance, a signal which is described by two harmonics, i.e. the first and the third, while the second is absent. So far, six parameters would be needed to match the model to this signal, instead of only four (two for the first harmonic and two for the third). In order to avoid these too complex models, those parameters, which do not contribute to the description of the data are removed as follows:

1. The model selection criterion of the unaltered model is calculated (see 6.2.6).
2. The amplitude of all harmonics, used to match the signal, are calculated. That set of parameters, corresponding to the lowest amplitude, is set equal to zero. The remaining signal and T.B.D. parameters are optimized again as described in step 6 of 6.2.3.

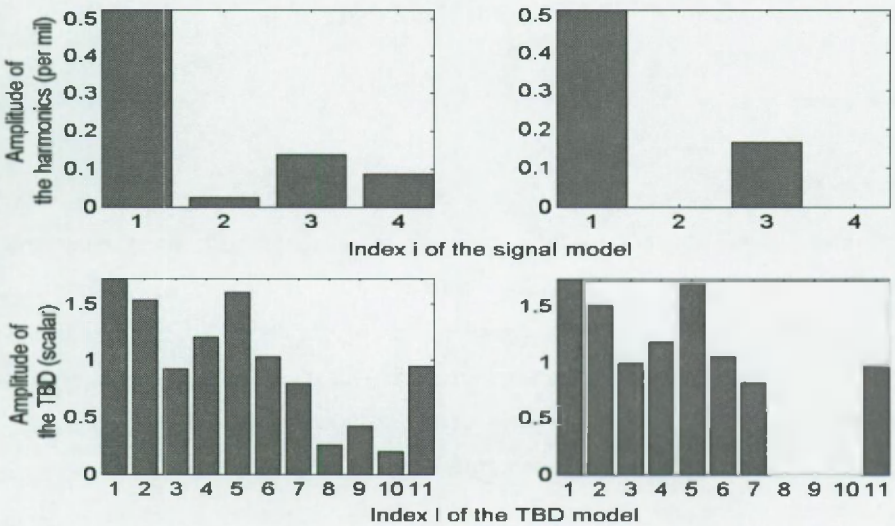

Figure 5: Parameters estimated on the $\delta^{18}O$-signal from a *Saxidomus gigantues* (Clam 1): at the left before and at the right after parameter removement step. The first row shows the amplitude of the harmonics of the signal model (equation 6-1) and the second row shows the amplitude of the time base (equation 6-6).

3. The same procedure is performed on the T.B.D. parameters.
4. The model selection criteria (see 6.2.6) are calculated for both less complex models and compared with that of the unaltered model. The model corresponding to the lowest model selection criterion is retained
5. The latter model replaces the 'unaltered model' in step 1; next steps 1 to 4 are repeated until the selection criterion calculated in step 1 is the lowest.

To illustrate that this step in the procedure is not exotic, the amplitude of the T.B.D. parameters estimated on the $\delta^{18}O$-signal, measured in a *Saxidomus giganteus* from Washington state are shown in Figure 5 before and after the parameter removement step (Clam 1, which is discussed later (see 6.3.1)). Note, finally, that an additional benefit of this procedure is that the algorithm is now able to identify if no periodicity or no T.B.D. is significantly present in the data-record, e.g. if all signal or T.B.D. parameters are insignificant compared to the stochastic noise and, thus, one should be removed. So, the

model selection criterion enables, at least for certain signal records[1], to validate the basic assumption on which the model is build, i.e. is the signal periodic?

6.2.6 Model selection criterion

So far, we have a set of optimized models, where all redundant parameters have been removed. In this paragraph we will show how to select the best model within this set. Increasing the model complexity will decrease the systematic errors, however, at the same time the model variability increases. Hence, it is not a good idea to select the model with the smallest cost function within the set because it will continue to decrease when more parameters are added. At a certain complexity the additional parameters no longer reduce the systematic errors but are used to follow the actual noise realization on the data. As the noise varies from measurement to measurement, the additional parameters increase only the model variability. To avoid this unwanted behavior the cost function is extended with a model complexity term that compensates for the increasing model variability. Summarized, the model selection criterion should be able to detect undermodeling (= too simple model) as well as overmodeling (= too complex model). This model complexity term is dependent upon the signal-to-noise ratio and the availability of a noise model. In the examples worked out in this paper the criterion to be minimized had the following expression

$$\frac{K(\hat{\theta})}{N} e^{p_c(n_\theta, N)} \text{ with penalty } p_c(n_\theta, N) = \frac{\ln(N)(n_\theta + 1)}{N - n_\theta - 2} \tag{6-7}$$

with $K(\hat{\theta})$ the residual cost function, N the total number of observations and n_θ the number of parameters. Notice that introducing a model selection criterion eliminated all interferences from the user, which makes the proposed method objective and user independent. Practically, this means that only a model set has to be defined, i.e. the maximum values for h and b must be set. Next equation (6-7) can be used to select the best model within this set. A detailed description of the model selection criteria can be found in chapter 5.

1. 'nearly periodic signals', disturbed by e.g. time jitter [Schoukens *et al.*, 1997], will probably be identified as periodic and in these cases the validation fails.

6.2.7 Growth rate

Once the equidistant samples are dated, the accretion rate can be calculated with

$$\text{accretion rate}(n) = \frac{D_S / T_s}{1 + \left(\frac{dg(n)}{dn}\right)} \tag{6-8}$$

The denominator is an estimate of the time gap between observation $n-1$ and observation n, while the nominator, D_S, is the distance between both physical observations. Consequently, the ratio between both is the accretion rate.

6.2.8 Local time reversal problem

The last problem encountered is that for an optimized model, it can occur that the sequence of the observations is altered. This would occur if the derivative of the T.B.D., $g'(n)$, would become more negative than -1, in other words when the accretion rate becomes locally negative. Although the proposed solution is surely not optimal, it can manage the problem by removing those observations, where the accretion rate had become negative and by optimizing the model again on the remaining observations. This way we were able to estimate an accretion rate which is always positive. An example where this intervention was necessary is Clam 3 (see 6.3.1). The signal seems not periodic in the oldest part of the clam and method tries to match the model by employing a negative growth rate. Therefore, the first ten observations are dropped and the model is matched again on the remaining part.

In theory, this step can be repeated if the recalculated growth rate would be negative again, but, so far, this situation did not arise. This step was performed after the model selection criterion, although it should be performed before it, but the latter would at least double the calculation time. Note, finally, that an estimated negative accretion rate can be found if the 'true' accretion rate would be small and positive, with a large uncertainty. The most robust solution is to use a constraint optimization method guaranteeing that $1 + g'(n) > 0$, which can most easily be implemented in combination with a global optimization approach. However, these adaptations are not yet implemented.

Another alternative consists in the use of positive polynomials, which cannot become negative, due to their construction.

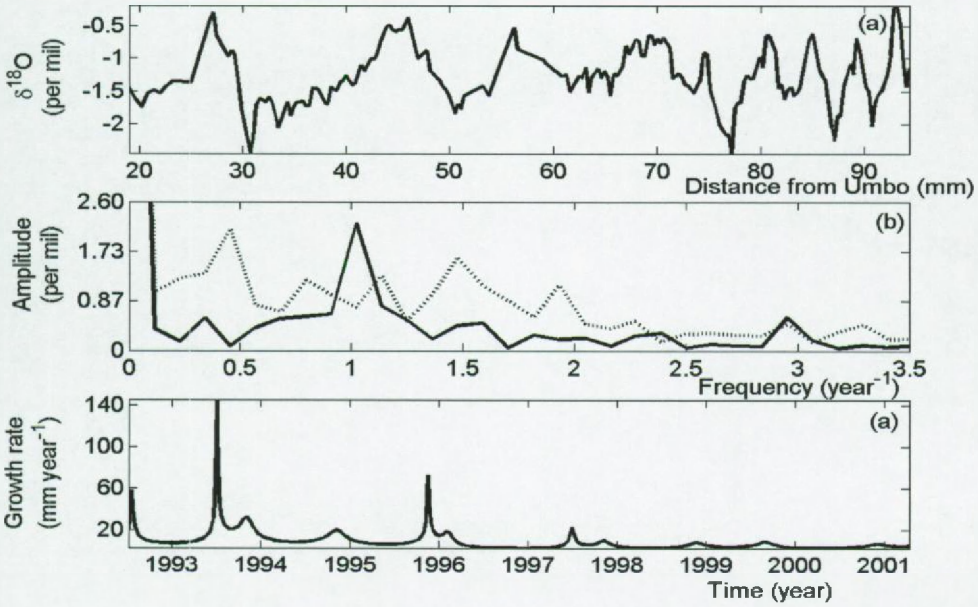

Figure 6: Clam 1: (a) raw data, (b) Fourier spectra before (dotted line) and after the TBD method (full line), and (c) growth rate.

6.3 Case studies

Records, which were already processed in previous chapters are processed again, so that similarities and differences can be compared.

6.3.1 Case study 1: Saxidomus giganteus

As a first example, the $\delta^{18}O$-signals measured in Saxidomus giganteus are processed. Three specimens were sampled from the West coast of the USA in Washington state, named Clam 1 to Clam 3. The large winter-summer variations in temperature are reflected in these signals (Figure 6a to 8a), and just like Wilkinson and Ivany [2002] did, this periodicity will be used to date the observations. Compared with their approach, the parametric time base distortion does not assume a sinusoidal signal, but the broader class of periodic signals (where the

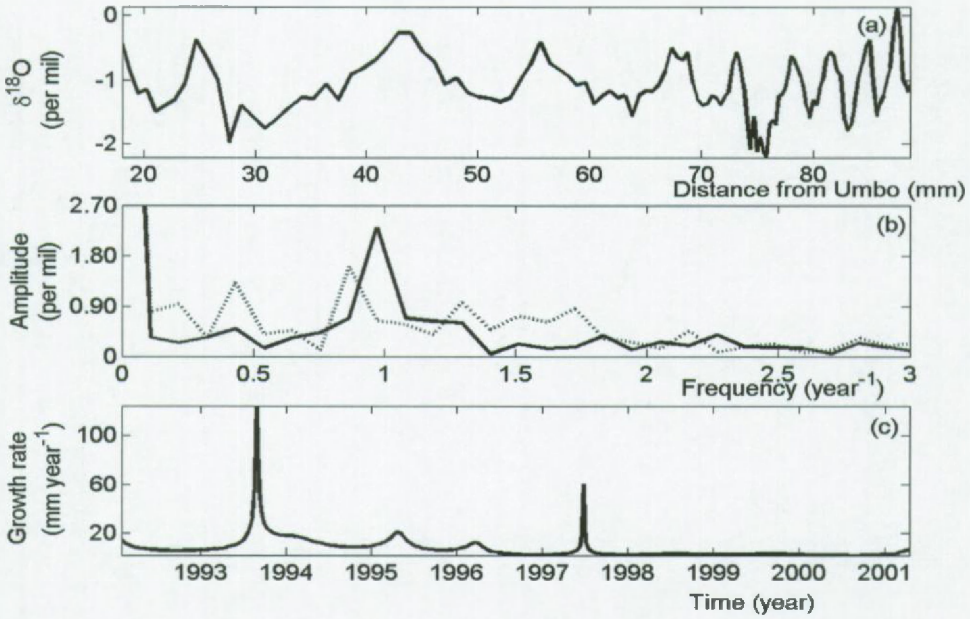

Figure 7: Clam 2: (a) raw data, (b) Fourier spectra before (dotted line) and after the TBD method (full line), and (c) growth rate.

sinusoidal signal is just a special case). In addition tuning parameters, like the width of the window (see chapter 3), are redundant. The variation in growth rate allowed in the model, will depend on the value of the residual cost function and the number of parameters and can differ from one example to another.

To test the algorithm, the maximum model complexity was limited to $(h, b) = (5, 24)$, ringing was reduced by adding $q = 8$ virtual samples at each border. The results are summarized in Table 1: in clam 1 the first and third harmonic are significant, while the

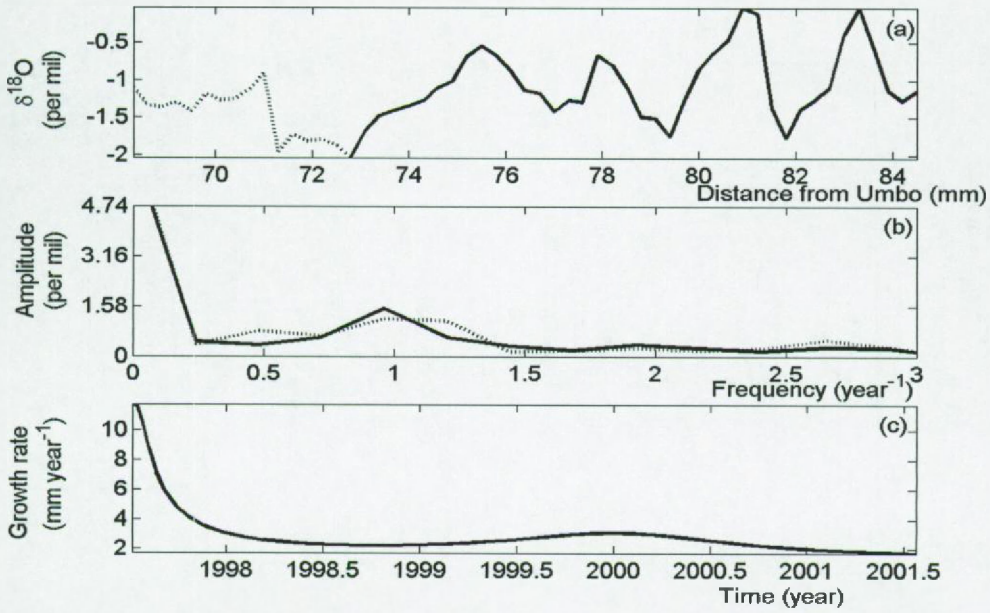

Figure 8: Clam 3: (a) raw data, (b) Fourier spectra before (dotted line) and after the TBD method (full line), and (c) growth rate.

Table 1: Summary of parameters and special procedure used on the *Saxidomus giganteus* examples.

	Clam 1	Clam 2	Clam 3
harmonics	1st, 3rd	1st	1st
number of T.B.D. parameters	16	12	4
Parameters Removed?	yes	yes	no
Negative growth solution used?	no	no	yes

second is absent. This could already be seen in Figure 5. In clams 2 and 3 only the first harmonic is significant. The number of time base distortion parameters was 16, 12 and 4, respectively. In clams 1 and 2 several parameters were removed, as described in paragraph 6.2.5. For clam 3 the emergency procedure against time reversal was necessary: the ten oldest observations are ignored (dotted line in Figure 8a). This is not surprising: the record is no longer periodic at these observations and therefore the reconstructed growth rate is probably largely biased.

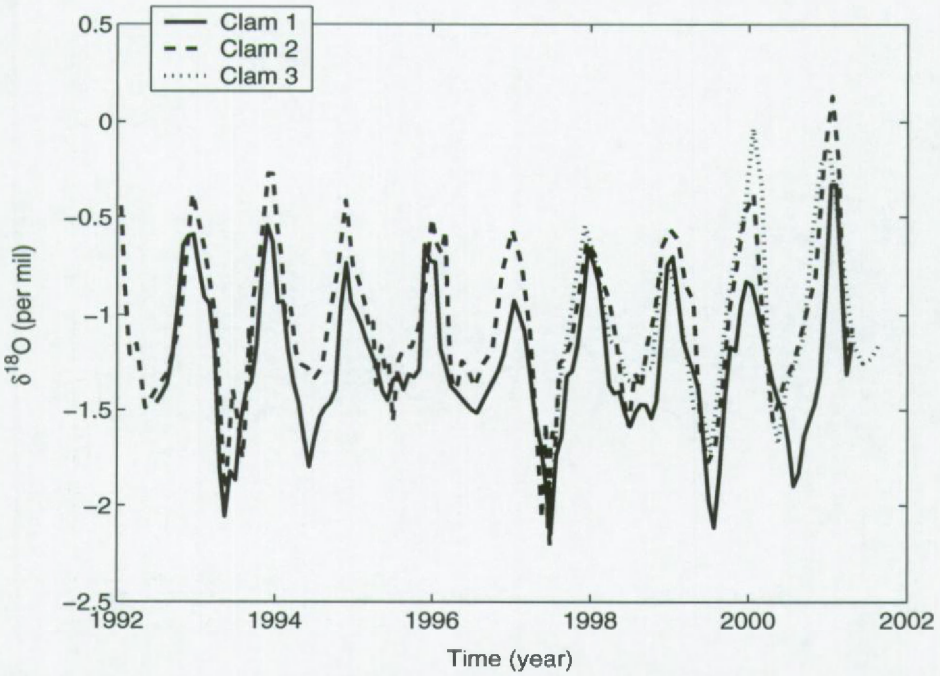

Figure 9: $\delta^{18}O$-records from the three clams after time base correction.

The three time series are compared in Figure 9. The correlation coefficients are given in Table 2.

Table 2: the correlation coefficients, R, between the three clams.

correlation, R (%)	1 vs. 2	1 vs. 3	2 vs. 3
Initial estimated TBD	91	90	88
parametric estimated TBD	88	80	88

Although the correlation is lower for the parametric time base, the difference in correlation coefficients is not significant (number of samples is 43). Note that $\delta^{18}O$ values of the summer of 1998 are less negative than those of 1997 and 1999 in all these three shells, that the values of the winter of 2001 corresponded to the highest values in all records. So, although we have described these signals as periodic, the magnitude of the maxima and

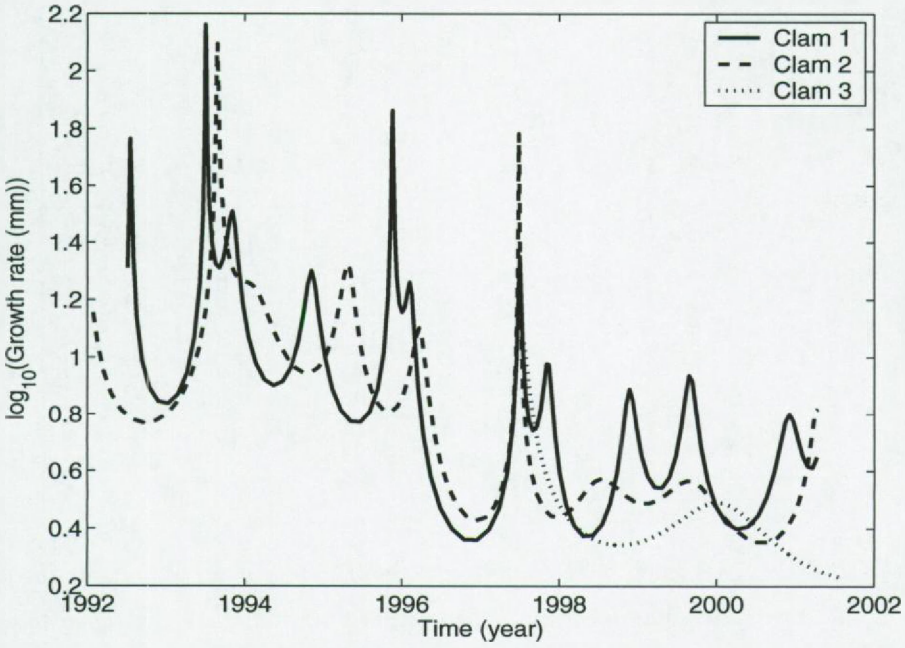

Figure 10: Logarithm of the growth rate of the three clams.

minima within one signal differs. These differences seems not due to stochastic variations. Consequently, an improvement of this procedure could take amplitude modulations into account. This more flexible model could describe periodic signals with varying amplitudes more precisely. These amplitude modulations seems not to differ much when different samples are compared. In other words, we can conclude that the $\delta^{18}O$-signals recorded in these three clams are very similar. So, it is most probably forced by environmental parameters and biological influences, which would vary from sample to sample, can be ignored.

The accretion rates are compared in Figure 10. Note that clam 1 and clam 2 have similar growth rate patterns. Both high peaks correspond and the periods of high growth rate (1992-1996) correspond, while some smaller variations seem out of phase (e.g. the smaller peak around 1995). A possible explanation for this are errors still present in the time base. An alternative explanation is the fact that the accretion rate is an non-linear function of the time

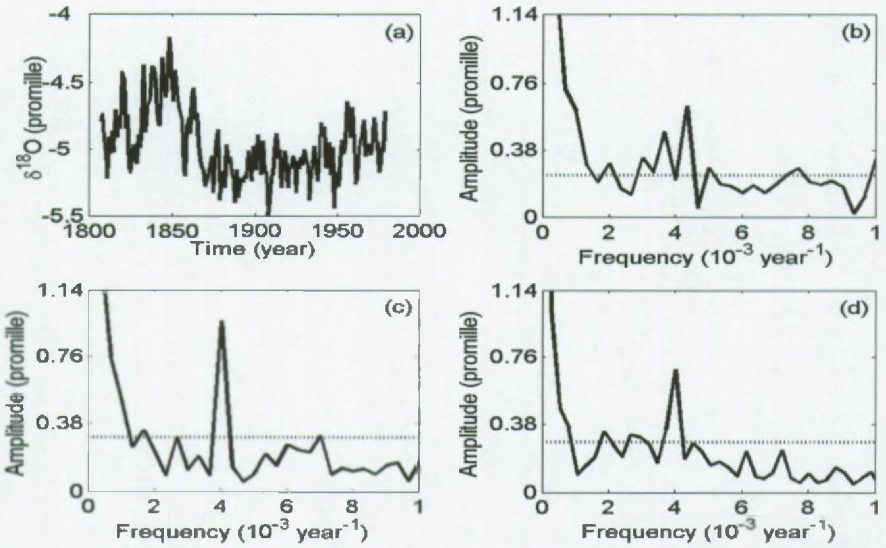

Figure 11: $\delta^{18}O$-signal from the Vanuatu Coral stable isotope data: (a) data on the initial time base, (b) corresponding spectrum, (c) spectrum after time base correction based on phase demodulation and (d) based on parametric T.B.D. approach. The dotted line represents the 95% uncertainty bound (assuming white circular noise).

base distortion parameters. Consequently, errors on these parameters can have a larger influence on the accretion rate than on the time base distortion itself.

In order to model these signal variations better, by removing stochastic variations, a new model can be proposed, where the time base distortion parameters differs for one sample to another, but the signal parameters are optimized on all samples simultaneously. The signal model can than be used to a monitor the environment, where individual and stochastic disturbances are filtered off.

6.3.2 Case Study 2: Vanuatu Coral stable isotope data

A next example handles the Vanuatu Coral stable isotope record [Quinn, 1996]. Details about the sample procedure can be found on http://www.ngdc.noaa.gov/paleo/corals.html. The record was initially dated with the use of annual growth bands. The signal record is shown in Figure 11a and its spectrum in Figure 11b. The dotted line represent the 95% uncertainty

bound, which is estimated by subtracting the trend (sixth order Legendre polynomial) and the periodic signal (one harmonic) from the raw data. From this residual the uncertainty was estimated [Pintelon and Schoukens, 1996] (assuming circular normally distributed white noise). As was previously shown an initial estimate of the improved time base can be found by applying the non-parametric phase demodulation method (see chapter 3). The corresponding spectrum is shown in Figure 11c. In a next step, a parametric model for the signal and T.B.D. are employed. Therefore, the most complex model consisted of $h = 3$ harmonics, $b = 20$ T.B.D. parameters and $q = 10$ extra samples at the borders. The selected model had one harmonic and six T.B.D. parameters (no observations were ignored because of negative growth rates). The raw, trended, data are dated with the improved time base. In Figure 11d its spectrum is shown. Comparing both spectra shows a higher noise level for the latter. The reason therefore is that the non-parametric estimate of the time base used a much wider window (12 lines compared to six T.B.D. parameters). Therefore, much more variation in the time base is taken into account. Unfortunately, this wide spectral window overestimates the T.B.D. and the model selection criterion selected a simpler model, which reduces the stochastic noise influence in the spectrum. At first sight one could prefer the non-parametric phase demodulation method, because of a better signal-to-noise ratio, but then the extra T.B.D. parameters are mostly used to follow the stochastic noise realization, which would differ if the record was measured again. Therefore, maybe surprisingly, the simpler model is chosen. An alternative improvement consists in estimating the uncertainty on the raw data, so that a model selection criterion with known noise model could be used. The latter penalizes more complex models less than this model selection criterion.

To conclude, an initial time base was constructed by Quinn *et al.* [1996] based on annual growth bands. Both the phase demodulation and this parametric approach can improve this time base with respect to the harmonic content. The fact that the signal-to-noise-ratio is higher for the non-parametric approach, suggests that this model was too complex, compared to the disturbing noise.

6.3.3 Case Study 3: Kenyan mangrove tree

In this case study, the vessel density in a *Rhizophora mucronata* is examined (see chapter 4 or [Verheyden *et al.*, submitted]). The stem disc was collected in November 1999 from Gazi

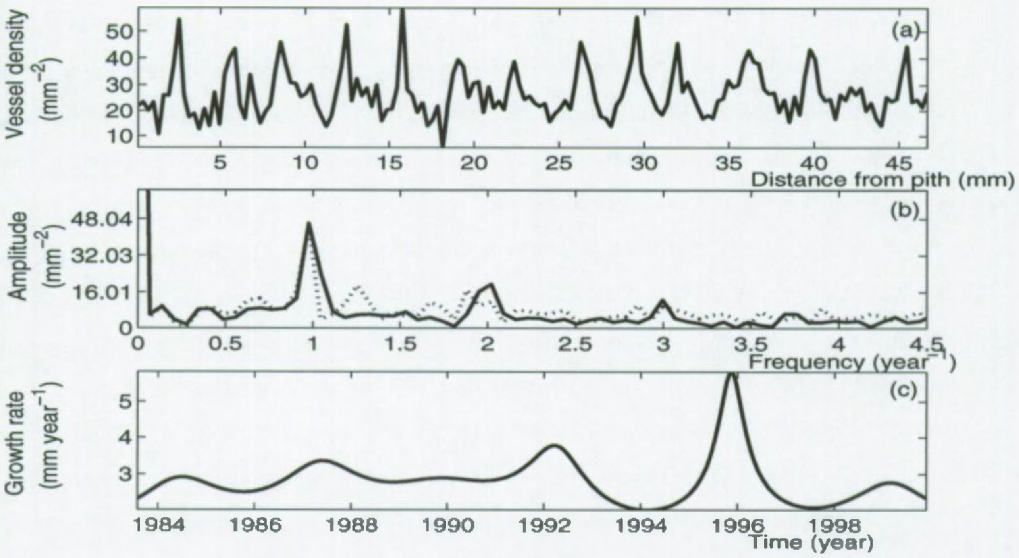

Figure 12: Vessel density measured in a mangrove tree: (a) the vessel density as function of the distance from pith, (b) spectrum without time base correction (dotted line) and with a correction (full line) and (c) growth rate as function of time.

Bay, Kenya and is now part of thexylarium of the Royal Museum for Central Africa (RMCA), Tervuren, Belgium. The sample was air-dried and transverse sections were sanded. The vessel density was measured directly on the polished wood discs, along a radial transect from bark to pith in adjacent windows. Window size was set to 300 μm height and 2100 μm width. The vessel density was measured at an optical magnification of 12x using image analysis software and recalculated to the number of vessels per squared millimeter. This experiment was carried out in order to examine which of the environmental parameters influenced the vessel density. However, this discussion lays behind the scope of this thesis. The raw data as a function of the distance from pith are shown in Figure 12a.

In the scope of this chapter, we will estimate the time base, assuming that the record is periodic. For that, the maximum model complexity was $(h, b) = (10, 40)$. Both borders were enlarged with, $q = 13$ observations. The model with the lowest model section criterion consists of three harmonics and 12 time base distortion parameters. the presence of the three harmonics is in correspondence with the three peaks that appear in the spectrum of

the observations after the time base correction. The shape of the growth rate cannot, at first sight, be related to environmental parameters, neither to the age effects. If much more complex growth rate models would be used (e.g. $b = 26$), yearly patterns in growth rate become visible, but without noise models, these variations are insignificant.

6.4 Conclusion

In this chapter an algorithm was explained in details, which reconstructs variations in growth rate assuming a periodic signals model. Attention was given to the strategy used to avoid local minima; reducing the Gibbs phenomenon and the model selection. Next, the estimation of the time base is exemplified on some case studies. The main benefit of this method is that it uses objective criteria to estimate the complexity of the time base. The main disadvantage is the assumption of a period signal model. Many signals are not properly described by this kind of model and should be described by more complex models like an amplitude modulated periodic signal or an periodic signal with a trend, etc... However, this strategy can be applied on such more complex signals.

References

[1] De Brauwere A., 2003. Interpretation of optimization algorithms applied to biogeochemical models. Master Thesis Vrije Universiteit Brussel, Faculty of Science.

[2] Martinson D.G., Menke W. and Stoffa P., 1982. An inverse approach to signal correlation. Journal of geophysical research 87(B6), 4807-4818.

[3] Pintelon R. and Schoukens J., 1996. An improved Sine-Wave Fitting Procedure for Characterizing Data Acquisition Channels. IEEE transactions on instrumentation and measurement, 45(2), 588-593.

[4] Quinn T.M., Crowley T.J. and Taylor F.W., 1996. New stable isotope results from a 173-year coral record from Espiritu Santo, Vanuatu. Geophysical Research Letters, 23[23], 3413-3416.

[5] Schoukens J., Pintelon R. and Vandersteen G., 1997. A Sinewave Fitting Procedure for Characterizing Data Acquisition Channels in the Presence of Time Base Distortions and Time Jaulin L., Kieffer M., Didrit O., and Walter E., 2001. Applied Interval Analysis, Springer.

[6] Verheyden A., De Ridder F., Schmitz N., Beeckman H., Koedam N, subm. High-resolution time series of vessel density reveal bi-modal distribution and correlation with climate variables in Kenyan mangrove trees. Submitted to Oecologia.

[7] Wilkinson B.H. and Ivany L.C., 2002. Paleoclimatic inference from stable isotope profiles of accretionary biogenic hardparts - a quantitative approach to the evaluation of incomplete data. Palaeography, palaeoclimatology, palaeoecology 185, 95-114.

CHAPTER 7

A COMPARATIVE STUDY OF THE TIME BASE CONSTRUCTION METHODS USED IN PALEO-ENVIRONMENTAL ARCHIVES

Keywords - Time series, anchor point method, tie point method, control point method, dynamical programming, phase demodulation, accretion rate, time base distortion.

Abstract - Several methods developed to reconstruct time bases for proxy records are tested on simulations. Each of these methods is based on different assumptions and depending on the presence of stochastic and/ or systematic errors their performance will differ. In this chapter, the methods are classified and validated in a systematic manner. Three types of orbital tuning methods are evaluated, i.e. (i) anchor point method, (ii) a correlation maximization method and (iii) the parametric method developed by Martinson *et al.* This method was refined in order to improve convergence during the optimization. These methods are compared with (iv) the time domain method developed by Wilkinson and Ivany, with (v) a frequency domain method, based on a phase demodulation and with (vi) the parametric method developed in chapter 6. Two types of sensitivity studies were performed: first, the sensitivity against stochastic noise was tested and next the robustness against some model errors was evaluated. From these simulations, it can be concluded that the orbital tuning methods are more sensitive toward model errors and parametric models perform better in the presence of both types of errors.

7.1 Context

In this chapter, the two proposed methods, i.e. (i) the non-parametric phase demodulation method and (ii) the parametric time base distortion approach, are compared with other methods, described in literature. Where are these two approaches situated according to solutions proposed in literature and what is their performance?

In order to bring some clarity in the range of proposed methods, we have first divided them in three classes: (i) orbital tuning methods, (ii) signal model methods and (iii) hard models. In the first class the investigator assumes that a dated time series is known. Next, the time base is build so that the measured proxy matches best to this known time series. This approach has two weak points: what information can still be extracted from a time series which is forced on a certain profile? Furthermore, a large bias is possible, if model errors are present. The signal model class makes an assumption about the class of functions to which the time series belongs, but within this class, it still has some freedom, because some parameters have to be tuned. Finally, for hard models, the investigator uses particular knowledge about the accretion rate or signal. Therefore, only another specialist in that particular field can judge on the quality of these assumptions and this model class is not further discussed here.

Notice that the parametric time base distortion, developed in chapter 6, is similar to the famous, but hardly used, method proposed by Martinson *et al.* in 1982. The parameters in the signal model are fixed, but the idea of a parametric representation of the time base was already present in their work. They proposed to model the total time base by e.g. a Fourier basis, which complicates the initial value problem. Possibly, this has been one of the reasons why this approach has not been used more often in the past. This problem is bypassed by separating the distortion in the time base from the average accretion rate, as we suggested. In addition, the phase demodulation is very well suited to generate initial values. A second problem, which was not considered, is a criterion to fix the number of parameters in the time base.

7.2 Introduction

During the last 25 years several methods, based on different assumptions, have been proposed to construct time series and accretion rates. In this chapter these methods are classified, analyzed, criticized and compared. A short description of each method is given, where special attention is directed to the assumptions made in each method and to the tuning parameters used.

First of all, one should realize that without any assumption about the signal model, it is not possible to derive a dated time series starting from a given distance series (proxy). Special attention is directed to the assumptions made in each method, after a short description of each method is given.

The methods used to reconstruct time bases are divided into three classes (Figure 1 structures the discussed models, limitations and solutions):

1. *Orbital tuning methods* generate an accretion profile by assuming a similarity between the distance-record with a postulated target function. So an assumption is made about the time record. In a next step, this similarity is used to optimize a time base, which can be performed in many different ways. We discuss three strategies, i.e. the anchor point method [Paillard *et al.*, 1996], the correlation maximization [Yu and Ding, 1998; Lisiecki and Lisiecki, 2002] and Martinson *et al.*'s method [Martinson *et al.*, 1982a; Martinson *et al.*, 1982b; Martinson *et al.*, 1987].
2. *Signal models* exploit some symmetry property of the time series and generates an estimated time base and a time series. The exact profile of the time series is not known or assumed to be known in advance, but the class to which the time series belongs is postulated by the user. We will focus on periodic functions. Three methods are discussed: one works in the time domain (Wilkinson-Ivany's method [Wilkinson and Ivany, 2002; Ivany and Wilkinson, 2003]), the other in the frequency domain (phase demodulation method (see chapter 3)) and the last one is a parametric time base distortion approach developed in chapter 6.
3. *Hard models*: the accretion rate is derived from biological, physical and/or chemical

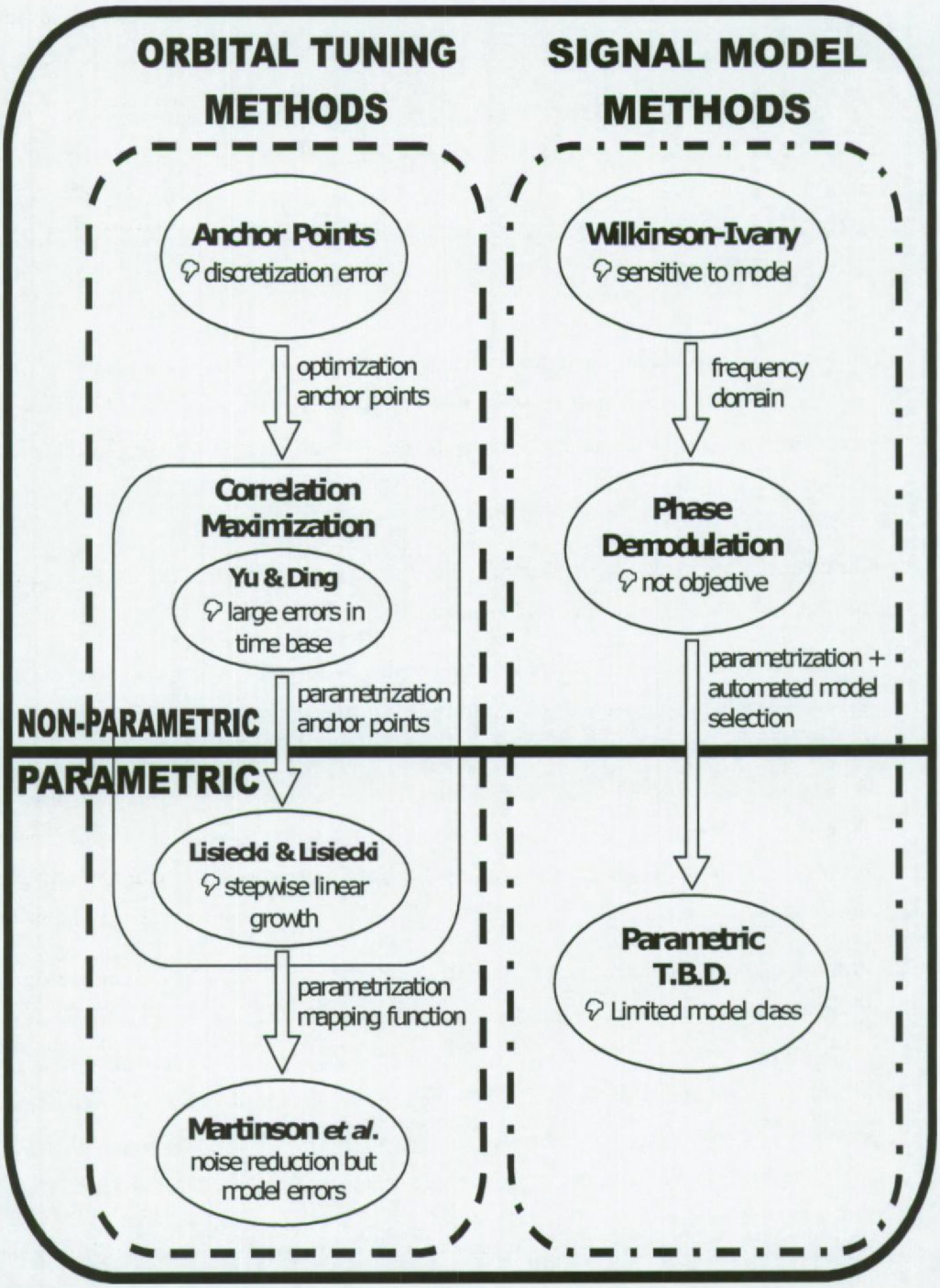

Figure 1: Structure of the different approaches, the limitations and solutions.

properties of the specimen. Examples are the accretion rate model derived by Kaufmann [2003] or the ice accumulation models derived by Jouzel *et al.* [1996]. Once the accretion rate is calculated, the time series can be derived. This class is not validated here.

The methods can also be classified in parametric (number of parameters is independent of the amount of data) and non-parametric methods (number of parameters increases with the amount of data). According to our knowledge the only parametric method proposed in literature is Martinson *et al.*'s method [1982a]. All others are non-parametric.

The precision and accuracy of a time series will largely depend upon the assumptions made. Therefore, special attention is given to the robustness of the method. This is done by testing each method on the same simulations, where

1. only stochastic errors are present to illustrate the sensitivity to noise; and
2. model errors are present to check the influence of a violation of the assumptions in the absence of noise.

Because the sensitivity of a specific method depends on the model error introduced, several different model errors have been used: first, the methods are tested on a periodic signal, with the orbital tuning methods tuned on a sinusoidal target function; next, an amplitude modulated signal is processed; further, a trended signal is processed; and finally a sinusoidal signal with an hiatus is tested. All these model errors can be found in real world dating applications. However, one should keep in mind that much more model errors can be imagined and that this list of four model errors is far from complete. Another limitation is that we have used sinusoidal proxies and target functions, which simulate annually resolved archives, like corals, sclerosponges, bivalves, etc... but that in sediment cores for example, the proxy and target function can be more complicated. To which extend our conclusions will hold when applied to such records is hard to predict.

We have updated some of the methods proposed in literature, such that each algorithm can be compared based on the assumptions made and not on the way these algorithms have been implemented.

7.3 Class 1: Orbital tuning methods

For these methods the investigator starts on the one hand from a proxy, as a function of the distance and on the other hand from a known profile of time series, so-called target or template function. The accretion relates both and is not known in advance. In general, this function is called the mapping function and three sub-classes of methods will be described, which can be used to construct or estimate the accretion, i.e. anchor point, correlation maximization and Martinson *et al.*'s method.

7.3.1 General idea

"Numerous investigators have shown that the earth's climate responds linearly, to some unresolved extent, to variations in the earth's orbital geometry. Since this climate response is continuously recorded in deep-sea sediments by climatically sensitive parameters, the orbital/climate link provides a unique opportunity for establishing a late Pleistocene geo-chronology. This can be achieved by considering the known history of the orbital forcing as an orbital "metronome", and the related portion (climate response) embedded in the geological data, as a "metronomic record". The metronomic record is distorted while being recorded in deep-sea sediments by such things as changes in sedimentation rate. Provided the lag between the orbital forcing and climatic response is constant, tuning the climatic response to keep pace with the orbital metronome will yield an absolute chronology" (after [Martinson *et al.*, 1982a]).

7.3.2 . Anchor, tie or control point method

7.3.2.1 Idea

In this non-parametric method the date of some observations is known, these observations are called the anchor, tie or control points. To our knowledge this method is very frequently used, but, surprisingly enough, we were unable to find a description of the properties of this method in literature [Paillard *et al.*, 1996]. The dates of the observations between these anchor points can be estimated employing an interpolation technique (mostly a linear interpolation, which means a linear growth is assumed between the anchor points). The

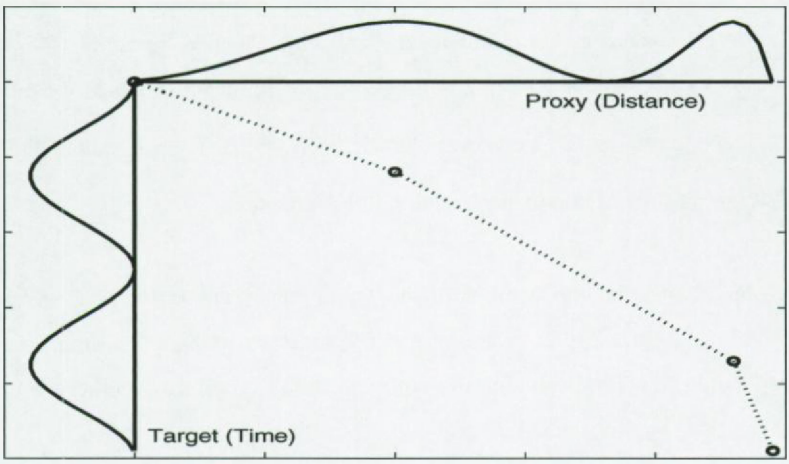

Figure 2: Visualization of the Anchor Point method. Horizontally the proxy is shown as function of a distance grid; Vertically the target function is shown as function of time. The date of some observations is known (in this example the first and last observation and the maxima, shown by 'o') and between these anchor points a linear accretion is assumed (dotted line).

dating of the anchor points can be based upon the target function or on other events registered in the proxy, like growth bands, well dated volcanic eruptions, etc.... Compared with the other orbital tuning methods, this method tunes the time-series by the judgement of the investigator (which makes it subjective).

Figure 2 illustrates this method: horizontally a proxy record as function of a distance grid is shown; vertically the target function is shown. To focus the mind the proxy record could be the $\delta^{18}O$-signal recorded in a bivalve and the target function could be the sea surface temperature, directly measured as function of time. The annual maxima can be used as anchor points (shown by the 'o' in Figure 2) and in between a constant accretion rate is assumed (shown by the straight dotted lines along the bissectrice in Figure 2).

7.3.2.2 Assumptions made by the anchor point method

1. The date of each anchor point is known, with absolute certainty. This means that no procedures are implemented in order to reduce the influence of stochastic noise;
2. The first and the last observation are dated, again with absolute certainty;
3. Between the anchor points the accretion rate is constant.

Tuning parameters: the number of anchor points is chosen by the investigator. For the example of Figure 2, one investigator can decide to use the annual maxima, while another one would use maxima and minima. In opposition to most of the other methods, any record can be dated by this method, as long as some anchor points are known.

7.3.2.3 Stochastic noise

All methods will be tested on a simulation where only stochastic error and no model errors are present. Any measured record is disturbed by this kind of errors and the better a method works, the less stochastic errors will be propagated into the time base. The 'true' signal is sinusoidal, without variations in accretion rate. Consequently, all reconstructed variations in accretion rate are errors generated by the particular method itself. The signal spans seven years and is sampled 50 times (full line in Figure 3 - frequency of 1 year^{-1}, sample period of 7/50 year and a distance grid of 7 mm, corresponding to a constant accretion rate of 1 mm year^{-1}). This signal is disturbed by white normally disturbed noise with a signal-to-noise-ratio (SNR) of 11.6 (see 'o's om Figure 3).

Figure 4 shows the accretion rate[1] estimated by the anchor point method (full line). The dotted line at one mm year^{-1} is the theoretical accretion rate. Because the first and last observation are dated, the number of annual maxima is known. The root-mean-square (RMS) of the error is 9%. This error can have several reasons:

1. the accretion rate is estimated as $D_s/(t_n - t_{n-1})$, where D_s is the distance between two subsequent samples (assumed to be constant) and t_n is the time instance estimated by the respective method.

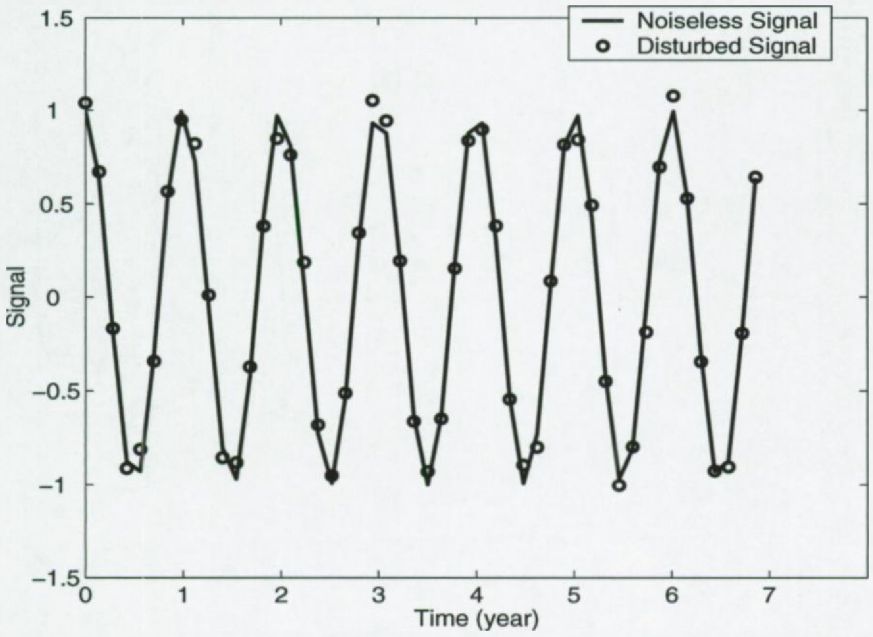

Figure 3: The simulated signal (full line) is shown, disturbed by normally distributed white noise ('o').

1. Suppose a record without disturbing noise. Because this record is sampled at discrete positions, the exact maxima will probably fall between two subsequent samples. This causes a discretization error, i.e. an error related to the finite sampling resolution. So even in the theoretical case of no stochastic noise, the anchor point method will not work perfectly.

2. Besides this discretization error, the disturbing noise changes the numerical value of the observations and consequently a wrong observation can be selected as the maximum, which can cause an additional error.

A logical extension of this model could consist of a signal model, e.g. a periodic signal assumption, so that interpolation could be implemented in order to estimate the maxima. This would avoid the discretization error and the influence of a particular noise realization. However, these extension would finally lead to the re-invention of Martinson *et al.*'s method.

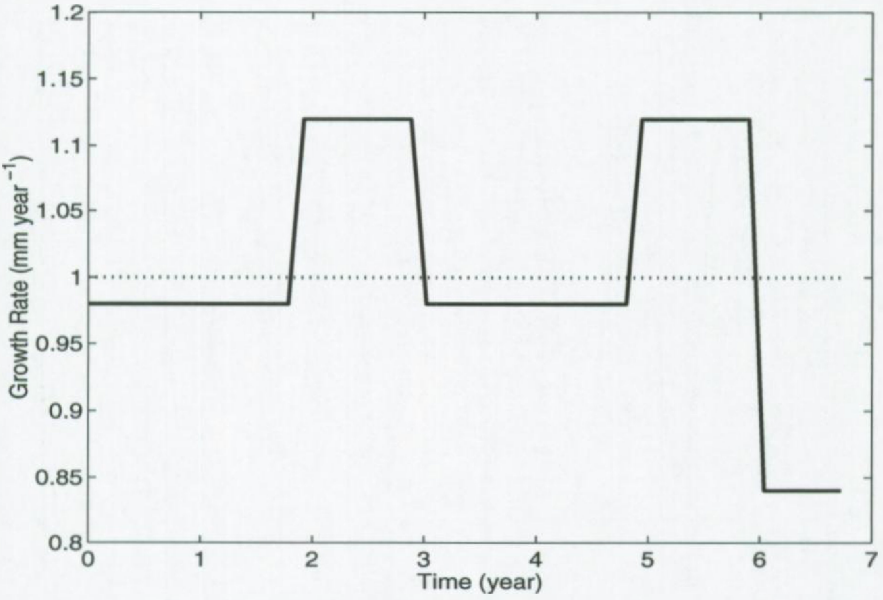

Figure 4: Accretion rate, estimated by the anchor point method on a simulation (dotted line is the true accretion rate, full line is the estimated accretion rate) disturbed by stochastic noise.

To conclude, because the dates of the anchor points are defined by the user, noise on the measurement is ignored, which makes this method vulnerable to stochastic fluctuations on the measured data. In addition a discretization error is unavoidable.

7.3.2.4 Model errors

Figure 5 shows the different simulated record signals (full lines) and the reference signals assumed by the orbital tuning methods (dotted lines). No stochastic noise was added in order to separate model errors clearly from stochastic errors. The signal parameters are identical to those used in the stochastic simulation. The accretion rate was constant again, except for (d) where an hiatus is introduced. So except for the hiatus, any variation in the accretion rate is a bias.

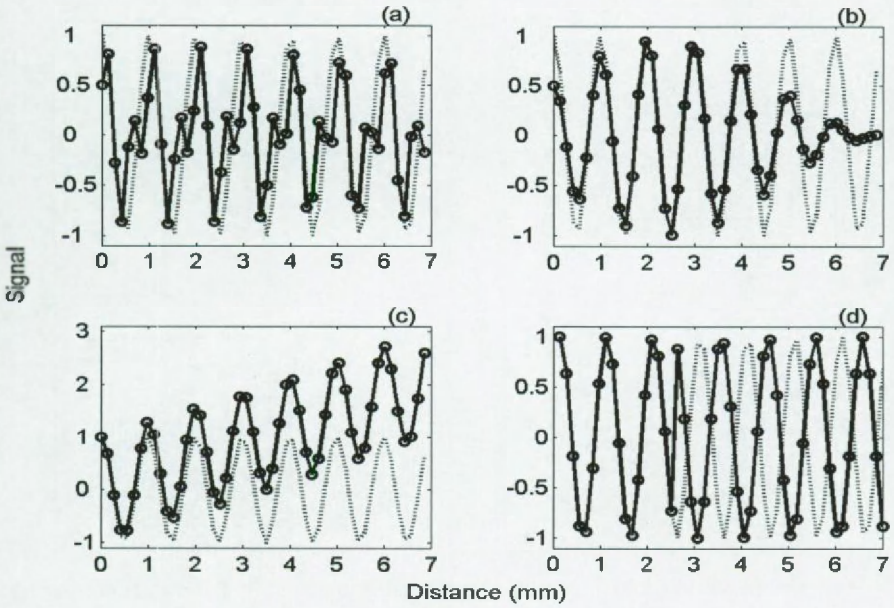

Figure 5: Some model errors: (a) a periodic signal (full line) instead of a sinusoidal signal (dotted line), (b) an amplitude modulated signal, (c) a signal with a trend and (d) a signal with a hiatus.

Periodic Signal:

In the first case (Figure 5a) the anchor point method is tested in the presence of a second harmonic. The yearly maxima were chosen as anchor point and the reconstructed growth rate is shown in Figure 6a. Note that

1. The first and the four last observations were dropped, because the record did not start and stop at a maximum, so these observations could not be dated.
2. Even without stochastic noise the anchor point method cannot estimate the accretion rate exactly, because of the discretization error. To make this more clear, the number of observations between two subsequent maxima are counted. For the first five years seven observations were made each year; while for the sixth year eight observations were made. Consequently, the estimated accretion rate is higher in this year.
3. The RMS error is 5%. So, despite this discretization error the anchor point method can

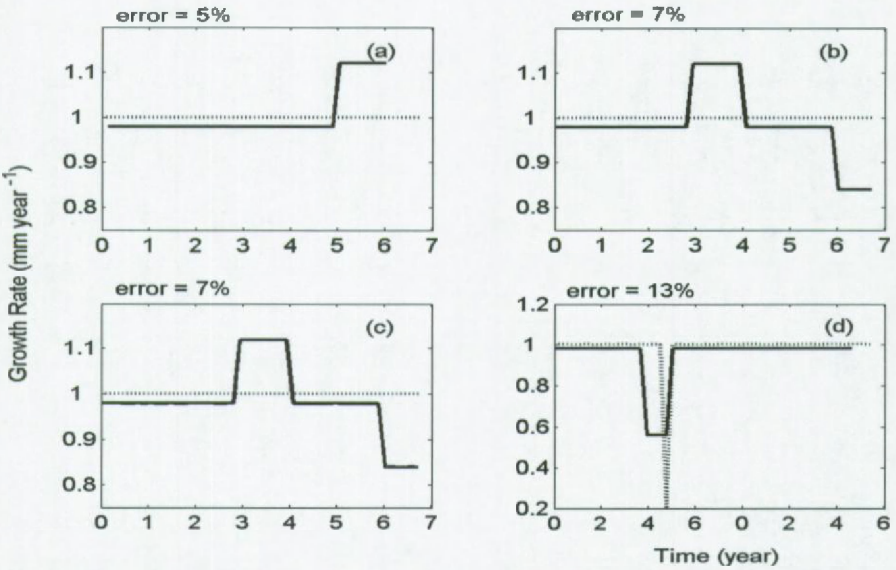

Figure 6: Accretion rate, estimated by the anchor point method on a simulation (dotted line is the true accretion rate, full line is the estimated accretion rate) where the signal is disturbed by (a) a periodic signal instead of a sinusoidal signal, (b) an amplitude modulated signal, (c) sinusoidal signal + a linear trend and (d) hiatus.

handle periodic signals.

Amplitude modulated signal:

Figure 5b shows an amplitude modulated signal. Each year the maximum was chosen as anchor point and the reconstructed accretion rate is shown in Figure 6B. The same conclusion as for the periodic signal can be drawn: the root mean square error is 7%. As can be seen in Figure 6B, most of the mis-match is due to a discretization error, but on the whole the anchor point method seems to perform well on amplitude modulated signals.

Trended Signal:

Figure 5c shows a trended signal: a linear trend was added to the sinusoidal signal. The same anchor points as in the amplitude modulated signal are used and thus the same accretion rate was found: root mean square error of 7%, caused by the discretization error.

Hiatus:

Figure 5d shows a signal with a hiatus. Technically, the hiatus is caused by removing the 18th observation and can be interpreted as a sudden discrete growth stop during the time that this observation should have been formed. Because the hiatus did not remove a maximum, the same maxima were selected by the anchor point method and the hiatus was reconstructed within the precision allowed by the number of anchor points. The position of the hiatus was identified, but the magnitude was averaged over the period between the two neighboring anchor points. This results in a root mean square error of 13%.

7.3.2.5 Conclusion

The anchor point method is not very robust against stochastic noise, but it performs reasonably well in the presence of model errors, at least those which were tested here. However, the main limitations are

1. that a discretization error limits its precision in all cases;
2. the method's subjectivity, caused by the user's decisions: which and how many observations are used as anchor points and which corresponding dates are chosen.

7.3.3 Correlation maximization method

The two disadvantages of the anchor point method can be circumvented by the implementation of an optimization. The correlation maximization methods and Martinson *et al.*'s method are such optimizations. The latter is discussed in the next paragraph. In the correlation maximization, a number of anchor points is chosen, e.g. in [Yu and Ding, 1998] all dates are used as anchor points, while only a limited number is used in [Lisiecki and Lisiecki, 2002].

7.3.3.1 Idea

The date corresponding to each anchor point is adapted, so that the correlation between the proxy-record and a given target function is maximized. If the number of observations

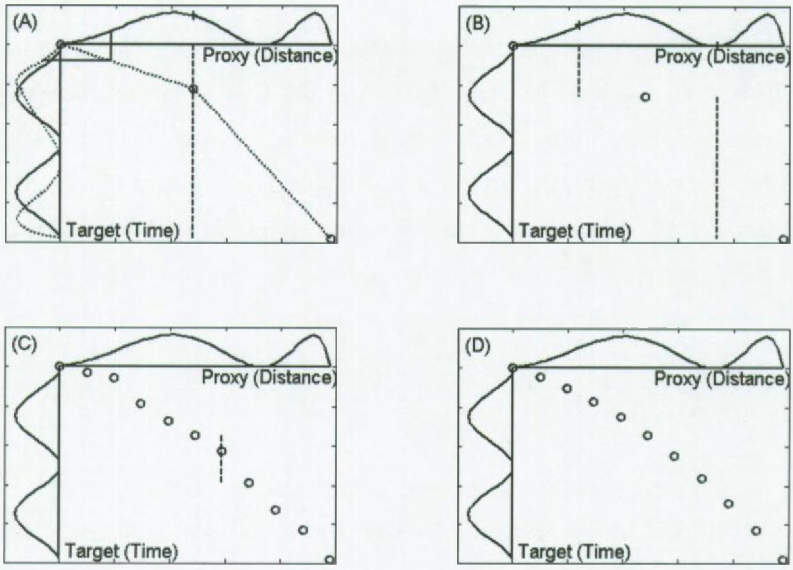

Figure 7: Visualization of the dynamical modeling method. Horizontally the proxy is shown as function of a distance grid. Vertically the target function is shown as function of time.

increases, the number of parameters (time instances to be estimated) changes, so this method is non-parametric.

The basic strategy behind the correlation maximization is (based on [Yu and Ding, 1998])

1. Step A: The first and last observation are used as anchor points. In the first step (Figure 7a) the date of the median observation (shown by '+') is estimated, by varying its date (shown by the vertical dashed line) between that of the first observation to that of the last observation. For each step the correlation with the target function is calculated, assuming a linear accretion between the median and respectively the first and last observation. The date achieving the highest correlation is selected.
2. Step B: the proxy-record is sub-divided into two parts: the first spans from the first observation to the dated median observation and the second spans from the median observation to the last one. The same procedure is repeated for each sub-record (Figure 7b): the date of the median observation, within the sub-record, is changed from the first to the last

date and for each date the correlation is calculated. Identically to Step A, the dates corresponding to the highest correlation are selected. This procedure is repeated until all observations are dated. Step A and Step B can be skipped if a good initial guess for the dates is available.

3. Steps C and D: due to the assumption of a constant accretion rate between the first/last observation and the median observation, the error on the date of the first dated observation (median) will be larger than on the last one (because in the latter one, the time span over which a linear accretion is assumed is much smaller and thus the potential error is also smaller). Therefore, the date of each observation is estimated again in the interval between its neighboring observations (Figure 7C). This procedure is repeated until the correlation has converged toward a maximum (Figure 7D).

The procedure proposed by Yu and Ding [1998] is similar to the one described above, but it starts from a set of anchor points as an initial accretion profile, while here no initial values are provided. Next Steps C and D are used to maximize the correlation.

Lisiecki and Lisiecki [2002] sub-divide the proxy record until the sub-records have reached a certain length (tuning parameter). Further, a mis-match between the sub-record and target function is allowed, and may change from sub-record to sub-record in order to tie both closer in certain regions. Reformulated, the number of anchor points is fixed in advance. Consequently, this adaptation makes this approach parametric, e.g. if the sample frequency would be doubled the number of anchor points would remain the same.

7.3.3.2 Basic assumptions made in the correlation maximization

1. The target function is known;
2. The first and last observation are dated;
3. The proxy and the target function are related by a static linear relation (so that the correlation can be used as criterion);
4. The order of the observations is not altered during the procedure.

7.3.3.3 Tuning parameters

None for the method described by Yu and Ding.

For Lisiecki's algorithm:

1. the number of observations within one sub-record is chosen by the investigator, i.e. he/she fixes the number of anchor points to be optimized in advance;
2. a certain mismatch, due to noise on the observations, between a sub-record and the target function is allowed.

7.3.3.4 Stochastic noise

The simulation described in Figure 3 is used. The RMS error is 35%, which may seem surprisingly high. The reason for this high error is that the effect and presence of stochastic noise is neglected. Consequently, a large part of the noise will be transformed into the accretion rate, which maximizes the correlation between the signal and target function as demanded by the algorithm. As an unwanted result the error on the accretion rate will become larger.

A remedy is already proposed by Lisiecki and Lisiecki [2002]: lower the number of anchor points. This method ends up somewhere in between the anchor point and the method proposed by Yu and Ding, depending on the chosen tuning parameters. Figure 8 shows the results when seven anchor points are used (one for each year, equal to the number of anchor points used in the anchor point method). The error is now acceptable and the difference with the true accretion rate falls within the uncertainty bounds.

7.3.3.5 Model errors

Tested on the simulation examples with model errors, shown in Figure 5, the correlation maximization method performs poorly on most model errors, i.e. on the periodic signal (RMS error of 10^7%), the amplitude modulation (206%) and on the trended signal (65%). Only for the hiatus, is its performance is very good (0.35%). The target function (dotted lines

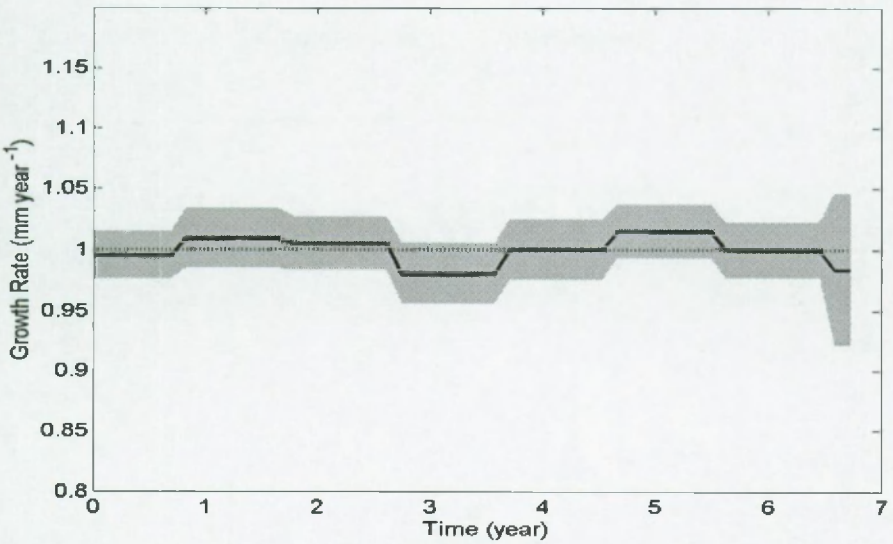

Figure 8: Accretion rate, estimated by the Lisiecki and Lisiecki method on a simulation (dotted line is the true accretion rate, the fill line is the estimated one) disturbed by stochastic noise. The grey band is the 95% uncertainty bound estimated with a Monte-Carlo simulation based on 100 runs.

in Figure 5) was sinusoidal, while the signal was not for the first three cases. So, maximizing the correlation leads often to large errors when the target function and observation profile differs. In Figure 9 only the reconstructed accretion rate is shown for the hiatus simulation. The hiatus is described very well by this method, because both, the signal and the target function, can be matched without introducing new errors. For the other methods the model error changed the similarity between the signal and target function. The 0.35% error that remains, is a numerical error and can further be reduced.

Again, the correlation maximization method proposed by Lisiecki and Lisiecki performs better in the presence of model errors, as can be seen in Figure 10. In the presence of overtones (a) the error is significant (19%), while it is negligible in the presence of amplitude modulations (b) or trends (c). In the presence of an hiatus (d), the error is slightly larger than for the anchor point method, which may be surprising, because here no discretization error is made. This difference is due to the position of the anchor points, which are now further away form the hiatus.

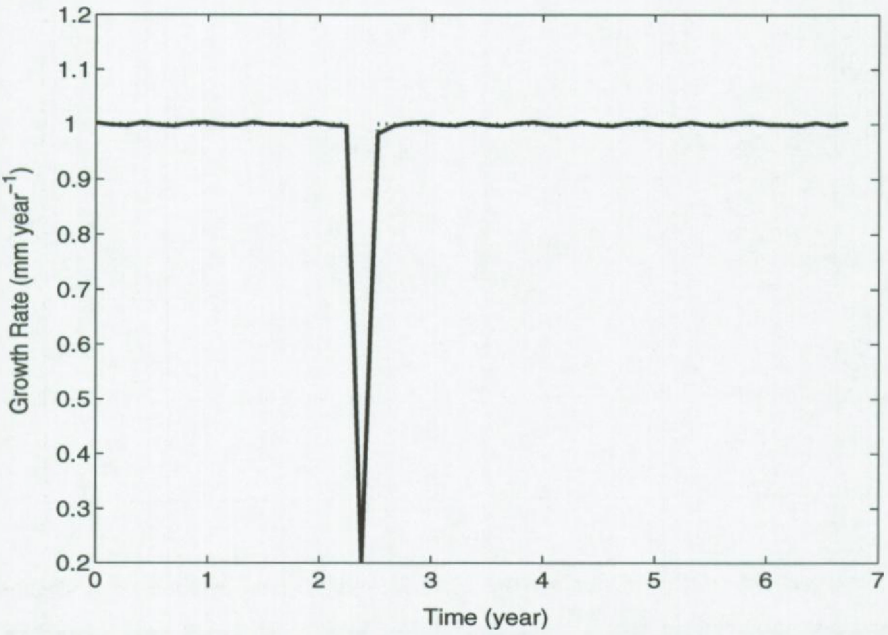

Figure 9: Estimated accretion rate by the correlation maximization method in the presence of a hiatus: dotted line is the true accretion rate, the full line is the estimated accretion rate.

7.3.3.6 Conclusion

Whereas the anchor point method fixes each anchor point in advance the correlation maximization methods optimizes these dates. The price for this optimization is that a target function must explicitly be given. The method proposed by Yu and Ding optimizes the date of each individual observation. Consequently, the stochastic measurement noise on one single observation can have a large influence on the estimated accretion rate. Because every real measurement is disturbed by stochastic noise, this method is not recommended. This problem is remedied by Lisiecki and Lisiecki, who limited the number of anchor points. In addition, the method proposed by Yu and Ding does not perform well in the presence of model errors, while the correlation maximization of Lisiecki and Lisiecki achieves a reasonable performance in reconstructing growth rates.

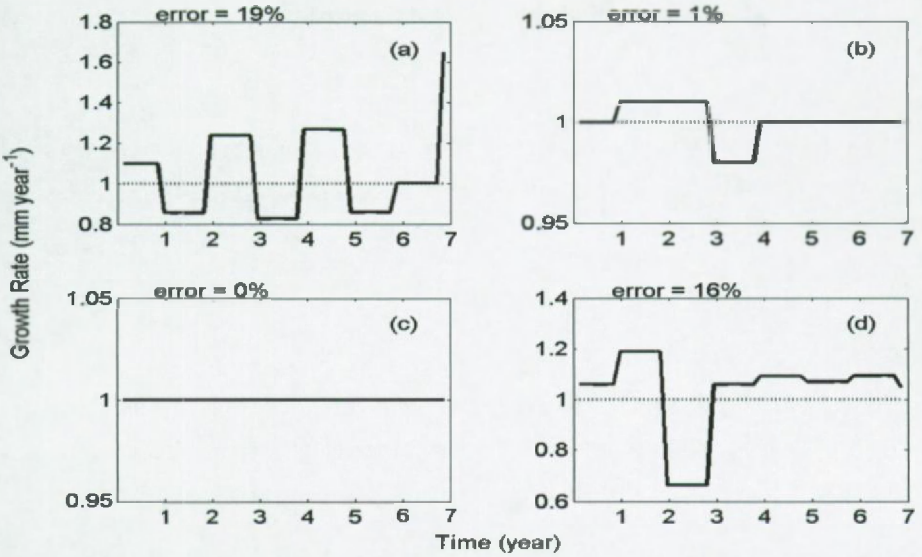

Figure 10: Accretion rate, estimated by the Lisiecki and Lisiecki method on a simulation (dotted line is the true accretion rate, the full line is the estimated one), where the signal is disturbed by (a) a periodic signal instead of a sinusoidal signal, (b) an amplitude modulated signal, (c) a sinusoidal signal + a linear trend and (d) an hiatus.

7.3.4 Martinson et al.'s method

So far, the growth rate in between anchor points is assumed to be constant. The procedure proposed by Martinson *et al.* [1982] will avoid such a simplification by employing a parametric time base model. This method starts from a measured proxy record along a distance grid and a known target function along a time grid. The mapping function, related to the accretion rate, is unknown and is expanded in a set of basis functions. The parameters corresponding to each of these basis functions are optimized. Because the number of parameters is much smaller than the number of observations, the influence of stochastic noise is minimized and model errors will partly be suppressed.

7.3.4.1 Idea

Consider for instance an periodic target function, described by

$$s(t) = A_0 + \sum_{k=1}^{h} A_k \sin(k\omega t) + A_{k+h} \cos(k\omega t) \tag{7-1}$$

where t is a time variable, A are known coefficients, ω is a known radial frequency and h is the number of harmonics, which is defined by the user. The analog signal, $s(t)$, is sampled at time instances

$$t_n = nT_s + g(n)T_s \tag{7-2}$$

with $n \in \{0, \ldots, N-1\}$, with N the number of samples, T_s the sample period and $g(n)$ the unknown mapping function. The latter is parametrized by a sum of 'simple' functions [Martinson *et al.*, 1982a] with unknown coefficients, e.g. a Fourier series

$$g(n) = \sum_{l=1}^{b/2} B_l \sin(2\pi l T_s) + B_{l+b/2} \cos(2\pi l T_s) \tag{7-3}$$

with B the unknown coefficients and b the number of parameters. The higher b is, the more flexible the mapping function is, but, on the other hand, the higher the sensitivity for noise will be. The unknown parameters, vector B, can be estimated by minimizing a non-linear least squares cost function

$$\hat{B} = \arg\min_{B} K_{\text{nls}} \tag{7-4}$$

with

$$K_{\text{nls}} = \frac{1}{2} \sum_{n=1}^{N} \left[s(n) - \left(A_0 + \sum_{k=1}^{h} A_k \sin(k\omega t_n) + A_{k+h} \cos(k\omega t_n) \right) \right]^2 \tag{7-5}$$

Figure 11 illustrates Martinson *et al.*'s method:

1. Two observations need to be dated (e.g. the first and last one, shown by the 'o's). This is necessary to implement a scale for the time grid (is one period 100 kyear or 1 year?).

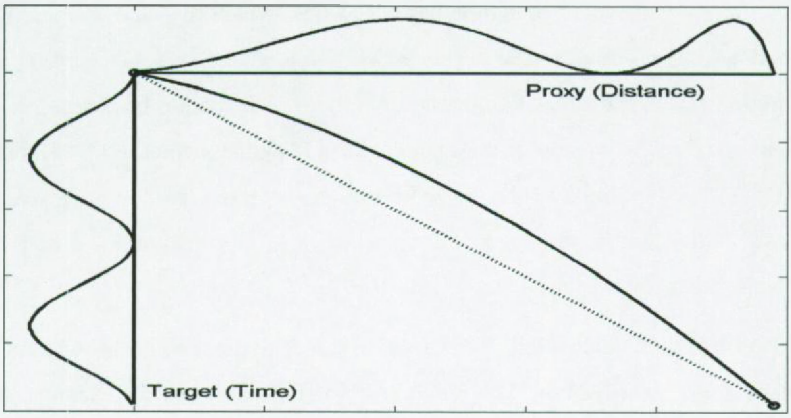

Figure 11: Visualization of Martinson et al.'s method. Horizontally the proxy is shown as function of a distance grid; Vertically the target function is shown as function of time. The accretion rate is expanded in a set of basis functions with unknown coefficients. These coefficients are estimated employing a least squares estimator. Starting from a constant accretion rate (dotted line), the parametrized accretion rate can be estimated (full line). Additionally two observations need to be dated (e.g. 'o's in the first and last observation).

2. An initial estimate of the time base (mapping function) is needed to initialize the parameters: e.g. the dotted line on Figure 11.
3. A least squares cost function is minimized to optimize the parameters. The corresponding accretion rate is shown by the full line on Figure 11.

7.3.4.2 Shortcomings and improvements

1. Which initial values should be used for the unknown coefficients, B? This problem is complicated by local minima in cost function surface [Martinson *et al.*, 1982a]. Here is proposed to filter the proxy and target function with a low pass filter in order to match the low frequency components first. In this manner local minima are first suppressed by the low-pass filter and next by-passed by improved initial values. In [Martinson *et al.*, 1987] Martinson *et al.* propose an alternative by using an independent radiometric chronology to generate initial values[1].

   We have used a different method to estimate the mapping function. Martinson *et al.*

estimated the time base in an absolute sense, while we start form a average constant accretion rate, disturbed by a time base distortion, which can be expanded in a set of basis function (e.g. Fourier series or Legendre polynomials). Because this average constant accretion rate is not estimated, the absolute time scale cannot be estimated. Consequently, the age of the proxy record is no longer estimated by the method. The advantage of such an adaptation is that initial values for these time base parameters can be generated by e.g. the phase demodulation technique. In principal, the anchor point method or correlation maximization could be used too.

2. How many coefficients should be chosen? The second problem is encountered empirically in [Martinson *et al.*, 1982a]: "The number of coefficients could for example be equal to the number of periods in the proxy". No considerations were used to estimate the number of parameters. We will use this rule to test this method.

3. A problem not handled is the Gibbs phenomenon: when the accretion rate is modeled by a Fourier series, it is implicitly assumed that the mapping function is continuous at the borders. This is in general not true. Therefore, an error is made. This problem was mentioned in chapter 4.

4. Because this is a parametric approach, it is straightforward to calculate uncertainty bounds on the estimated parameters

$$C_B = \sigma^2 (J^T J)^{-1} \text{ with } J = \left(\frac{\partial s}{\partial B}\right)_{\hat{B}} \tag{7-6}$$

where C_B the co-variance matrix on the estimated parameters, J the Jacobian, T transpose and σ^2 the variance on the measurements. If the latter are unknown, it can be estimated from the residual cost function ($\sigma^2 = 2K(\hat{\theta})/(N - n_B)$, with N the number of observations and n_B the number of parameters). The uncertainty on the time instances can be estimated by

1. However, these radiometric observations can only date the record if a corresponding target function is used. If, in addition, a parametric mapping function is used, the problem is moved to the new data set. The best solution for this new problem is to use two signal models, i.e. one for the radiometric data and one for the original signal, and one time base model. This way, the influence of stochastic noise is reduced.

$$C_{\text{time}} = (JC_B J^T)^{-1} \text{ with } J = \left(\frac{\partial g}{\partial B}\right)_{\dot{B}} \tag{7-7}$$

and on the accretion rate by

$$\sigma^2_{\text{accretion rate}} = \frac{D^2_s}{(t_n - t_{n-1})^4}\sigma^2 t_n + \frac{D^2_s}{(t_n - t_{n-1})^4}\sigma^2 t_{n-1}. \tag{7-8}$$

7.3.4.3 Basic assumptions made by Martinson *et al.*'s method

1. The target function is known;
2. the time base can be modeled by a limited number of 'simple' functions;
3. the relation between the proxy and the target function is static and linear[1].

Tuning parameters: the number of coefficients used to parametrize the accretion rate and the stop criterion in the optimization algorithm.

7.3.4.4 Limitations

The initial guess for starting values can limit the applicability of this method. In addition, the target function must be known, which is often not straightforward. If the target function is a periodic function, initial values for the time base can be gathered by e.g. the phase demodulation method. If other target functions are used, no general procedure seems to be available.

7.3.4.5 Stochastic noise

The simulation described in Figure 3 is used. The number of parameters used to reconstruct the mapping function and growth rate was eight, resulting in a RMS error of 0.9%. Compared with the method of Yu and Ding (35%) this value is low, due to the parametrization of the problem, while it is comparable with the value of Lisiecki and

1. A generalization to linear dynamical models is made by Brüggeman [1992].

Lisiecki (1%). The difference is due to the numerical precision. If Martinson *et al.*'s method would be implemented with a model selection criterion to estimate the desired number of parameters, this error would be even lower, i.e close to zero.

A Monte-Carlo simulation has been run to estimate the uncertainty on the time base parameters, B. This is compared with the uncertainty estimated in equation (7-6).

The results are shown in Table 1. In the case of stochastic noise, the estimated uncertainties

Table 1: values and uncertainties for the time base parameters of Martinson *et al.*'s method. (*) the true parameter values are not valid for this model error.

	Estimation of	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
'true' value		0	0	0	0	0	0	0	0
Stochastic noise	**Value**	0.04	0.00	0.02	0.01	-0.02	-0.01	0.00	0.00
	Uncertainty	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
	M.C. uncertainty	0.019	0.019	0.019	0.021	0.019	0.018	0.018	0.020
Periodic	**Value**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Uncertainty	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Amplitude modulated	**Value**	-0.02	-0.03	-0.03	-0.02	-0.01	-0.01	0.00	0.00
	Uncertainty	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
Trend	**Value**	0.22	0.35	0.59	-0.33	0.24	0.09	-0.53	-0.43
	Uncertainty	0.38	0.36	0.37	0.38	0.37	0.38	0.36	0.36
Hiatus (*)	**Value**	2.66	0.43	0.30	-0.98	0.48	-2.05	-1.04	-0.28
	Uncertainty	0.25	0.26	0.25	0.24	0.24	0.25	0.25	0.25

are slightly smaller than those estimated by a Monte-Carlo simulation. This can be due to the linearization. The estimated parameter values are much smaller than the corresponding uncertainties, which is also an indication that these are not significant. Figure 12 shows the estimated accretion rate and its 68% uncertainty bound. Again, none of the variations seems

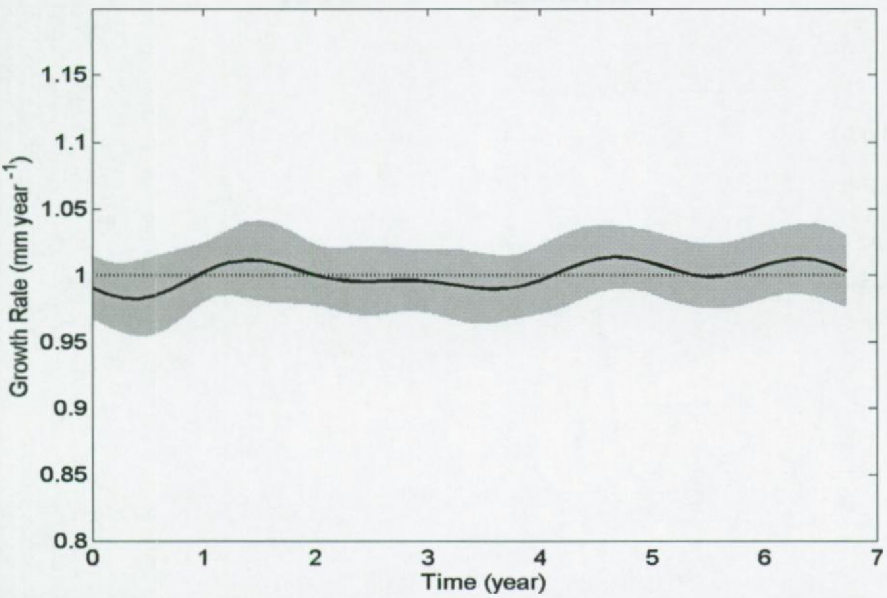

Figure 12: Estimated accretion rate by the Martinson *et al.*'s method in the presence of stochastic noise: the horizontal dotted line is the 'true' accretion rate, the full line is the estimated one. The grey band is the 95% uncertainty bound estimated with a Monte-Carlo simulation based on 100 runs.

to be significant, so if a automated model selection criterion had been used, these variations would have been rejected.

7.3.4.6 Model errors

Tested on the simulation examples, shown in Figure 5, Martinson *et al.*'s method performs well: the RMS error of the accretion rate is 0% when the proxy record consists of two harmonics, even if the target function is only sinusoidal. The estimated accretion rate is shown in Figure 13A. The same conclusion can be drawn for the amplitude modulated signal (RMS error of 1% - Figure 13B). On the trended signal, Martinson *et al.*'s method performs worse (RMS error of 53% - Figure 13C). On an incomplete record, Martinson *et al.*'s method is able to reconstruct the hiatus (RMS error of 8%). The reason for this error is the discontinuity in the growth rate caused by the hiatus, and consequently in the mapping function, $g(n)$, which can then hardly be described by a small number of Fourier

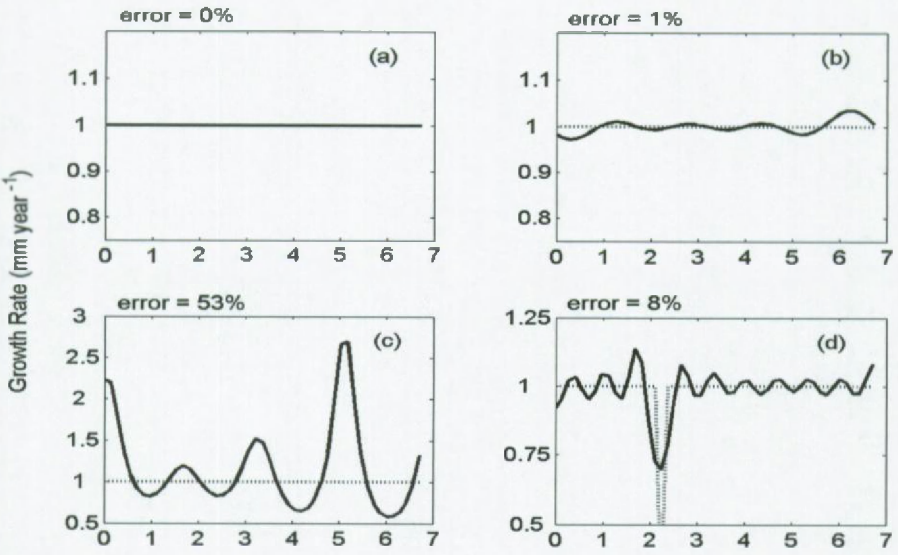

Figure 13: Estimated accretion rate by the Martinson *et al.*'s method in the presence of model errors: dash-dotted line is the true accretion rate and the full line is the estimated accretion rate: (a) a periodic signal instead of a sinusoidal signal, (b) an amplitude modulated signal, (c) sinusoidal signal + a linear trend and (d) hiatus.

coefficients. In order to overcome this problem, splines could be used as basis functions. These 'discontinuous' functions are very well suited to describe the time base in the presence of an hiatus.

When the uncertainty bounds are calculated in the presence of model errors, no clear relation can be found between the mismatch and the corresponding uncertainty bound. A possible explanation is that the uncertainty itself is estimated from the residual cost function, which expresses the mismatch between the signal model and the target function. The latter mismatch is not directly related to the accretion rate. More precise estimates of the uncertainty bounds on the accretion rate can be gathered if the uncertainties on individual observations are measured. some parameter values are significant compared to their uncertainty (Table 1). This significance is obviously a bias due to the model error.

7.3.4.7 Conclusion

No reliable method initial values, nor a model selection criterion is proposed by Martinson *et al.* [1982a] and most importantly, as for all orbital tuning methods the target function must be known. On the other hand, if the target function is similar to the proxy, this method, with some minor adaptations, performs well.

7.4 Class 2: Signal model methods

If the time series of the measured proxy is not similar to the target function, model errors are introduced, when one uses one of the previous methods. Each particular orbital tuning method can estimate a largely biased time base for particular types of model errors. Class 2 consists of signal models, which will lean on an assumed property of the unknown target function, like the target function is sinusoidal, periodic, exponential,... Exploiting this property, the time series and accretion rate can be estimated. The investigator still has to make certain assumptions about the target function to estimate a time base. But, in contrast with orbital tuning methods, an ensemble of target function is proposed, instead of one single function. Such an ensemble can consist of one target function with unknown parameters or of a group of target functions, with varying complexity and with unknown parameters. This makes signal models more flexible than orbital tuning methods. Two distinct non-parametric methods exists in literature for this class of methods. The first acts in the time domain, the second in the frequency domain. Finally a parametric method is discussed.

7.4.1 Time domain method developed by Wilkinson and Ivany

The first non-parametric approach, originally proposed by Wilkinson and Ivany [2002], assumes that the time series is sinusoidal.

7.4.1.1 Idea

A window moves over the proxy record by one observation per step. For each window a least squares estimator estimates the amplitude, A, angular frequency, ω, phase, ϕ, and position (offset or DC component), DC, by matching

$$S_{\mathrm{WI}}(m,n) = DC(m) + A(m)\sin[\omega(m)nT_s + \phi(m)] \tag{7-9}$$

in a least squares sense on the samples in the window (T_s is the sample period, $m \in [0, \ldots, N-1-c]$ the window number, with c the width of the window and

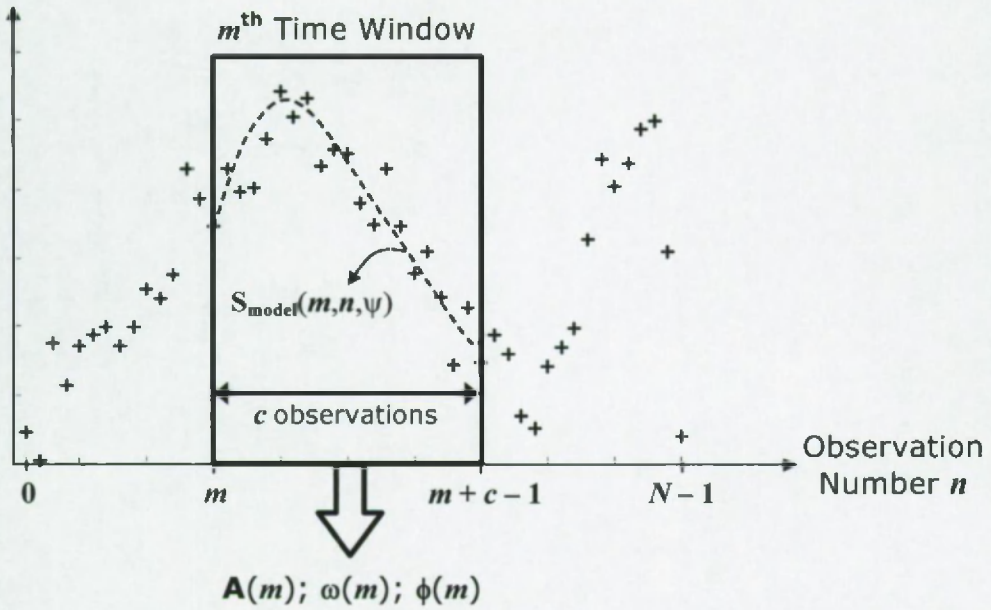

Figure 14: Illustration of the parameters used to estimate the amplitude, frequency and phase of Wilkinson and Ivany's method. The observations are shown by the '+', the dotted line is the model and the rectangular is the time window.

$n \in [m, \ldots, m+c-1]$ the sample number within the window (see Figure 14)). When the accretion rate was large this will be reflected in the frequency parameter, which would be low and vise versa. By moving the window over the observations, the frequency is estimated over the total length of the record and the varying accretion rate can be calculated.

7.4.1.2 Improvements

1. It is better to remove the offset, DC, before the remaining parameters are matched on each window. The offset estimated in each window will slightly differ from this average offset, due to the particular noise realization in that window. These differences will cause extra variations in the estimated amplitude and phase [Schoukens and Renneboog, 1984]. Therefore, we propose to estimate the offset on the total record. This has the advantage that the noise is averaged over much more observations. In addition, in a next step the amplitude must be fixed anyway (see equation 7-14), so an estimation of the variations in the offset is redundant.

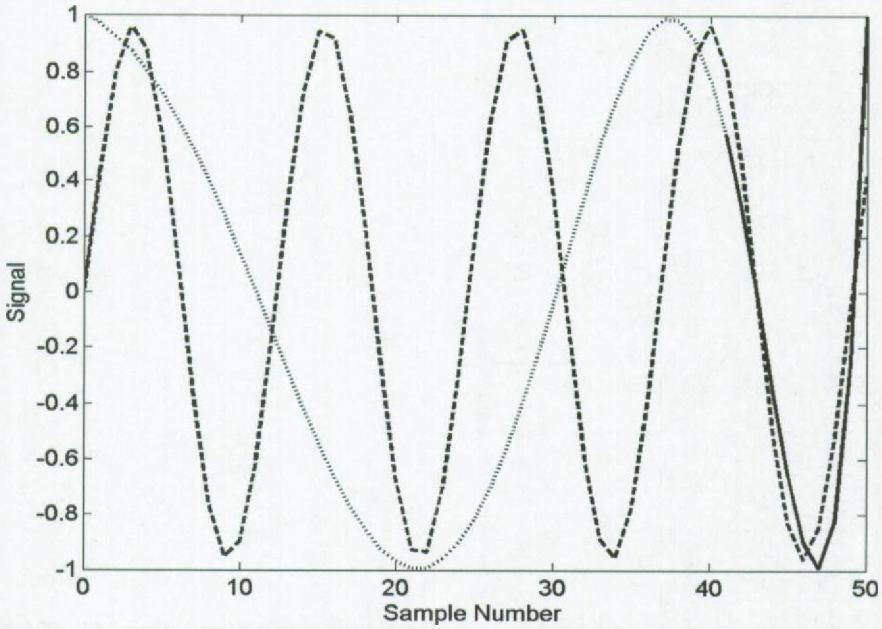

Figure 15: The dotted and the full line represent a proxy-signal. On the full line a sinusoidal signal is matched. This sinusoidal is shown by the dashed line.

2. The method proposed in [Wilkinson and Ivany, 2002] does not take into account that a part of the variations in accretion rate will appear in the phase modulation. This is illustrated in Figure 15. The proxy record is a cosine (a sines with phase $\pi/2$), shown by the dotted and the full line. The last window of ten samples is represented by the full line. On this window, a sinusoidal model (equation 7-9) is matched. This model is extrapolated over the total measurement window. At the origin, it can be seen that the phase is now zero. Indeed, a change in the time base does not only influence the frequency, but also the phase. So, if we want to reconstruct the time base, both the frequency and phase have to be taken into account. To reformulate these ideas, assume a sinusoidal signal, $s_{\sin}(n)$, described by

$$s_{\sin}(n) = A_0 \sin\{\omega_0[n + g(n)]T_s + \phi_0\} \tag{7-10}$$

with sample number $n \in [0, \ldots, N-1]$, sample period T_s, A_0, ω_0 and ϕ_0 respectively the constant amplitude, radial frequency and phase and the time base distortion $g(n)$. Note that we exclude frequency and phase modulations, i.e. ω_0 and ϕ_0 are no function of the sample

number, n. If this were the case, the time base would not be uniquely reconstructed. The measured observations,

$$s(n) = s_{\sin}(n) + e(n) \tag{7-11}$$

are modeled by equation (7-10); the mismatch between both, the observations and the model, is expressed by the error term $e(n)$. In order to estimate the time base distortion, $g(n)$, we move with a window of width c (expressed in number of observations) over the record in $m \in [0, \ldots, N-c-1]$ steps. For each window,

$$s_{\text{model}}(m, n, \psi(m)) = A(m)\sin[\omega(m)nT_s + \phi(m)] \tag{7-12}$$

is matched, i.e. the parameter vector $\psi(m) = [A(m), \omega(m), \phi(m)]$ is estimated ($n \in [m, \ldots, m+c-1]$ is the sample number within the window, $A(m)$, $\omega(m)$ and $\phi(m)$ are respectively the modulated amplitude, radial frequency and phase, which can change with the window number, m (see Figure 14). The window moves over the proxy record in steps of one observation and at each position, m, the parameters are estimated by

$$\hat{\psi}(m) = \arg\min_{\psi(m)} K_{\text{nls}}(m) \tag{7-13}$$

with

$$K_{\text{nls}}(m) = \frac{1}{2}\sum_{n=m}^{m+c-1}[s(n) - s_{\text{model}}(m, n, \psi)]^2 \tag{7-14}$$

In order to reconstruct the variation in accretion rate equation (7-10) is set equal to equation (7-12) and the resulting expression is solved for the time base distortion, $g(m)$. This is only possible if an additional assumption is made, i.e. the amplitude is not modulated

$$A(m) = A_0 \tag{7-15}$$

which leads to

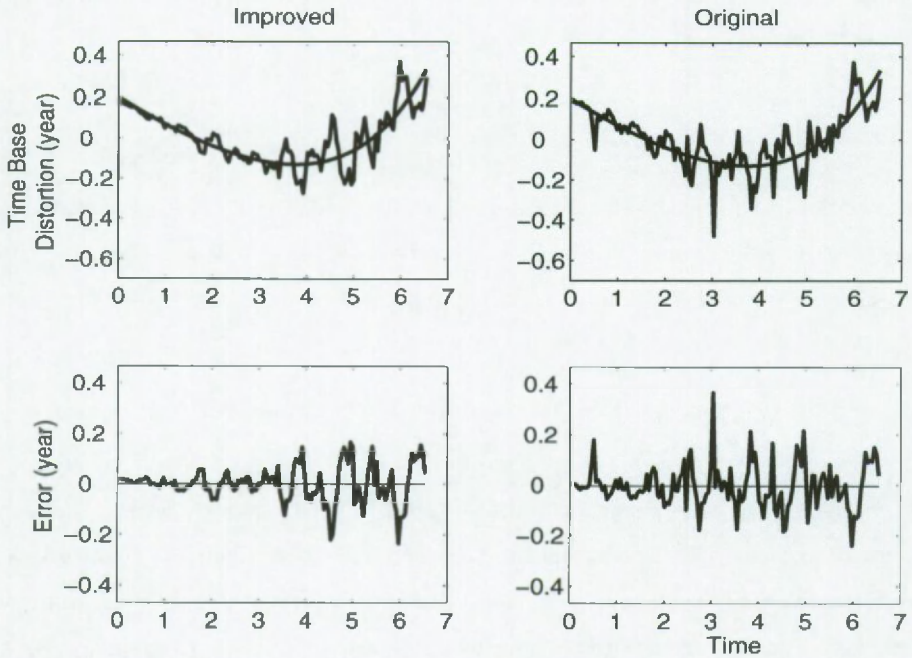

Figure 16: Estimated time base distortion with Wilkinson-Ivany method versus true time base distortion: left with the improved method and right with the original method.

$$g(m) = m\left(\frac{\omega(m)}{\omega_0} - 1\right) + \frac{\phi(m) - \phi_0}{\omega_0 T_s} \tag{7-16}$$

Note that $m \in [0, \ldots N-1-c]$, so the c last observations are not dated. In order to find a unique solution for the time base distortion, $g(m)$, two additional hypotheses are used:

1. The average time base distortion is zero, i.e. ϕ_0 can be tuned in order to fulfill this assumption. This assumption can be replaced by: the first observation is dated, which fixes ϕ_0 too. the latter would be the case if e.g. a bivalve is sampled and the most recent skeleton part is formed on the moment of sampling; and
2. No linear trend is present in the time base distortion, which fixes ω_0.

The improved results are illustrated on a simulation (See Figure 16), consisting of a sinusoidal signal (150 observations sampled over seven years), disturbed by stochastic noise (SNR of 2.6), no model errors, an exponentially decreasing accretion rate (as is reflected in

the time base distortion shown on Figure 16), and the width of the window is ten. Figure 16 shows further that the adapted estimation of the accretion rate decreases the RMS error from 8.4% for the original to 6.9% for the improved estimation. From other simulations, it followed that the difference between these errors increases when the window becomes smaller and/or when the SNR decreases.

7.4.1.3 Assumptions made by Wilkinson and Ivany

1. the time series of the proxy record is modeled by a sines within a window of width c;
2. the variations in accretion rate are not faster than the width of the window;
3. at least two observations are dated.

The width of the window. Enlarging the window will decrease the noise influence on the parameters, but at the same time it will lower the time resolution of the variations in accretion rate. Decreasing the window width will improve the time resolution of the accretion rate, but at the same time, the noise influence on the accretion rate will increase.

7.4.1.4 Stochastic noise

The simulation described in Figure 3 is used. The accretion rate was estimated with a window of ten observations and the resulting RMS error was 3% (see Figure 17). Note that the accretion rate could not be estimated for the last ten observations, because of the width of the window.

7.4.1.5 Model errors

Tested on the simulation examples, shown in Figure 5, Wilkinsons and Ivany's method performs quite poorly, as can be seen in Figure 18. The RMS error for the periodic signal is 94% (Figure 18a), for the amplitude modulated signal 16% (Figure 18b), for the trended signal 47% (Figure 18c) and for the hiatus 8% (Figure 18d). The window width was again ten samples. Decreasing this would lower the errors for the amplitude modulated signal, because the amplitude variation would be smaller in a smaller window, and for the hiatus, because less windows would cover the hiatus.

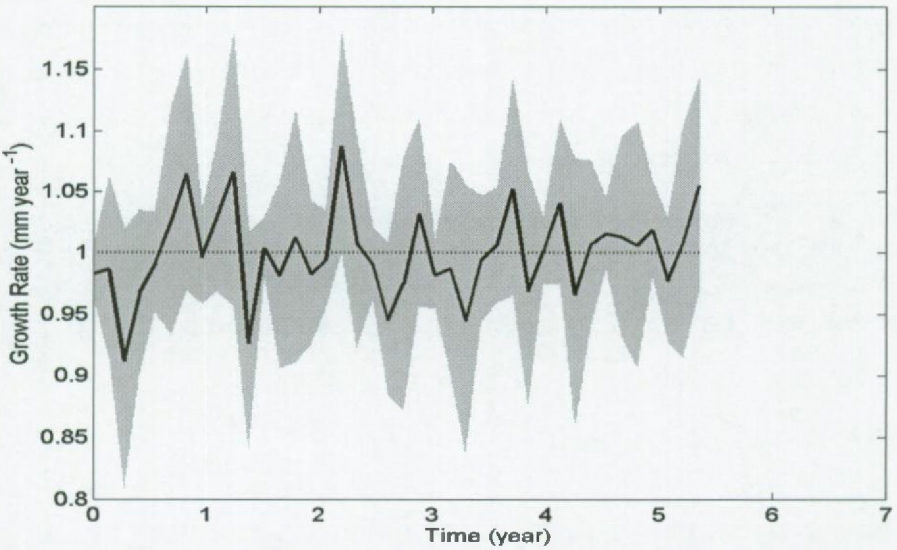

Figure 17: Estimated accretion rate by Wilkinson-Ivany's method in the presence of stochastic noise: the horizontal dotted line is the 'true' accretion rate, the full line is the estimated one. the grey band is the 95% uncertainty bound estimated with a Monte-Carlo simulation based on 100 runs.

7.4.1.6 Conclusion

The adapted Wilkinson-Ivany's method performs well in the presence of noise, but is not always satisfactory in the presence of model errors.

7.4.2 Frequency domain method: a phase demodulation approach

Before the phase demodulation approach is discussed, it is useful to argument why an approach in the frequency domain has advantages. Figure 19 shows the Fourier spectra of the different simulated records (Figure 5). In the time domain, both the target function and the model error appear at every time instance, which complicates their separation.

In the frequency domain the model errors in cases (a) (overtone) and (c) (trend) appear at different frequencies than the target function. For the model errors of cases (b) (amplitude

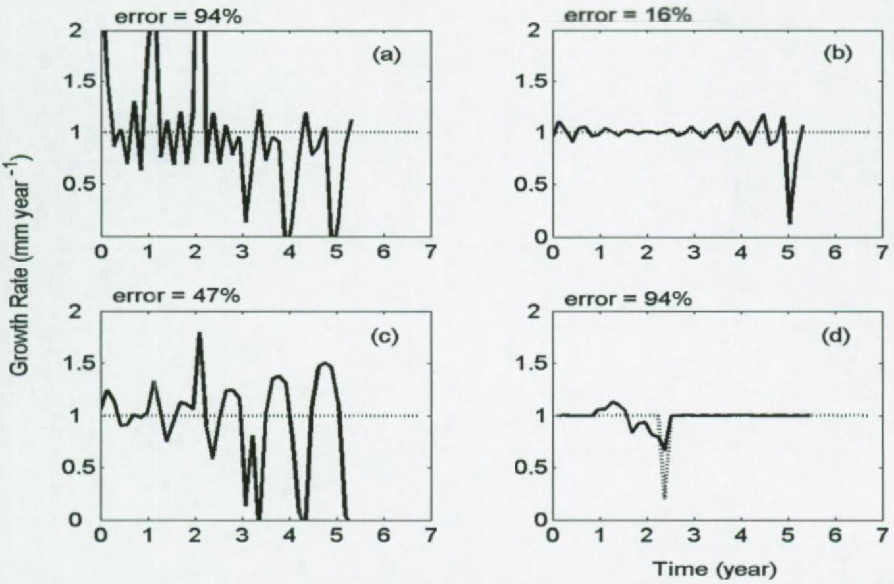

Figure 18: Estimated accretion rate by Wilkinson-Ivany's method in the presence of model errors: dotted line is the true accretion rate, the full line is the estimated accretion rate: (A) a periodic signal instead of a sinusoidal signal, (B) an amplitude modulated signal, (C) sinusoidal signal + a linear trend and (D) hiatus.

modulation) and (d) (hiatus), the model error appear around the harmonic and can thus still influence the time base. Note that

1. these spectra show only half of the record. The phase of the different frequencies is not shown;
2. Many noise properties are defined in the frequency domain, like white noise (power density is constant over a finite frequency range). So in order to identify the different components of a signal model, the spectral interpretation can be beneficial.

7.4.2.1 Idea

The phase demodulation method uses the frequency domain to identify the time base. As will be shown, the robustness against stochastic noise and model errors is larger than the similar method in the time domain. An extended discussion of this method can be found in chapter 3. It is assumed that the time series is periodic and band limited. Consequently, it can be modeled by a limited number of harmonics in the frequency domain. The variations in

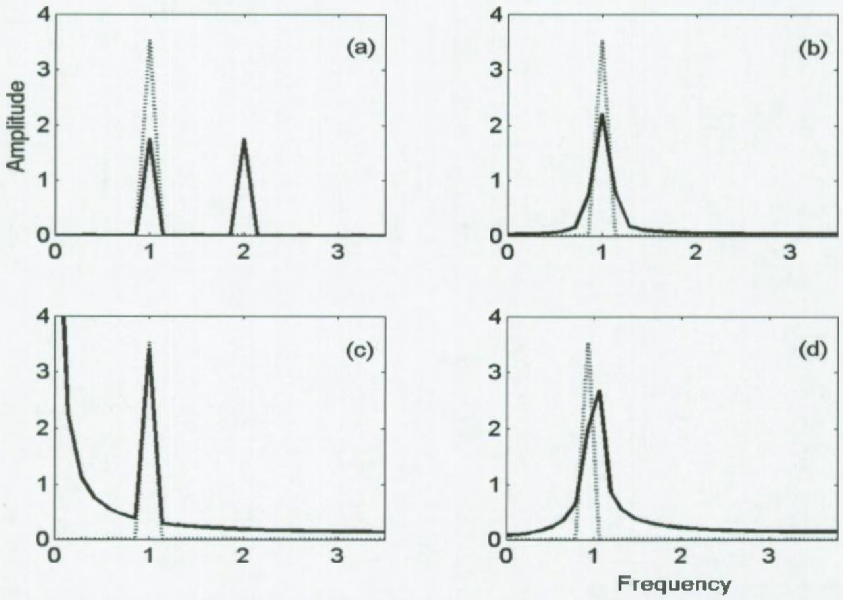

Figure 19: The spectra of the different proxy records (full line) and the assumed target functions (dotted line) are shown: (a) a periodic signal instead of a sinusoidal signal, (b) an amplitude modulated signal, (c) sinusoidal signal + a linear trend and (d) hiatus.

accretion rate will act as a phase modulation on these harmonics and a phase demodulation technique will reconstruct the accretion rate according to the following procedures

1. The spectrum, $S(\omega)$, of the proxy-record, $s(n)$, is calculated (DFT: Discrete Fourier Transform). This is only possible if we have an initial estimate of the time instances of the samples. Therefore, the time base is initialized employing an average constant growth rate

$$S(\omega) = \mathrm{DFT}[s(nT_s)] \tag{7-17}$$

2. The T.B.D. is isolated by windowing the spectrum, $S(\omega)$, with a rectangular window, $c(\omega)$, around the first harmonic. This cuts out an estimate, $\hat{G}(\omega)$, of the spectrum

$$G(\omega) = C(\omega)S(\omega) \tag{7-18}$$

with $C(\omega) = 1$ if ω is in the frequency range of the window and $C(\omega) = 0$ elsewhere.

3. The estimation of the time base distortion, $\hat{g}(n)$ (for a definition see equation (7-2)), is

completed by employing the Inverse Discrete Fourier Transformation (IDFT)

$$\hat{g}(n) = \frac{\varphi\{\exp[-j(\omega_0/\omega_s)2\pi n]\mathrm{IDFT}[G]\}}{\omega_0} \tag{7-19}$$

with $\varphi(u)$ the phase of u, $j = \sqrt{-1}$, ω_0 the fundamental radial frequency (2π/age (e.g. year^{-1})) and ω_s the radial sampling frequency $2\pi/T_s$, with T_s the sample period (average time gap between two subsequent observations). The exponential term shifts the frequency by $-\omega_0$.

To summarize, transforming the time signal to the frequency domain, enables reconstruction of the time base distortion, $g(n)$, by a relatively easy calculation. If the time base distortion is known, employing equation (7-19), every observation can be dated.

7.4.2.2 Assumptions made by the phase demodulation method are

1. proxy time series is periodic and band limited;
2. accretion rate is band limited;
3. at least two observations are dated, in order to have a time scale or an age of the record.

The width of the frequency domain window, c. Enlarging the window's width will improve the time resolution of the accretion rate, but at the same time the noise influence will be larger. Decreasing the window width will decrease the resolution, but at the same time the noise influence decreases. These properties are opposite to those in the time domain.

7.4.2.3 Stochastic noise

The simulation example, shown in Figure 3, is used. The accretion rate was estimated with a window of 1 year^{-1} (7 DFT lines). This would be the maximum allowed when a second harmonic is present in the signal, because decreasing the window size will even lower the noise influence, while increasing the window size will mis-interpret the time base distortion around the second harmonic (which is absent in this simulation). An RMS error of 0.5% is found on the estimated accretion rate (see Figure 20). This low error results from the spectral

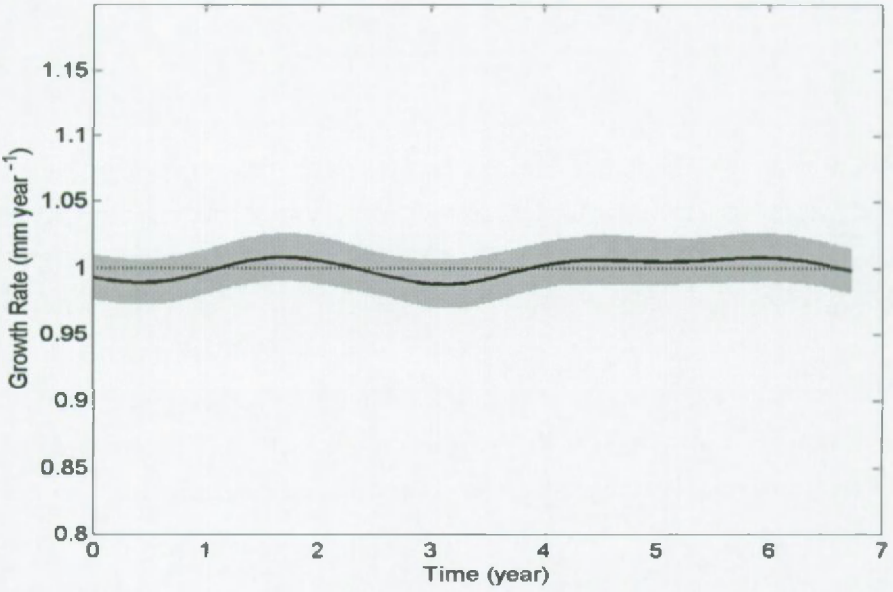

Figure 20: Estimated accretion rate by the phase demodulation method in the presence of stochastic noise: the horizontal dotted line is the 'true' accretion rate, the full line is the estimated one. The grey band is the 95% uncertainty bound estimated with a Monte-Carlo simulation base on 100 runs.

interpretation: the white noise is spread over all frequencies. Only seven lines are used to separate the time base distortion from this noise, while the noise on all other lines is not used. In addition, it is known that the time base distortion is symmetrical around the first harmonic. Therefore, the noise on both sides can be averaged out.

7.4.2.4 Model errors

Tested on the simulation example, shown in Figure 5, the phase demodulation method works well, because it was not assumed that the signal was sinusoidal. Figure 21 shows the estimated accretion rates (again a spectral window of seven lines was used to separate the time base distortion): the second harmonic was not covered by the window used to cut out an estimate of the time base distortion (T.B.D.). Consequently, no errors are made (Figure 21a). The same is true for the amplitude modulated signal (Figure 21b). Most of the trend appeared at the low frequency components of the spectrum and did not overlap with the window

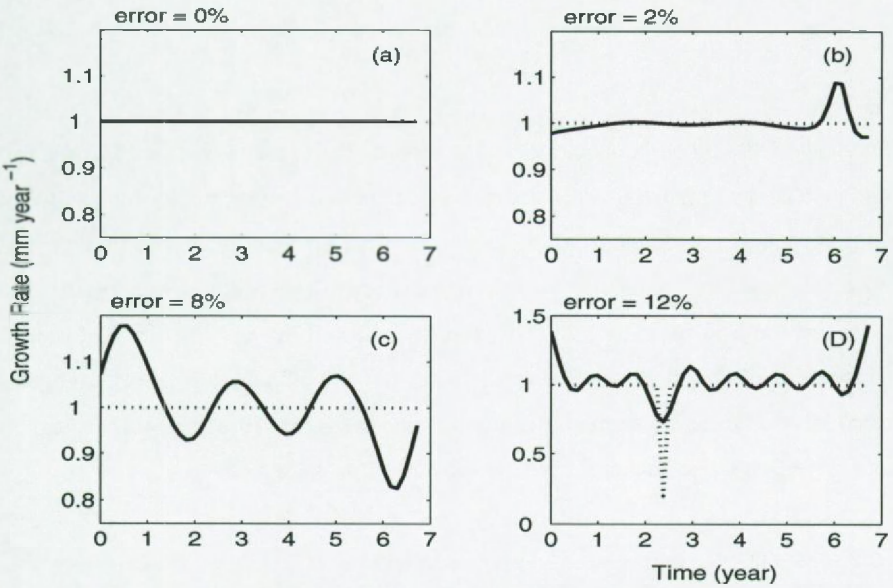

Figure 21: Estimated accretion rate by phase demodulation method in the presence of model errors: dotted line is the true accretion rate, the full line is the estimated accretion rate: (a) a periodic signal instead of a sinusoidal signal, (b) an amplitude modulated signal, (c) sinusoidal signal + a linear trend and (d) hiatus.

(Figure 21c). The hiatus produced the largest error because this discontinuity is spread out in the frequency domain (RMS error of 12% - Figure 21d), i.e. it is not bandwidth limited (Figure 19d), and is consequently only partly covered by the window.

7.4.2.5 Conclusion

The phase demodulation technique performed well in the presence of stochastic noise and in the presence of the model errors tested here. However, this is partly due to the subjective choice of using a window of seven lines to separate the time base distortion. One could have chosen a different window and could draw different conclusions from the same data. Therefore, a parametric model, combined with an automated model selection criterion will finally be discussed.

7.4.3 Parametric time base distortion approach

7.4.3.1 Idea

This method, described in chapter 6, is a combination of Martinson *et al.*'s method for the estimation of the time base, combined with a parametric representation of the signal, which will make it less vulnerable to model errors. In the framework of this discussion, we have limited the model complexity to periodic signal models. This is a parametric method, whose parameters need to be initialized. This can be achieved by, e.g., the phase demodulation method. In addition, this method employs an automated model selection criterion, which is able to estimate the complexity of the model from the residual cost function. Such automated model selection procedures makes the whole more objective.

It is assumed that the time series is described by equation (7-1). In contrast with Martinson *et al.*'s method, the amplitudes, A, frequency, ω, and number of harmonics, h have to be estimated and are not chosen by the operator. The time variable, t, is unknown, due to the variations in accretion rate, and can be decomposed into a constant time step and a distortion of this constant step, as in equation (7-2). This distortion of the time base can be developed into a Fourier series (equation (7-3)). The time base distortion, $g(n)$, depends on the coefficients, B, and on the number of coefficients, b, which should also be estimated.

Firstly, the investigator has to set the most complex signal and time base model, which are defined by the numbers $h_{\max}$ and $b_{\max}$. The number of parameters of this most complex model is $2h_{\max} + b_{\max} + 1$ and this number cannot be higher than the number of observations, otherwise the system is ill-conditioned. Most probably, the data profile is not that complex and the investigator may decide to limit the most complex model to a smaller complexity.

Secondly, all models with a lower complexity are optimized individually, e.g. the parameters of a model with complexity (h, b) are grouped in a vector,

$$\theta = [A_0, A_1, \ldots, A_h, B_1, \ldots, B_b, \omega]. \tag{7-20}$$

The optimal values in a least square sense are found by

$$\hat{\theta} = \arg\min_{\theta} K_{nls} \tag{7-21}$$

with K_{nls} defined in equation (7-5). Uncertainty bounds on the estimated parameters can be estimated by

$$C_\theta = \sigma^2 (J^T J)^{-1} \text{ with } J = \left(\frac{\partial s}{\partial \theta}\right)_{\hat{\theta}} \tag{7-22}$$

where C_θ is the co-variance matrix on the parameters (with $\theta = [\omega, A, B]$, where ω and A are defined in equation (7-1) and B in equation (7-3)), J the Jacobian matrix and σ^2 the variance on the measurements. The latter can be estimated from the residual cost function

$$\sigma^2 = 2K(\hat{\theta})/(N - n_\theta) \tag{7-23}$$

with N the number of observations and n_θ the number of parameters. The uncertainty on the time instances can be estimated by equation (7-7) and on the accretion rate by equation (7-8).

Thirdly, the optimal complexity, i.e. h and b, has to be estimated. This is achieved by assigning a value to each model which expresses how suitable it is to describe the data. this value cannot simply be a measure of the fit, since increasing the model complexity will decrease the systematic errors. However, at the same time the model variability increases. Hence, it is not a good idea to select the model with the smallest cost function within the set because it will continue to decrease when more parameters are added. At a certain complexity the additional parameters no longer reduce the systematic errors but are used to follow the actual noise realization on the data. As the noise varies from measurement to measurement, the additional parameters increase only the model variability. To avoid this unwanted behavior for the model selection criterion the cost function is extended with a model complexity term that compensates for the increasing model variability. Summarized, the model selection criterion should be able to detect undermodeling (= too simple model) as well as overmodeling (= too complex model). This model complexity term is dependent upon the signal-to-noise ratio and the availability of a noise model. In the examples worked out in this paper the criterion to be minimized had the following expression

$$\frac{K(\hat{\theta})}{N} e^{p_c(n_\theta, N)} \text{ with penalty } p_c(n_\theta, N) = \frac{\ln(N)(n_\theta + 1)}{N - n_\theta - 2} \tag{7-24}$$

with $K(\hat{\theta})$ the residual cost function, N the total number of observations and n_θ the number of parameters. Notice that introducing a model selection criterion eliminated all interferences from the user, which makes the proposed method objective and user independent. Practically, this means that only a model set has to be defined, i.e. the maximum values for h and b must be set. Next, equation (7-24) can be used to select the best model within this set. A detailed description of the model selection criteria can be found in chapter 5.

7.4.3.2 Assumptions

1. The signal is periodic and band limited;
2. At least two observations are dated.
3. If the uncertainty on the individual observations is not known, it is assumed that the noise is white, i.e. independent and identically distributed over all observations.

The first assumption can be replaced by another property of the signal, like a periodic signal with trend, or even a polynomial signal, multiple non-harmonically related periodicities, etc.... the main limitation is that the phase modulation method cannot in all cases be used to generate initial values. So, other methods must be able to generate sufficiently good initial values.

Besides the stop criterion in the optimization algorithm, no tuning parameters are used. The complexity of the time base distortion model and of the signal model are set by the user in order to limit the computation time, but with enough calculation time the maximal number of parameters could be chosen equal to the number of observations, so that the cost function would be zero. It would not be possible to estimate more parameters and the model selection criterion will select the best model from this general model set.

7.4.3.3 Stochastic noise

The simulation example, shown in Figure 3, is used. The values and uncertainties of the signal parameters are summarized in Table 2. Assuming white noise, the method proposed above is able to detect that no variations in accretion rate occurred, i.e. the model selection criterion rejected all time base distortion parameters. Consequently, the RMS error and the uncertainty on the time base were zero (see Figure 22).

Table 2: estimated values and uncertainties of the parametric approach tested in the presence of stochastic noise.

	ω (rad/year)	A_0	A_1	A_2
'true' value	6.28	0	1	0
Estimated value	6.27	0.01	0.99	0.02
$\hat{\sigma}$	0.01	0.01	0.03	0.03
$\hat{\sigma}$(MC)	0.05	0.01	0.02	0.02

7.4.3.4 Model errors

Tested on the simulation examples, shown in Figure 5, the parametric time base distortion approach performed well. White noise was assumed. For the periodic signal (Figure 23a), the amplitude modulated signal (Figure 23b) and the trended signal (Figure 23c) the combination of a parametric signal and time base distortion model, $g(n)$, with a model selection criterion was able to reconstruct the correct accretion rate. Only for the hiatus this parametric time base distortion approach produced a quite large error (RMS error of 22% - Figure 23d). The reason is the same as for the phase demodulation method: the hiatus is a discontinuity in the time domain, which is spread over all frequencies in the frequency domain. The model selection criterion rejects the more complex models, which results in a simple time base distortion model with a large error. This error can, at least partly, be avoided if a noise model is available, i.e. the individual uncertainties on the observations are known, where the model selection criterion does not penalize more complex models so strongly. An

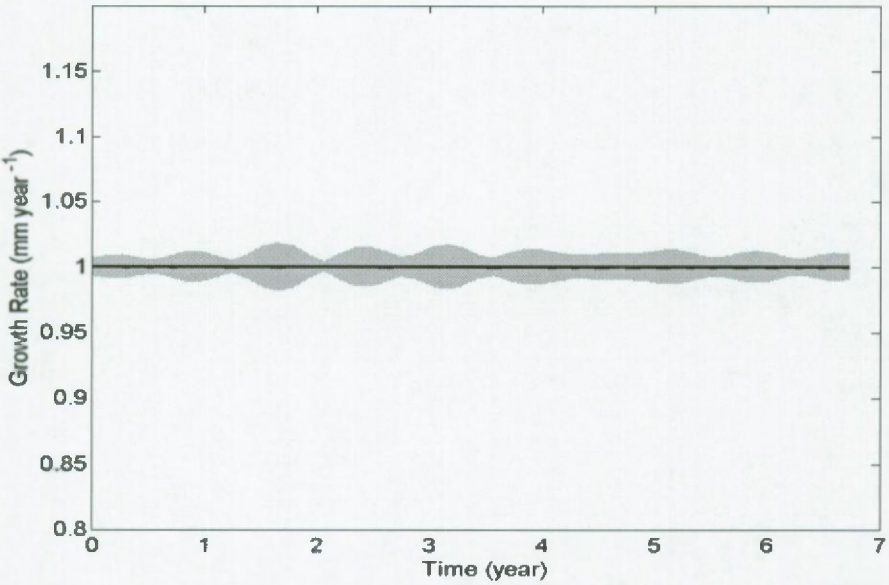

Figure 22: Estimated accretion rate by the parametric time base distortion method in the presence of stochastic noise: the horizontal dotted line is the 'true' accretion rate, the full line is the estimated one (both overlap). The grey band is the 95% uncertainty bound estimated with a Monte-Carlo simulation base on 100 runs.

alternative solution would use B-splines, which are able to describe such hiatuses with a small number of parameters.

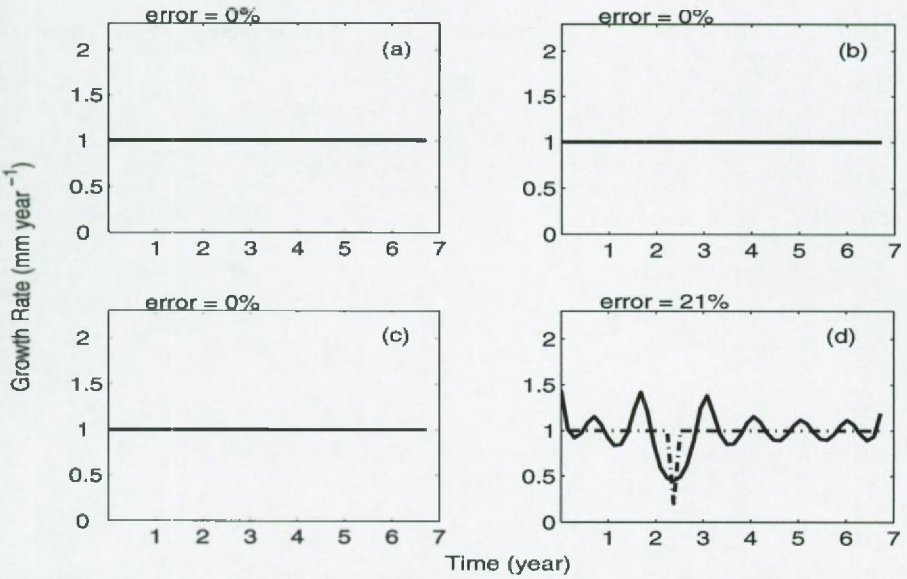

Figure 23: Estimated accretion rate by parametric time base distortion approach in the presence of model errors: dash-dotted line is the true accretion rate, the full line is the estimated accretion rate and the dotted line its estimated uncertainty: (a) a periodic signal instead of a sinusoidal signal, (b) an amplitude modulated signal, (c) sinusoidal signal + a linear trend and (d) hiatus.

The estimated parameters and their uncertainties are given in Table 3.For the first model

Table 3: Overview of the estimated parameters and their uncertainties in the presence of model errors.

Model error	Parameter	Estimated value	Estimated uncertainty
Amplitude modulated	ω (rad/year)	6.25	3.54
	A_0	0.0085	23.13
	A_1	0.4427	5.76
	A_2	0.4467	5.68
Trend	ω (rad/year)	6.20	0.34
	A_0	1.01	0.14
	A_1	0.93	0.65
	A_2	0.53	1.12
Hiatus	ω (rad/year)	5.97	0.08
	A_0	0.00	0.05
	A_1	0.01	0.27
	A_2	0.90	0.06
	B_1	0.76	0.13
	B_2	-0.52	0.12
	B_3	0.91	0.23
	B_4	0.18	0.16

error, i.e. the periodic signal, no model error was present for this approach and the uncertainties on the parameters are consequently zero and this model error is consequently not included in Table 3. All time base distortion parameters were rejected for the periodic, the amplitude modulated and the trended model error. Consequently, these time base

distortion parameters are not included in Table 3. The mismatch for the growth rate in the case of a hiatus is 21%, which is not negligible. An explanation for this is that many parameters are needed to model a hiatus (discontinue function), with a Fourier series. On the other hand, the automated model selection criterion will penalize for this increasing complexity. On the balance, the indicated, intermediate complex model is selected. A possible solution could be the use of splines, which are B-splines, able to describe this type of model errors with only a limited number of parameters.

The estimated uncertainty on the signal parameters for all model errors can sometimes be large. This is because the variance is estimated from the residual cost function. All variation, which is not covered by the model is supposed to be noise. In the presence of model errors this noise can become severe. This problem can be circumvented by measuring the uncertainties, instead of estimating them. The latter would also change the automated model selection criterion, so that more complex models would be selected, which would decrease the error on the accretion rate in the case of an hiatus.

7.4.3.5 Conclusion

The parametric time base distortion approach performs best, compared with other approaches in the presence of stochastic noise as well as in the presence of most model errors tested here. This is partly a result of the model selection criterion, used to tune the complexity and thus to filter the stochastic noise.

7.5 Class 3: Hard model method

In some particular and well defined cases, biological, chemical and physical prior knowledge can be used to predict the accretion rate. The properties of two of these models are discussed, but a complete analysis is not performed. One such example is given for the stalagmite growth by Kaufmann [2003]: "In some special cases prior knowledge about the accretion rate can be taken into account. For example the growth of stalagmites can be approximated by a simple mathematical model, which depends on growth rate and equilibrium radius. These two parameters are controlled by the climate. Temperature variations derived from ice and deep-sea core data, together with models for changes in precipitation and soil cover, are used to derive stalagmite stratigraphies, which reflect the palaeo-climate variations imposed." These types of hard models are not discussed here.

7.6 Comparison of the methods

The methods described above are compared on three criteria:

1. Influence of stochastic noise;
2. Use of tuning parameters; and
3. Sensitivity to model errors and assumptions used.

7.6.1 Influence of stochastic noise

The most important property of an accretion rate reconstruction method is that it must be relatively insensitive to stochastic noise, because every record is disturbed by it. If this noise is propagated into the reconstructed time base, the time series becomes unreliable. In general, non-parametric methods are more sensitive to stochastic noise than parametric methods, as is illustrated in Table 4. The best performing methods are the phase demodulation, Martinson *et al.*'s method and the parametric time base distortion approach.

Table 4: Overview listing the compared methods and answering the questions: What is the noise sensitivity (estimated from the example of Figure 3)? Is the method parametric? Does it use tuning parameters? Does it use the orbital tuning?

Method	R.M.S. (error)	Mean standard deviation	Parametric?	Tuning parameters?	Orbital tuning?
Anchor points	9%	4%	no	yes	yes
Yu and Ding	35%	987%	no	no	yes
Lisiecki and Lisiecki	1%	1.2%	yes	yes	yes
Wilkinson and Ivany	3%	3.5%	no	yes	no
Phase Demodulation	0.7%	0.8%	no	yes	no
Martinson *et al.*	0.9%	1.3%	yes	yes	yes
Parametric Time Base Distortion Approach	0%	0.5%	yes	no	no

Two of these three procedures are parametric models, which are able to filter, at least partly, the signal from the stochastic noise. The phase demodulation method is non-parametric but also in this method the noise is filtered, by an averaging effect of the time base distortion around the first harmonic.

7.6.2 Use of tuning parameters

Tuning parameters have to be defined by the investigator. Consequently, the same record, converted to a time series with the same method, is not uniquely defined. The choice of tuning parameters can differ depending upon the choice of the investigator and, therefore, it is better to eliminate tuning parameters whenever possible. Only the correlation maximization, proposed by Yu and Ding, and the parametric time base distortion approach do not use tuning parameters (Table 4). Actually for the parametric time base distortion all tuning parameters, i.e. h in equation (7-1) and b in equation (7-3), are automatically set within the algorithm using objective, statistical criteria.

7.6.3 Assumptions and model errors

It is impossible to transform a proxy-record into a time series without making additional assumptions. Depending upon the nature of the record, the assumptions made by one method may be better than those of an alternative method. For example, if the record does not show any pattern, the only method that can be used is probably the anchor point method, because all others somehow use the pattern of the time series. On the other hand, if the record seems to be, e.g. sinusoidal, orbital tuning methods can be used. In the real world, the investigator does not know which assumptions are correct and what is worse: the result of a violation of one or several assumptions is hard to predict. We have tried to test the methods on some possible violations of the assumptions. Table 5 summarizes the performances of the methods

Table 5: Overview listing the RMS errors of the methods with respect to evaluated model errors estimated from the simulation examples given in Figure 5.

Method	Periodic instead of sinusoidal	Amplitude modulated instead of sinusoidal	Trend + sinusoidal instead of sinusoidal	Hiatus
Anchor points	5%	8%	8%	13%
Yu and Ding	10^7%	206%	65%	0.35%
Lisiecki and Lisiecki	19%	1%	0%	16%
Wilkinson and Ivany	94%	16%	47%	8%
Phase Demodulation	0%	2%	8%	12%
Martinson *et al.*	0%	1%	53%	8%
Parametric time Base Distortion Approach	0%	0%	0%	22%

in the presence of the tested model errors. Note that the list of possible model errors is far from complete and that a method performing well on one particular model error, does not guarantee that it will work well on real world records. The general conclusions may be that

1. In general parametric methods seem to work better than the non-parametric ones;
2. Some the orbital tuning methods are sensitive to a mis-identification of the target function;

If the methods are assessed individually and mutually compared, the following remarks can be made

1. The anchor point method is quite robust against the tested model errors, but this is mainly due to the fact that in these examples the identification of the anchor points was not altered by the chosen model errors.
2. The anchor point method suffers from a discretization error, which means that even in the absence of model errors and disturbing noise, the anchor point method is not perfect.
3. Correlation maximization: the accretion rate is not limited by any expression, consequently it performs well in the absence of any errors, but in all other cases, these errors are propagated into the accretion rate.
4. The adapted Martinson *et al.*'s method is the best performing orbital tuning method. Only for

the signal with trend, it was not able to reconstruct the accretion rate.

5. Martinson *et al.*'s method is parametric, which inhibits the propagation of the stochastic noise and model errors into the accretion rate. Hereby it performs much better than the correlation maximization method.
6. The main disadvantage of the parametric methods, compared with the others is that it is possible that the order of the observations can be altered by these methods. However, this seems not to occur, as long as the number of time base distortion parameters remains small.
7. Martinson *et al.*'s method versus parametric time base distortion method: on the tested situations the latter performs better, except for the hiatus. Here, Martinson *et al.*'s method uses more time base distortion parameters. So the hiatus is better matched by Martinson *et al.*'s method.

7.6.4 Initial value problem

Parametric models, combined with local optimization methods will converge only if sufficiently good initial values are available (e.g. [Martinson *et al.*, 1982a] will hardly work because of a lack of a good initial guess for these values). On the other hand, the way these initial values are found is of no importance, so the non-parametric methods can be used to generate initial values. This has two advantages:

1. Take the correlation maximization as an example. In most cases this method seems to perform worse than the others, but if the main accretion rate variation is caused by an hiatus, this method performs best. However, the investigator cannot be sure in advance which method will perform best. If now, all methods are used to initialize the parameters for the parametric approach, the latter will select those initial values which lead to the lowest cost function. So, all other methods proposed can be used in the parametric approaches to generate initial values for the parameters and can thus help to avoid local minima.
2. So far, the phase demodulation method was used to initialize the parameters for the T.B.D. If other parametric models are used to describe the signal (e.g. a polynomial model), no initial values for the time base distortion are available, which makes this approach impractical. However, when combined with the other methods, used to initialize the T.B.D. parameters, these types of signal models can still be optimized.

7.7 Conclusion

To summarize, it is impossible to build a time base out of the blue. Instead, the investigator has to make some assumptions about the time series. Depending upon these assumptions the constructed time series and accretion rate evolution will be more or less sensitive to stochastic and model errors. Several methods proposed to construct time bases for proxy records have been tested and compared for their robustness in the presence of stochastic noise and in the presence of four types of model errors. From these tests we may conclude that non-parametric methods are more sensitive to stochastic disturbances. The tested models can be sub-divided into orbital tuning methods and methods with a signal model. The latter class is less sensitive to model errors. A last, but most important criterion, is the use of tuning parameters. If tuning parameters are used, different people can draw different conclusions from the same data. So, such subjective tuning parameters should be avoided if possible. From the set of methods we have tested here, the parametric time base distortion method has performed best. In this method, the signal model, the time base and the model complexity are tuned on the proxy record.

References

[1] Brüggeman W., 1992. A minimal cost function method for optimizing the age-depth relation of deep-sea sediment cores. Paleoceanography 7(4), 476-487.

[2] Ivany L. C., Wilkinson B. H. and Jones D. S., 2003. Using stable isotopic data to resolve rate and duration of growth throughout ontogeny: an example from the Surf Clam, *Spisula solidissima*. Palaios 18, 126-137.

[3] Jouzel J., Waelbroeck C., Malaize B., Bender M., Petit J.R., Stievenard M., Barkov N.I., Barnola J.M., King T., Kotlyakov V.M., Lipenkov V., Lorius C., Raynaud D., Ritz C., and Sowers T., 1996. Climatic interpretation of the recently extended Vostok ice records. Climate Dynamics 12(8), 513-521.

[4] Kaufmann G., 2003. Stalagmite growth and palaeo-climate: the numerical perspective. Earth and planetary science letters 214, 251-266.

[5] Lisiecki L. E. and Lisiecki P., 2002. Application of dynamical programming to the correlation of paleoclimate records. Paleoceanography 17(D4), 1049.

[6] Martinson D.G., Menke W. and Stoffa P., 1982a. An inverse approach to signal correlation.

Journal of geophysical research 87(B6), 4807-4818.

[7] Martinson D.G., Menke W. and Stoffa P., 1982b. An inverse approach to signal correlation (Reply). Journal of geophysical research 89(B4), 2501-2504.

[8] Martinson D. G., Pisias N. G., Hays J. D., Imbrie J., Moore T. C. and Shackleton N. J., 1987. Age dating and the Orbital Theory of the Ice Age: Development of a High Resolution 0 to 300,000-year Chronostratigraphy. Quaternary Research 27, 1-29.

[9] Paillard D., Labeyrie L. and Yiou P., 1996. Macintosh program performs time-series analysis, Eos Transactions AGU, 77(39), 379.

[10]Schoukens J. and Renneboog J., 1984. Spectrum analysis of signals with very low frequency components. Conference on precision electromagnetic measurements, August, 20-24, Delft, Netherlands, p. 1-4.

[11]Wilkinson B. H. and Ivany L. C., 2002. Paleoclimatic inference from stable isotope profiles of accretionary biogenic hardparts - a quantitative approach to the evaluation of incomplete data. Palaeography, palaeoclimatology, palaeoecology 185, 95-114.

[12]Yu Z. W. and Ding Z. I., 1998. An automatic orbital tuning method for paleoclimate records. Geophysical Research Letters 25, 4525-4528.

CHAPTER 8

NON-HARMONICALLY RELATED FREQUENCIES IN ENVIRONMENTAL ARCHIVES

Abstract - A mathematical inverse method is proposed, which estimates the spectral content of proxy signals, where the time base is not or hardly known. This approach can be used on signals consisting of multiple non-harmonically related periodic frequencies. The main idea is to expand the time base with a distortion term. We present a quantitative method for optimizing the signal parameters, while the time base distortion is estimated too. Once the time base distortion is known, an improved time base can be reconstructed. In opposition to existing methods, the signal and time base are not parametrized independently. Based on the residual cost function the optimal complexity of the time base and signal model are determined. Statistical properties of the algorithm are tested on simulations and next two real world example are processed.

8.1 Context

In this last chapter, we have tried to illustrate that the proposed approach can be expanded to more complex signal models. In addition, a possible critic is that we have used small, unknown records. In order to counter this criticism, the well-known Vostok ice core record is processed. To do so, the algorithms had to be extended to signal models with multiple non-harmonically related periodicities, so that the 100 kyear and the 40 kyear periodicity could be handled.

The data processed in the coral example were collected by Miriam Pfeiffer and Christian Dullo (Leibniz Institut fur Meereswissenschaften, Germany).

8.2 Introduction and situation of the problem

Often paleo-environmental archives are examined employing a spectral interpretation, like the Blackman-Tukey method [Harris, 1978], singular spectrum analysis [Ghil *et al*, 2002], wavelet analysis [Torrence and Compo, 1998], or interpretation of the auto correlation function [Broersen *et al.*, 2003], etc.... Indeed spectral analysis is an important tool in climate research, because it allows the variation of a time series to be separated into contributions associated with different time scales [Schulz and Stattegger, 1997; Schulz and Mudelsee, 2001]. However, all these interpretations are based on a known or a priori estimated time base. Errors in the time base can easily be mis-interpreted as e.g. narrow band noise [Schulz and Mudelsee, 2001]. In chapter 3, we have shown how this narrow band noise can be interpreted as a distortion of the time base and in chapter 6 a parametric approach for the class of periodic signals models was presented.

The relation between the signal record and time base is often ignored in time series analysis of geological records, because the investigator mostly constructs a time base first and interprets the time series afterwards. To link the measured distance signal with the time base, we introduce the concept of a time base distortion. For this, we start from an initially estimated time base. Such a base can be constructed, based on, e.g., growth bands in speleothems [Verheyden *et al.*, 2000; Finch *et al.*, 2001], corals [Sinclair *et al*, 1998; Fallon *et al*, 1999; Wei *et al.*, 2000; Marshall and McColloch, 2002; Kuhnert *et al*, 2002], on a physical ice accumulation model used to date ice cores [Jouzel *et al*, 1996], or, if no information is present, on the assumption of a constant accretion rate. It is possible that errors are made in such time bases. In order to correct these errors a distortion of the time base is allowed and in this paper is shown how this distortion can be estimated when one or multiple non-harmonically related periodic components are present in the signal.

A related problem is determining the optimal complexity of the signal model and of the time base distortion. The complexity is quantified by the number of parameters used in the respective model. The more parameters, the better the match with the measured record will be. However, at a certain complexity the additional parameters no longer reduce the systematic errors, but are used to follow the actual noise realization on the data. As noise

varies from measurement to measurement, the additional parameters increase only the model variability. In order to estimate the optimal model complexity the cost function is extended with a model complexity term that compensates for increasing model variability. This extension of the cost function is discussed in paragraph 8.5.

Two case studies are worked out: first a stable oxygen isotope signal in a coral is discussed (paragraph 8.7.8.7.1) and next the temperature record, reconstructed from the Vostok ice core record, is analyzed (paragraph 8.7.2). For both cases several problems are discussed:

1. The time instance of each observation is often doubtful. To overcome this problem a refined time base is proposed.
2. Which variation is significant? What is noise? And how can we separate both in an objective manner?
3. How many periodic signals models can significantly be separated from the stochastic noise?
4. One such periodicity consists of one or more harmonics. How many harmonics are significant within each periodicity?
5. What are the characteristics of the periodic components, like the numerical values of the frequencies, phases and amplitudes?
6. When a signal consists of multiple periodicities, investigators are often interested in one of these periodicities and want to separate it from the other periodicities and from the stochastic noise. The parametric approach presented here, consists of several building blocks, like the different periodic signals models and the noise. Once the model is matched, each of these blocks can be treated separately.

An example of the last problem is the separation of anomalies, like El Niño, from, e.g., the annual variations, in, e.g., a stable isotope record of corals or sponges. How this is done is illustrated in the first case study. Another example is the separation and identification of the

different variabilities due to solar input in ice cores [Petit *et al.*, 1999], which is discussed in the last case study.

This chapter is organized as follows. Firstly, the mathematical formulation of the signal and time base are elaborated. Secondly, the parameter estimation and, thirdly, the model complexity are explained and tested on a Monte-Carlo simulation. Finally, the two real world examples are discussed.

8.3 Problem formulation

The aim of this section is to formulate the mathematical model. The record under investigation is assumed to consist of the sum of non-harmonically related periodic signals models. Thus, the record under test will be modelled as function of time. The discrete time signal model, $s_{\text{model}}(t_n)$, is given by

$$s_{\text{model}}(t_n) = A_0 + \sum_{m=1}^{M} s_m(t_n) \tag{8-1}$$

where A_0 is the offset, M is the total number of non-harmonically related frequencies in the signal and n the observation number. One such periodic component, $s_m(t_n)$, is given by

$$s_m(t_n) = \sum_{k=1}^{h_m} A_{m,k}\cos(k\omega_m t_n) + A_{m,k+h_m}\sin(k\omega_m t_n) \tag{8-2}$$

where h_m is the number of harmonics in the periodic component m, the parameters A_m are stored in a vector of length $2h_m$, ω_m is the unknown angular frequency of the periodic component m and

$$t_n = (n + g(n))T_s \tag{8-3}$$

is the unknown time base with T_s the sample period. t_n represents the date of each observation and is decomposed in an average constant time step, i.e. nT_s (assuming a constant accretion rate), and a time base distortion, $g(n)T_s$, which expresses variations in accretion. In order to quantify these variations, the distortion $g(n)$ is expanded in a Fourier series

$$g(n) = \sum_{l=1}^{b/2} B_l \cos(2\pi n l) + B_{l+b/2} \sin(2\pi n l) \qquad (8\text{-}4)$$

where B is a vector of length b, containing the unknown time base coefficients. How many coefficients are needed can be chosen by the user, but an objective selection criterion is proposed in paragraph 8.5. Note that

1. any variation in growth or accretion rate can be expressed by equation (8-3);
2. polynomial models for the time base distortion were tested too, but in all cases tested so far, the residual cost function seems to be higher for these models than for Fourier models for the same number of parameters. Therefore, this paper is limited to Fourier modeled time base distortions.

The complexity of one model is given by the number of periodic signal models, M, the number of harmonics in each periodic component, h_m, and the number of time base distortion parameters, b. The complexity of the models which are optimized varied from no periodicity at all to $M_{\max}$ periodicities: $M \in \{0, \ldots, M_{\max}\}$. Each of these periodicities consists of one to a given maximum number of harmonics: $h_m \in \{1, \ldots, h_{m,\max}\}$. Also the number of parameters of the time base distortion varied from zero to a given maximum: $b \in \{0, \ldots, b_{\max}\}$. In theory, the maximum number of parameters can be equal to the number of observations, corresponding to a cost function equal to zero, but in practice, the complexity is usually limited in order to avoid unnecessarily long computation times.

8.4 Parameter estimation

In this section, first the general structure of the cost function approach used to estimate the model parameters is discussed (8.4.1). Next a strategy to gather initial values is proposed (8.4.2).

8.4.1 Cost function

The parameters to be estimated for one particular model are grouped in a vector, θ,

$$\theta = [A_0, \omega_1, \omega_2, \ldots \omega_m, A_1, A_2, \ldots, A_m, B] \tag{8-5}$$

with $A_i = [A_{i,1}, A_{i,2}, \ldots, A_{i,2h_i}]$, $B = [B_1, B_2, \ldots, B_b]$, $m \in \{1, \ldots, M\}$ and $i = \{1, \ldots, h_m\}$. The optimal value of the parameter vector, denoted by $\hat{\theta}$ is found by minimizing a least squares cost function

$$K(\theta) = \frac{1}{2} \sum_{n=1}^{N} [s(n) - s_{\text{model}}(t_n)]^2 \tag{8-6}$$

where $s(n)$ is a vector of N observations and $s_{\text{model}}(t_n)$ is defined in equation (8-1). Due to the nonlinearity of the model w.r.t. the unknown parameters, a numerical method must be used to find the cost function minimum. for the results presented in this chapter, a Levenberg-Marquardt optimization algorithm was used.

Special attention was given to the stop criterion of this algorithm. If, on the one hand, one is interested in the parameter values, the optimization can be stopped when the changes in the parameter values have dropped beneath a certain numerical value. If, on the other hand, one is interested in the optimal model complexity, the iteration should be stopped if the cost function has converged, which is not necessarily the same as a parameter convergence. A third alternative is to stop the optimization if the Jacobian and the residual are orthogonal. In

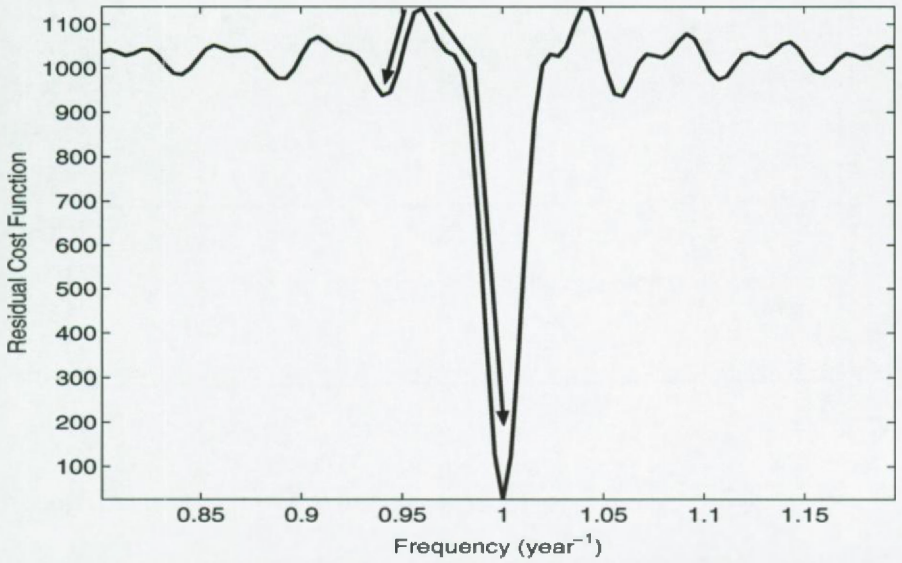

Figure 1: A cross section of the cost function as function along one fundamental frequency. The true frequency is one year^{-1}. The arrows indicates to which local minimum the Levenberg-Marquardt algorithm will converge starting from the initial value at the beginning of the arrow.

order to be sure that all conditions are fulfilled, all three criteria had to be smaller than e.g. 10^{-8}.

8.4.2 Initial values and the optimization strategy

The Levenberg-Marquardt optimization algorithm preforms a local optimization, i.e. it can converge towards a local minimum. In this case, a set of different parameter values exists, that will match better on the same record, because they are associated with a lower cost function minimum. Consequently, the actually estimated parameter values are worthless and the automated model selection is unreliable. To make things even worse, this problem often occurs for this type of models with multiple fundamental frequencies. Therefore, special attention is given to the initialization of fundamental frequencies. The sensitivity of the optimization method w.r.t. the initial values is illustrated in Figure 1, where a cross section of the cost function is shown along one frequency[1]. An initial misfit of four percent, i.e. 0.96 year^{-1} instead of 1 year^{-1}, will already lead to a local minimum. To avoid this, we have used a

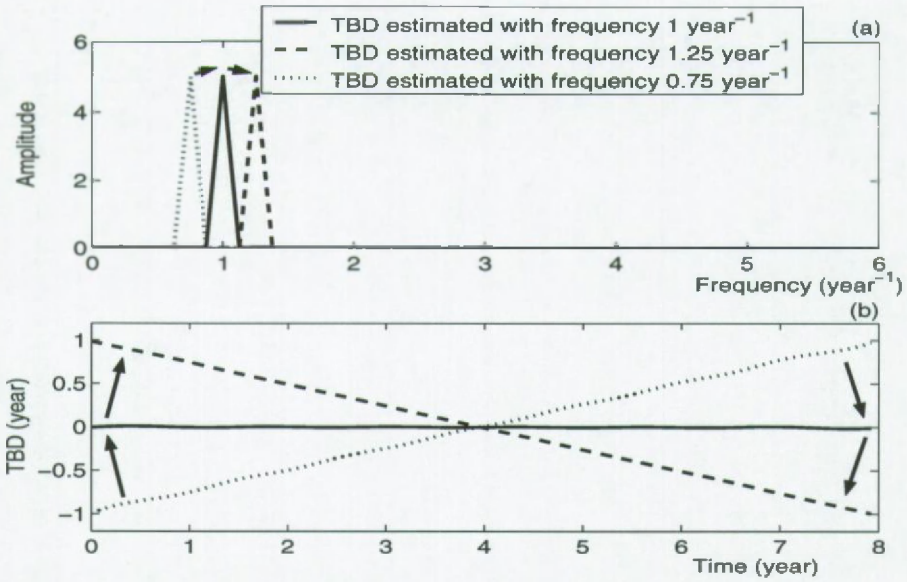

Figure 2: A signal of 100 observations is simulated, consisting of an annual frequency and no variations in accretion rate. From this signal, we estimated a TBD with a window of 12 lines around a frequency of 1, 1.25 and 0.75 year^{-1}. Notice the trends appearing in the TBD, whose slopes are inversely proportional to the mismatch in the frequency. By setting the slope equal to zero, the frequency can be estimated

property of the time base distortion: a linear trend appear with a slope proportional to the misfit of the initial fundamental frequency (see chapter 3 for a proof and Figure 2 for an illustration), i.e. its slope would be negative if the initial value of the respective frequency is larger than the true one and positive if it is smaller. The true frequencies are unknown, but the slope in the time base distortion can be identified. Because (i) this initial estimate of the slope is non-parametric, it is not sensitive to these local minima and can be used to correct the initial estimate of the fundamental frequencies; and (ii) Because this slope is estimated simultaneously with the time base distortion, the initialization of the initial of the frequency will not be influenced by the shape of the actual time base distortion, which is not the case if alternative methods are used to estimate the two non-harmonically related frequencies.

Initial values can be generated following the next procedure:

1. The parameter values used to construct this graph are those used in the simulation (vide infra).

1. Initial values for the fundamental frequencies, ω_m, are chosen by the user. These values are tuned until the slope in the time base distortion is smaller than a numerical stop criterion, e.g. 10^{-15} (no units);
2. Around the first harmonic of each periodic component a non-parametric time base distortion is estimated following the guidelines of [Pintelon and Schoukens, 1996]. To do this, a window width in the frequency domain has to be chosen. We suggest to use the difference between the two initial frequencies as measure for this width. These initial estimates of the time base distortion will consist of the 'true' time base distortion and some measurement noise. This noise is stochastic and will differ on each initial estimate around each fundamental frequency. Therefore, the average of all time base distortions is calculated. This way, the noise cancels out;
3. The parameters, B, can be initialized by matching equation (8-4) on this averaged time base distortion.
4. An improved time scale can be build using equation (8-3). On this time scale an initial time series of the signal can be build and the signal parameters, A_0, A_i and ω_m, can be estimated following the lines of [Pintelon and Schoukens, 1996];
5. The time base distortion parameters, B, can be refined by minimizing a least squares cost function, where the parameters A_0, A_i and ω_m are fixed;
6. Steps 4 and 5 are repeated in a relaxation algorithm until all parameters have converged. Such a relaxation algorithm is avoidable, but decreases the calculation time used in the next step;
7. Finally all parameters are optimized simultaneously in a Levenberg-Marquardt optimization algorithm.

8.5 Model complexity

Once all models with varying complexity are optimized, an automated model selection criterion is used to select the optimal model from this set. The residual cost function will always decrease when the model complexity increases. Therefore, it is not a good idea to use the residual cost function in order to select the proper model complexity. As explained in chapter 5, the cost function is extended with a penalty factor, which increases with the number of parameters

$$\frac{K(\hat{\theta})}{N}\exp\left(\frac{\ln(N)(n_\theta+1)}{N-n_\theta-2}\right) \tag{8-7}$$

where K is the residual cost function, N the number of observations and n_θ the number of parameters. Minimizing this criterion, will estimate the optimal model to describe the record and the time base. Indeed, this criterion trades off model complexity with model variability, such that a more complex model will only be accepted as optimal model if its decrease in cost function is large enough, i.e. compensates for the penalty factor. Because mostly no noise information is provided for paleo-environmental records, the discussion is limited to model selection criteria in the absence of noise information.

In order to test the performance of the model selection criteria and the procedures to avoid local minima, we will first discuss a Monte-Carlo simulation.

8.6 Simulation

In order to illustrate the performance of the proposed algorithms and the model selection strategy, a Monte-Carlo simulation has been run one hundred times. It consisted of two non-harmonically related periodic components with frequency 1 and 1.5 year^{-1} respectively. Each consisted of two harmonics, $A_1 = [1, 0, 0, 1]$, $A_2 = [0, 0.5, 0.5, 0]$, and the number of the time base distortion parameters was four, $B = [1, 1, 0, 0]$. The number of observations was one thousand and the time base spanned twenty years. Zero mean white noise was added in order to have a signal-to-noise-ratio (SNR) of two. The algorithm was initialized with a first frequency varying randomly from 0.75 to 1.25 year^{-1} and a second varying randomly from 1.25 to 1.75 year^{-1}. The most complex model consisted of four signal parameters for each of the two non-harmonically related signals, $h_1 = 4$ and $h_2 = 4$, and eight T.B.D. parameters, i.e. $b = 4$.

The correct model complexity was found 85 times out of the one hundred runs. The selection of a wrong model complexity was mostly a result of convergence toward a bad local minimum, resulting from the choice of the initial values. This is illustrated in Figure 3, where the histograms of the initial values for the two frequencies are shown and their relation with the selection of the correct model. These histograms show that the choice of the first frequency was not critical: the chance that the algorithm diverged seems independent of this initial value. However, if the initial guess for the second frequency was too far from the true one, the algorithm was not able to converge toward a good minimum. The means and standard deviations of the estimated parameters for the 85 correct models, are given in Table 1. From this table we can conclude that the estimated values converges to the true values.

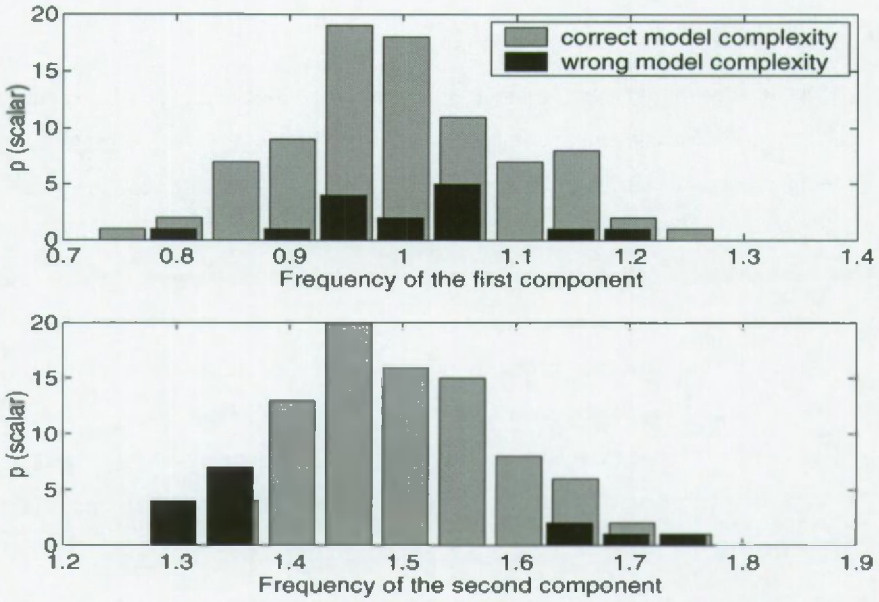

Figure 3: Histogram of the initial values for the first and second fundamental frequency (*p* counts the number of initial values in a particular class).

Table 1: Results of a simulation example for SNR = 2.

SNR=2	true or expected values	mean	standard deviations
f_1 (year^{-1})	1	1.00	10^{-3}
f_2 (year^{-1})	1.5	1.50	10^{-3}
A_0	0	0.00	0.02
A_1	[1,0,0,1]	[1.00,0.02,-0.01,1.00]	[0.02,0.10,0.05,0.03]
A_2	[0,0.5,0.5,0]	[0.01,0.49,0.50,-0.01]	[0.05,0.04,0.03,0.09]
B	[1,1,0,0]	[1.00,0.97,-0.01,-0.02]	[0.14,0.15,0.38,0.21]
$K(\hat{\theta})$	150.25	151.77	0.06

Also, the residual cost function behaved as expected (the expectation value of this cost function[1] was 151.25) and can thus be used for model selection purposes. Thus, even in the

presence of high stochastic noise (SNR=2), mostly the model selection criterion was able to predict the correct model complexity and the parameter values converged to the true values. If the SNR would increase, the performances of the automated model selection and optimization will further improve. So, we may conclude that the algorithm is quite robust in the presence of stochastic noise.

1. The expectation value can be calculated, because the variance of the noise is known, i.e. 0.55, corresponding to a SNR of two.

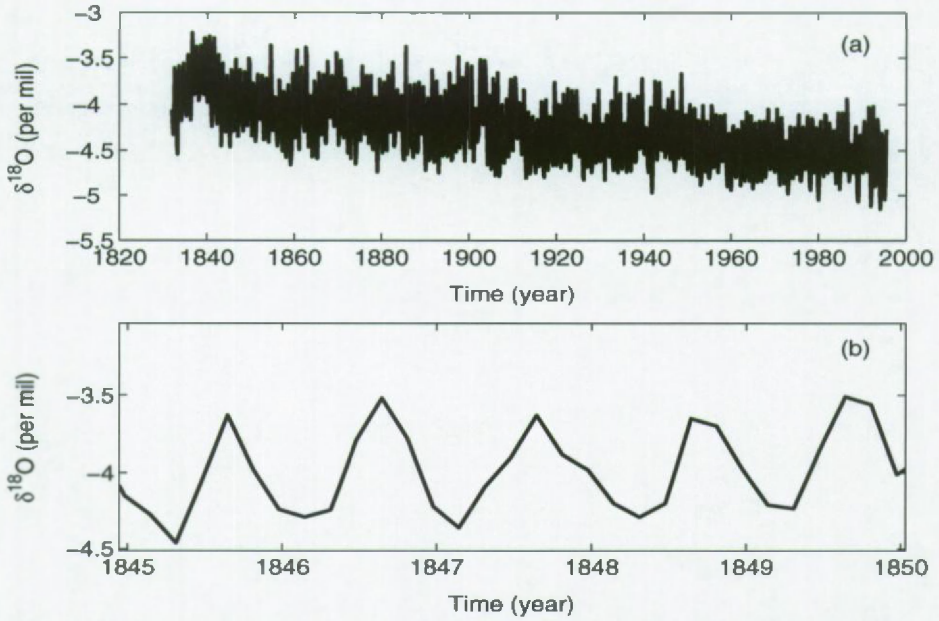

Figure 4: The raw data of the stable isotope record: (a) complete window in time domain and (b) a detailed window.

8.7 Experimental results

8.7.1 Coral stable isotope data

In a first case study the $\delta^{18}O$-record of a 163 year coral is examined. 983 observations were taken from this record and yearly maxima were used as anchor points to date the record and get a first glance of record's content. Figure 4 shows the data, where an annual variation and a trend can easily be distinguished. However, identifying anomalies, like El Niño variations, is difficult in Figure 4. Instead, a more precise description is desired and as a first step the Fourier transform is calculated (Figure 5): at the left side, the slow variations can be separated, but most of the variation seems to be caused by the two or three harmonics of the annual periodicity. Note that the frequency of the third harmonic is equal to the Nyquist

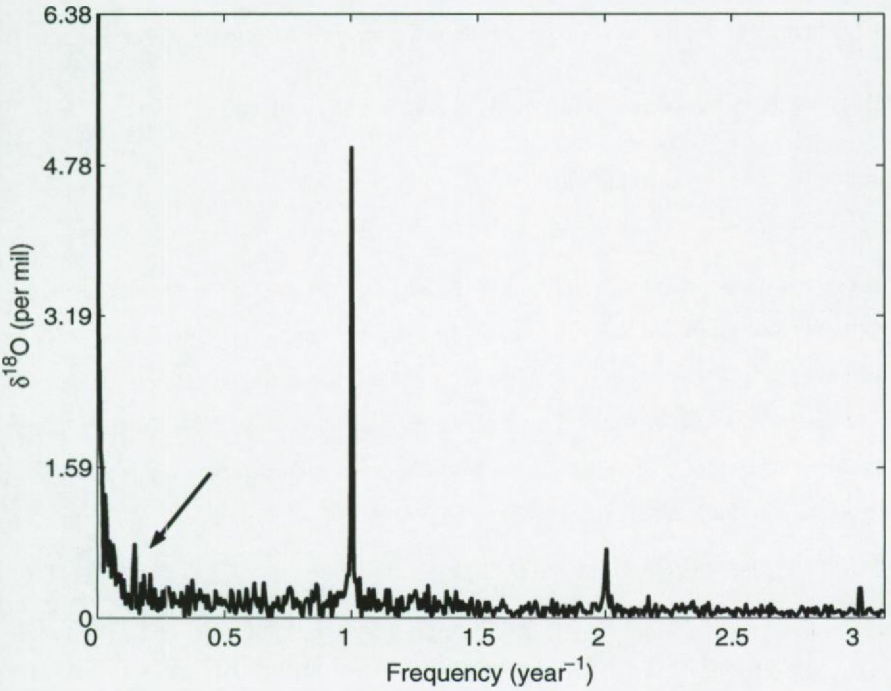

Figure 5: Spectrum of the raw data.

frequency, which is the highest frequency that can be seen in a Fourier spectrum with this sampling frequency. The variation due to El Niño is captured in the small peak right of the trend (see arrow in Figure 5). Its periodicity is about seven years. However, the peak is so small compared to the other variations of the record, that it may have been created by the choice of the anchor points or that is simply negligible, compared to the disturbing noise.

In order to take all these different effects into account, the record, $s_{\text{record}}(t_n)$ is formally decomposed into the signal model, $s_{\text{model}}(t_n)$, defined in (8-1), a trend or colored noise term, $v(t_n)$ and a white noise term, $e(t_n)$, taking the instrumental noise into account:

$$s_{\text{record}}(t_n) = s_{\text{model}}(t_n) + v(t_n) + e(t_n) \tag{8-8}$$

Estimating these different components will:

1. characterize the trend or colored noise;

2. if necessary the time base estimated with the anchor point method can be refined;

3. verify whether or not the El Niño peak is significant, compared to the disturbing noise;

4. verify how many harmonics of the yearly cyclicity are significant;

5. characterize the instrumental noise.

In order to characterize the trend, a $10th$ order Legendre polynomial was matched on the record. We have assumed that the changes in the trend are slower than the variations in the time base. Under these conditions the trend can be estimated and subtracted before the signal and time base model are matched. It would have been better to estimate the trend parameters simultaneously with the signal and time base distortion parameters, but this approach is too time consuming with the present computer speed.

Table 2: Results of the analyses of the coral record.

	No time base distortion		Time base distortion	
	One periodicity	**Two Periodicities**	**One periodicity**	**Two Periodicities**
Cost function (per mil)²	11.656	11.541	11.650	11.089
MDL value	0.0125	0.0126	0.0125	0.0120
SNR	1.54	1.55	1.54	1.58
Selected h_1	2	2	2	2
Selected h_2	-	1	-	1
Selected b	-	-	0	1
period 1 (year)	0.99	0.99	0.99	0.99
period 2 (year)	-	7.64	-	6.85

The most complex model from the model set consisted of $(M, h_1, h_2, b)_{\max} = (2, 3, 3, 14)$. The time base distortion can be interpreted as a correction on the time base generated by the anchor point method. In this context: the time base constructed with the anchor point method will be correct, if the time base distortion is zero. Otherwise some variations in the time base

are still significant. The benefit of starting from the anchor points method's time base is that the correction is smaller and that thus the algorithms converge faster. In order to visualize the different complexity levels, Table 2 shows the results of the optimal model with only one periodicity with and without a time base distortion and the optimal model with two periodicities, again with and without a time base distortion. On all these different levels of complexity two harmonics were found for the first periodicity. The two sets with two periodicities selected the second periodicity, but with only one harmonic. For sets with a time base distortion, the results differed: in the set with one periodicity no time base distortion parameters were significant, while two parameters were selected in the set with two periodicities. A remarkable difference is the numerical value of the second periodicity. If no time base distortion is used it is 7.64 year, which dropped to 6.85 year when a time base distortion is present in the model. When a time base distortion is modelled too, the variation around this El Niño periodicity is less spread, which is seen in the higher signal-to-noise-ratio (1.58 versus 1.55). The time base is only marginally influenced by the time base distortion: the RMS value of the time base distortion is ten thousand times smaller than the RMS value of the time base. So, the reason for selecting a time base distortion is mainly due to the slightly higher amplitude of the El Niño variation.

Figure 6(a) and (b) show the residual, after the annual variation is subtracted, i.e. the trend and the El Niño variation. The signal-to-noise-ratio of the El Niño peak is 2.9, where the signal is calculated as the amplitude of the peak and the noise as the standard deviation of the residual error, without the trend.

Two harmonics are selected for the yearly frequency (Table 2). Accordingly, the third peak at the Nyquist frequency was not significant. For the yearly periodicity and its first overtone, the signal-to-noise-ratios are respectively 18.0 and 2.4.

However, for the model selection criterion and for the estimation of the uncertainties, we have assumed that the residual noise was white noise. This assumption is violated as can be seen on Figure 6b: the noise slightly decreases with increasing frequency. It is very well possible that the third harmonic of the yearly periodicity would become significant if a different noise model was implemented.

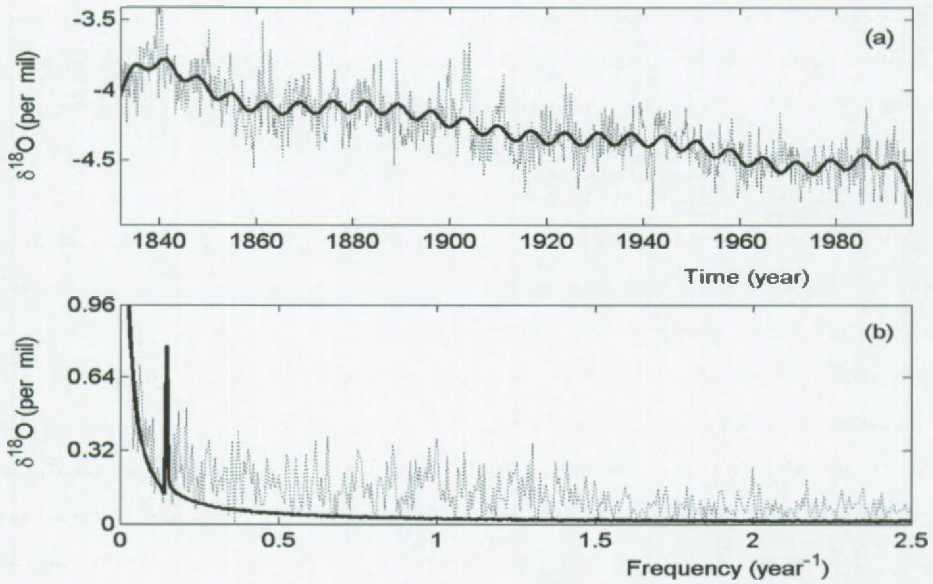

Figure 6: Residual when yearly variations are subtracted: (a) in the time domain and (b) in the frequency domain. The measurements are shown in grey and the model in black. Notice that the 6.8 year El Niño variation can much easier be identified in the latter.

To conclude, the initial estimate for the time base could slightly be improved. Two harmonics of the yearly periodicity have been identified, as well as a small El Niño peak. However, the white noise assumption may not hold, which could possibly bias the results.

8.7.2 Vostok ice core record

The 3310 m long Vostok ice core record is used to compare this approach with

1. the spectral components of the proposed GT4 time base [Petit *et al.*,1999]; and
2. the AR model proposed by Broersen *et al.* [2003].

For the approach proposed here, the periodic content is examined on the GT4 time base, with and without a time base distortion.

This 3310 m long proxy record spans a period of approximately 420 kyear. The core is analyzed with a spatial distance of one meter between the samples. An initial time scale is linked to the depth by combining theoretical flow and accumulation models for the ice, named the GT4 time base. This time base construction is based on three hypotheses:

1. Accumulation rate varied in proportion to the derivative of the water vapour pressure with respect to temperature;
2. Accumulation between Vostok and Dome B varies linearly with distance along the line connecting those two sites;
3. Vostok ice at 1534 m is formed 11 kyear ago and ice at 3254 m 390 kyear ago.

In this study we focus on the temperature variations as they were calculated from the deuterium observations. Additional information and the data can be downloaded from http:// www.ngdc.noaa.gov.paleo/). As can be seen on Figure 7, the sampling frequency decreases with the age/depth of the observations.

8.7.2.1 Without a time base distortion

A. Lomb-Scargle spectra and Blackman-Tukey window

Petit *et al.* [1999] find that "The power spectrum of ΔT_I (...) shows a large concentration of variance (37%) in the 100-kyr band along with a significant concentration (23%) in the obliquity band (peak at 41 kyr)", and concludes that "Spectral analysis emphasizes the dominance of the 100 kyear cycle [...] and a strong imprint of 40 and/or 20 kyear periodicities despite the fact that the glaciological dating is tuned by fitting only two control points in the 100 kyear band". To come to this conclusion, they have used a Blackman-Tukey method. This method is an adaptation of the discrete Fourier transform, which assumes that the observations are continuous at the borders of the measurement window. This assumption causes interferences between the different spectral components, which is called leakage (e.g. [B. Porat, 1997]). To remedy this undesirable effect of a rectangular measurement window many other windows are proposed. The Blackman-Tukey window is one of these windows that can be used.

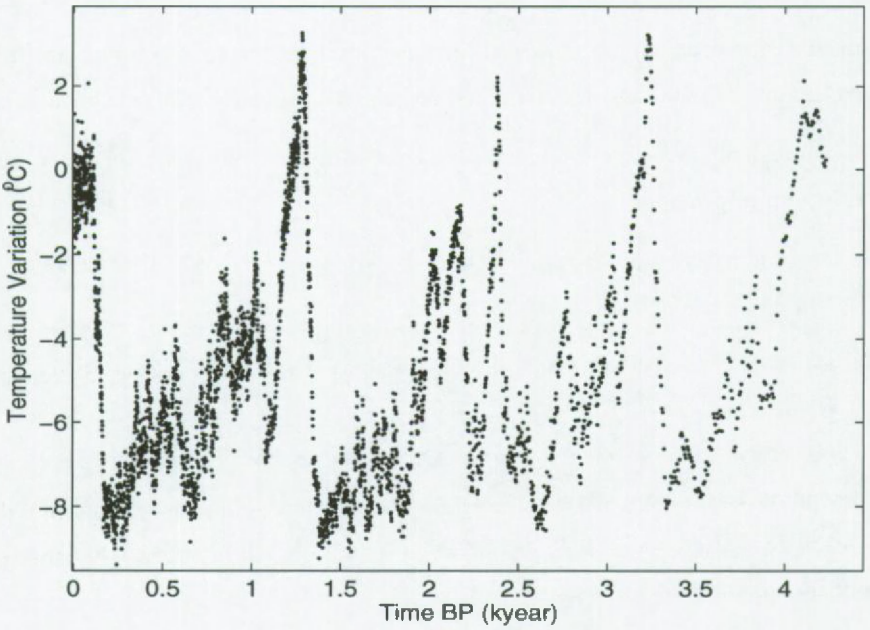

Figure 7: Variation in paleoclimatic temperature, extracted from the Antarctica Vostok ice core and ordered along an initially estimated time scale, i.e. GT4.

B. Auto-Regressive model

Conversely to the conclusion made by Petit *et al.*, Broersen *et al.* [2003], employed an auto-regressive (AR) model, combined with a model selection criterion. The AR-model is an empirical model for correlated time series (e.g. [Turcotte D.L., 1997]). The model and the model selection criterion are described in chapter 5, paragraph 5.4.3. On the same time series and came to the conclusion that: "No peaks at the low frequencies have been detected with the time series methods. This conclusion is in contrast with the spectra that have been computed with Fourier techniques like Lomb-Scargle spectra for missing data or modified periodograms for consecutive data [Petit *et al.*, 1999]."

So, both come to a opposite conclusion: Petit *et al.* uses the spectral interpretation to illustrate the dominance of the 100 kyear and 41 kyear periodicity, while Broersen *et al.* employs parametric time series analysis with a model selection criterion to illustrate that no

periodicity is significant at all. At first sight, one may see some cyclicity in the record, which can give the impression that Broersen *et al.*'s criteria have been too strict. However, the difference between 'seeing' and 'self fulfilling prophecy' can easily vanish and, therefore, objective model selection criteria, like proposed by Broersen *et al.* are needed to define which components in a signal can be identified as significant and which cannot.

C. Parametric signal model, coupled to a model selection criterion

The algorithms proposed in here are suitable for this purpose, because the model selection criterion is able to distinguish significant variations from stochastic noise. Notice, however, that we have to assume that the true model belongs to the model set.

Note that in order to reduce the calculation time and needed memory capacity, the data-record was resampled with a sample period of approximately 425 year (1000 observations), while the original one was approximately 127 year (3110 observations). This was achieved using a linear interpolation algorithm. This resulted in newly constructed observations, which are equidistant. The results can slightly be influenced because the model selection rules are a function of the number of observations.

The maximum number of non-harmonically related periodicities was $M_{\max} = 2$, the maximum number of harmonics in the first periodicity was $h_{1,\max} = 7$ and in the second $h_{2,\max} = 2$. If the time base was refined, the maximum number of time base distortion parameters was $b_{\max} = 10$.

First, the GT4 time base was used, i.e. the time base distortion parameters were fixed at zero, i.e $B = [0, 0]$. The lowest model selection criterion, $MDL_c = 1.80$ (see Table 3), corresponds to a model with two periodicities ($h_1 = 2$ and $h_2 = 1$). The $SNR = 2.86$[1], which is slightly better than in the model with only one periodicity: in this model, the 101.9 kyear periodicity consists too of two harmonics, $h_1 = 2$, its model selection criterion, $MDL_c = 2.32$, is slightly higher than that of the model with the 101.5 and 39.9 kyear

1. The signal-to-noise ration is defined as the ratio of the root-mean-square of the model and the root-mean-square of the residual (measurement minus signal model).

Table 3 Results of the analyses of the Vostok ice core record.

	No time base distortion		Time base distortion	
	One periodicity	Two Periodicities	One periodicity	Two Periodicities
Cost function ($°C^2$)	2213	1659	494	493
MDL value	2.32	1.80	0.578	0.591
Selected h_1	2	2	6	6
Selected h_2	-	2	-	1
Selected b	-	-	10	10
period 1 (kyear)	101.9	101.5	102.6	102.5
period 2 (kyear)	-	39.9	-	43.4

period. For that reason, we may conclude that if the GT4 time base is used, both cycles are significant (Figure 8). However, a lot of the remaining variation cannot be explained by instrumental noise only (e.g. the peaks between the 101.5 and 39.9 kyear cycle). Note, that a lot of this unexplained variation is found in the low frequency range, i.e. in the neighborhood of the harmonics. So, a white noise model, describing measurement noise will not be sufficient to describe this residual. However, the 23 kyear, for example, is not yet taken into account in this model.

8.7.2.2 With a time base distortion

In a next step, we will try to describe the remaining variation as errors in the time base. Therefore, the time base distortion parameters, B, were estimated too. The most remarkable outcome of the automated model selection criterion is that it selected a model with only one periodicity (with a period of 102.6 kyear). However, $h_1 = 6$ harmonics were selected and $b = 10$ time base distortion parameters are used to correct the GT4 time base (see Table 3). The spectrum of the signal on the new time base is shown in Figure 9. Note, first of all, that the part of the spectrum which is not explained by the 102.6 kyear periodicity has become much smaller: the root-mean-square of the residual error has dropped from 3.16 to 0.97 $°C$,

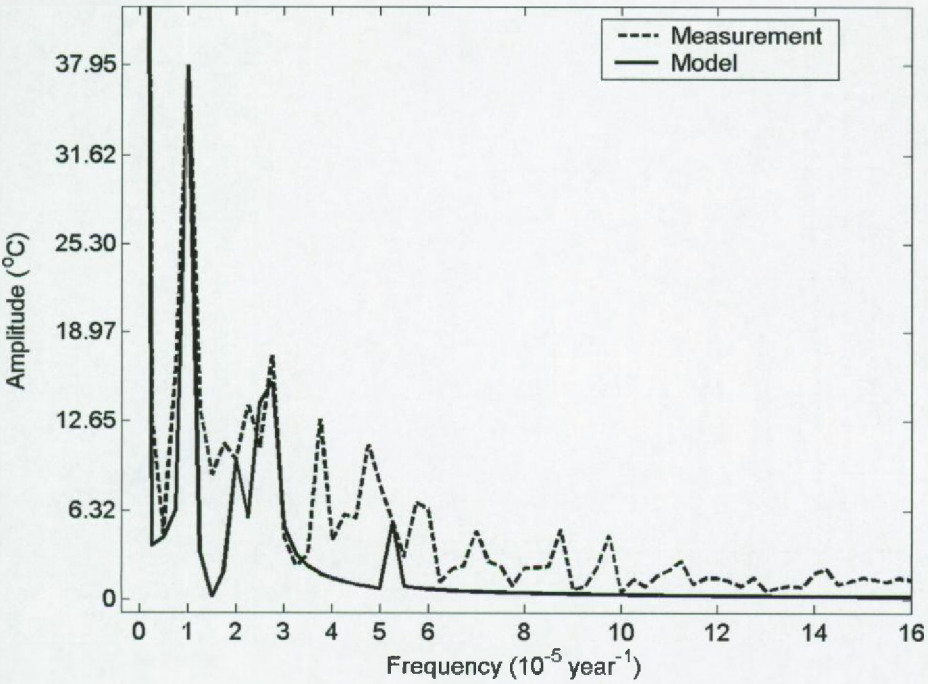

Figure 8: Spectrum of the temperature variation on the GT4 time base. The 100 kyear and 41 kyear periodicities can be seen at frequencies of 1 and 2.5 10^{-5} year^{-1}. The full line represents the model and the dashed line the measurements.

which corresponds to a $SNR = 5.49$. In addition, this residual is now more equally spread over all frequencies, as one could expect from measurement noise.

Let us now focus on the "41 kyear cycle". If we limit the model set, from which the model selection criterion has to choose a model, to those models consisting of at least the two cycles. The complexity of each cycle vary as defined earlier, i.e. the most complex model consisted of $(h_1, h_2, b)_{\max} = (7, 2, 10)$, with M fixed at 2. The model, which is selected under these condition, is described in Table 3. Both models, with one and two periodicities are almost identical. The only difference is a small 43.4 kyear periodicity, whose amplitude is ten times smaller than those of the 102.5 kyear cycle. The disappearance of the "41 kyear periodicity" can be explained as an artefact due to errors in the GT4 time base. The record on the GT4 and on the refined time base are shown in Figure 10: the root-mean-square value of

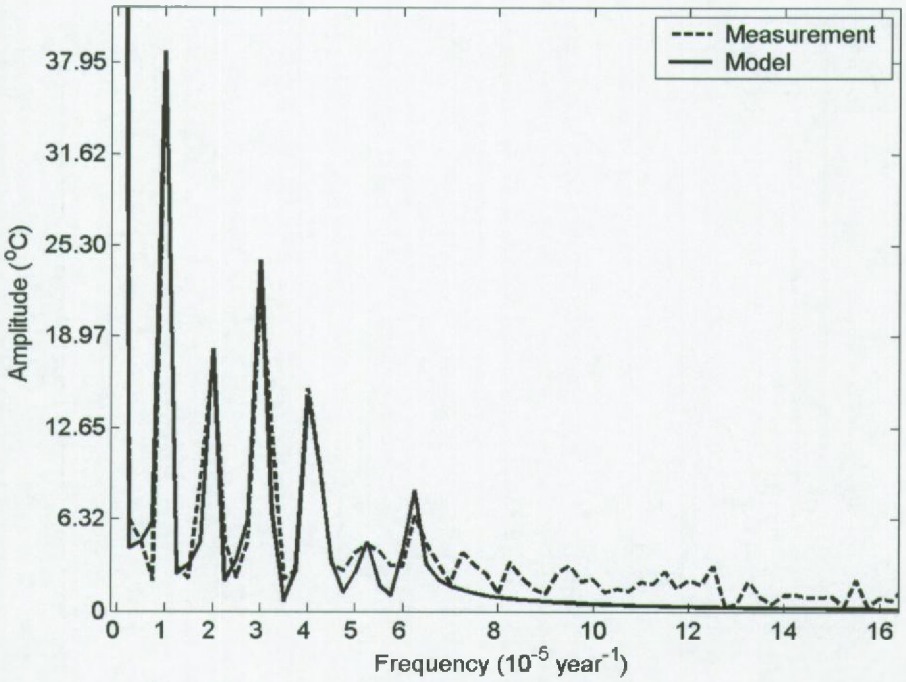

Figure 9: Spectrum of the temperature variation on the refined time base. The 100 kyear period can be seen at frequencies of 10^{-5} year^{-1}. Note that the 41 kyear period at a frequency of 2.5 10^{-5} year^{-1} has disappeared.The full line represents the model and the dashed line the measurements.

the time base distortion is 17.2 kyear, the largest distortions have an amplitude of about 40 kyear, which is about 10% of the total time base. This is illustrated in Figure 11, where the time-depth relation of the GT4 and the refined time base are shown.

An alternative explanation of the time base distortion is that the GT4 time base is correct, but that the frequency of the 100 kyear periodicity is modulated. The proposed model is not capable to distinguish distortions in the time base from phase or frequency modulations. However, even in that case the 41 kyear cycle is insignificant in this record and appeared only because of the distortion of the 100 kyear cycle, because a phase demodulation can be transformed into a frequency modulation. So, we cannot distinguish the time base distortion

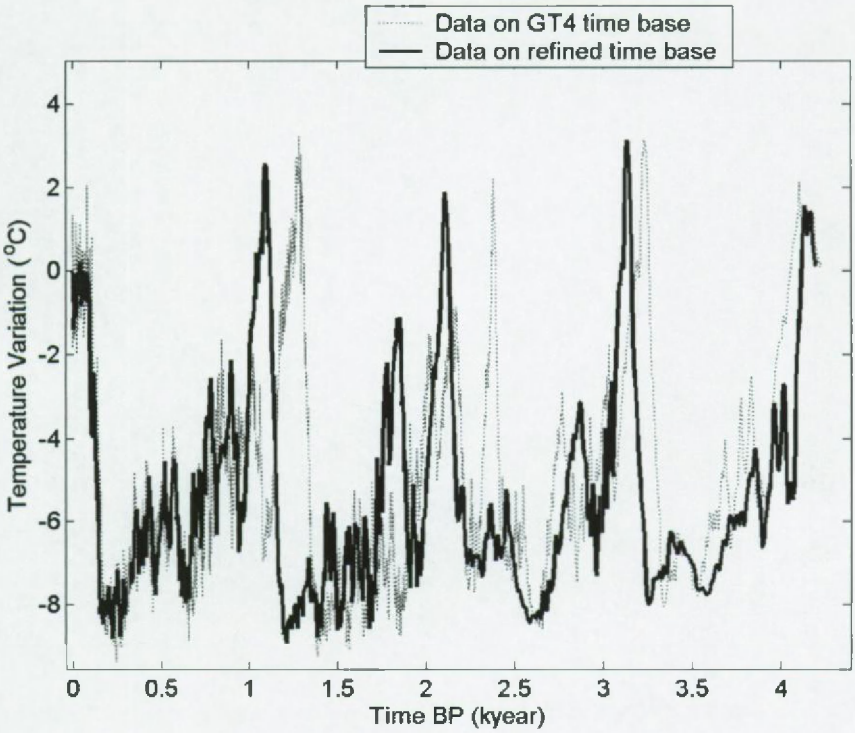

Figure 10: Vostok temperature data on GT4 time base and on the refined time base.

from a potential frequency modulation, but in both cases, the 41 kyear cycle is not significant.

To summarize, we have started from the assumptions that

1. the record can be described by at most two periodicities and stochastic white noise;
2. each periodicity consists of one or more harmonics;
3. the time base can be improved;
4. the time base distortion is bandwidth limited;
5. the model selection criterion chooses the optimal model from the model set and the optimization algorithms converged towards the correct minima.

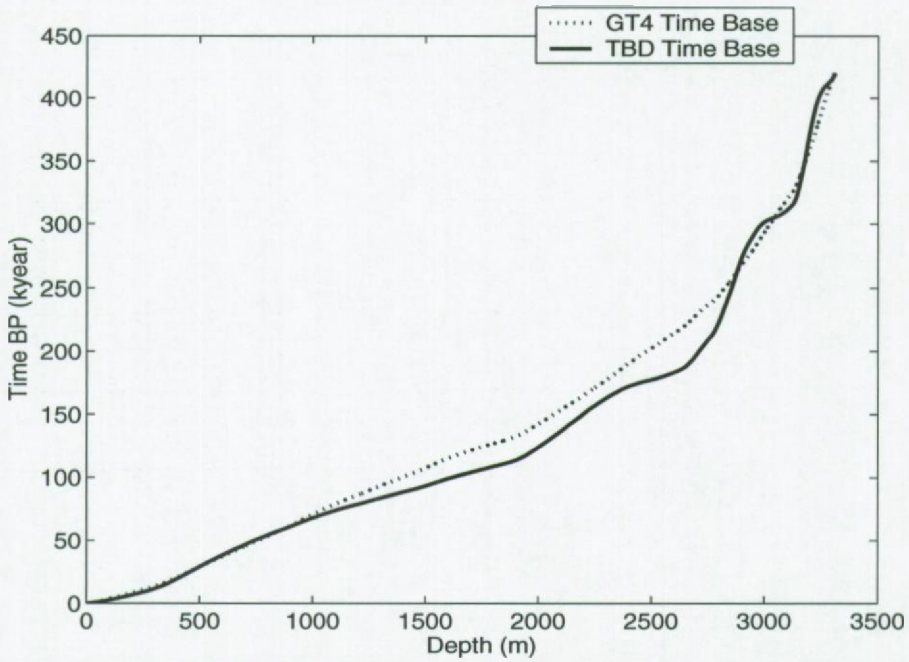

Figure 11: GT4 and the refined time bases (TBD) are shown as function of the depth. The time scale of 420 kyear is not altered, but the relative ages of the observations differs for both time models.

From these assumptions it can be concluded that the 100 kyear cycle is significant, but that the 41 kyear cycle is an artefact, which appeared in the original GT4 time base spectrum, due to errors in the time base.

So, if we come back to the discussion between Broersen *et al.* and Petit *et al.*, this approach ends up, somewhere in the middle: on the one hand, if we limited the discussion to the observations dated on the GT4 time base, the 100 kyear and 41 kyear periodicity are significant. On the other hand, it is shown that spectral interpretation of time series is very vulnerable to errors in the time base. Such an error is potentially made by both interpretations. Here we have shown that if the 100 kyear period is present, the 41 kyear cycle becomes insignificant. Table 2 summarizes this discussion.

Table 4: Summery of the different spectral interpretations of the temperature in the Vostok record.

Author	Model	Model selection criterion	Time base	Result	
				100 kyear period	40 kyear period
Petit *et al.*	Lomb-Scargle Blackman -Tukey	no	GT4	yes	yes
Broersen *et al.*	Auto-regressive	yes	GT4	no	no
this approach	Periodic signal	yes	GT4+distortion	yes	no

8.8 Conclusion

An algorithm has been proposed which can simultaneously identify multiple periodic components and a time base in paleo-environmental archives. The algorithm is expanded with an automated model selection step, which is able to select the most reliable model complexity and this as well for the signal model as for the time base model. The algorithm was tested on simulations and on two real world examples. First, a 163 year old coral record was examined. Herein, it was unclear whether or not a small periodic component with a periodicity of 6.8 year was significant, compared to much larger variations, due to the annual variation of the signal. The algorithm identified both periodic signal models. Next, the Vostok ice core record was processed. Herein the 100 kyear cycle was significant, while the 41 kyear cycle was not. This conclusion was based on the assumption that both cycles can be modeled by a periodic signal model with non-harmonically related frequencies.

References

[1] Broersen P.M.T., de Waele S., Bos R., 2003. Estimation of Autoregressive Spectra with randomly missing data. IMTC 2003 - Instrumentation and Measurement Technology Conference, Vail, CO, USA, 20-22 May 2003, pp. 1154-1159.

[2] Fallon S.J., McCulloch M.T., van Woesik R. and Sinclair D.J., 1999. Corals at their latitudinal limits: laser ablation trace element systematics in porites from Shirigai Bay, Japan. Earth and Planetary Science Letters 172, 221-238.

[3] Finch A.A., Shaw P.A., Weedom G.P. and Holmgren K., 2001. Trace element variation in speleothem aragonite: potential for palaeoenvironmental reconstruction. Earth and Planetary Science Letters 186, 255-267.

[4] Ghil M., Allen M.R., Dettinger M.D., Ide K., Kondrashov D., Mann M.E., Robertson A.W., Saunders A., Tian Y., Varadi F., and Yiou P., 2002. Advanced spectral methods for climatic time series. Reviews of Geophysics, 40(1), 1-41.

[5] Hasselmann K., 1976. Stochastic climate models Part I. Theory. Tellus, 28(6), 473-485.

[6] Harris F.J., 1978. On the use of Windows for harmonic analysis with the discrete Fourier

Transform, Proceedings of the IEEE, 66(1), 51-84.

[7] Jouzel J., Waelbroeck C., Malaize B., Bender M., Petit J.R., Stievenard M., Barkov N.I., Barnola J.M., King T., Kotlyakov V.M., Lipenkov V., Lorius C., Raynaud D., Ritz C., and Sowers T., 1996. Climatic interpretation of the recently extended Vostok ice records. Climate Dynamics 12(8), 513-521.

[8] Kuhnert H., Pätzold J., Schnetger B. and Wefer G., 2002. Sea-surface temperature variability in the 16th century at Bermuda inferred from coral records. Palaeogeography, Palaeoclimatology, Palaeoecology 179, 159-171.

[9] Marshall J.F. and McCulloch M.T., 2002. An assessment of the Sr/Ca ratio in shallow water hermatypic corals as a proxy for sea surface temperature. Geochimica et Cosmochimica Acta 66, 3263-3280.

[10]Schulz M., Stattegger K., 1997. SPECTRUM: spectral analysis of unevenly spaced paleoclimatic time series. Computers and Geosciences, 23, 929-945.

[11]Schulz, M., Mudelsee M., 2001. REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. Computers and Geosciences, 28, 421-426.

[12]Petit J., Jouzel J., Raynaud D., Barkov N., Basile I., Bender M., Chapellaz J., Davis J., Delaygue G., Demotte M., Kotlyakov V., Legrand M., Lipenkov V., Lorius C., Pépin L., Ritz C., Saltzman E. and Stievenard M., 1999. 420,000 years of climate and atmospheric history revealed by the Vostok deep Antarctic ice core, Nature, 399, 429-436.

[13]Pintelon R. and Schoukens J., 1996. An improved sine-wave fitting procedure for characterizing data acquisition channels, IEEE trans. Instrum. Meas., 5, 588-593.

[14]Porat B, A course in digital signal processing. John Wiley & Sons, Inc., New York, 1997.

[15]Sinclair D.J., Kinsley L.P.J. and McCulloch M.T., 1998. High resolution analysis of trace elements in corals by laser ablation ICP-MS. Geochimica et Cosmochimica Acta 62 (11), 1889-1901.

[16]Torrence C. and Compo G.P., 1998. A practical guide to wavelet analysis, Bulletin of the American Meteorological Society, 79(1), 61-78.

[17]Turcotte D.L., Fractals and Chaos in Geology and Geophysics (Second edition). Cambridge University Press, Cambridge, 1997.

[18]Verheyden S., Keppens E., Fairchild I.J., McDermott F. and Weis M., 2000. Mg, Sr and Sr isotope geochemistry of a Belgian Holocene speleothem: implications for palaeoclimatic

reconstructions. Chemical Geology 169: 161-144.

[19]Wei G., Sun M., Li X. and Nie B., 2000. Mg/Ca, Sr/Ca and U/Ca ratios of a porities coral from Sanya Bay, Hainan Island, South China Sea and their relationship to sea surface temperature. Palaeogeography, Palaeoclimatology, Palaeoecology 162, 59-74.

CHAPTER 9

CONCLUSION AND FUTURE PERSPECTIVES

Before addressing some topics for further research, it is worthwhile to review roughly what has been achieved throughout the previous chapters.

9.1 Recapitulation

In the second chapter, the calibration a LA-ICP-MS measurement instrument is refined. Therefore, we have developed an estimator for the drift based on multiple internal standards instead of one single internal standard. The main advantage of this approach is that the noise on these standards is partially cancelled out. Some additional benefits are that individual uncertainties on the observations can be estimated and that an internal quality check is performed. It is shown that this approach is a generalization of the classical single internal standard calibration.

The remaining part is devoted to the estimation of the time base for proxy records. This time base is divided in two parts. One takes the average constant accretion rate into account, while the other estimates the distortions from this average value. It is not possible to estimate the time base, without a model for the measured signal. Except for the last chapter, we have mainly focused on signals, modeled by periodic functions. For this class of signals it is possible to generate a reasonably good initial values for the time base distortion parameters. This is done by a phase demodulation technique in the frequency domain.

In the third chapter the concept of a time base distortion is discussed and a procedure to generate an initial estimate of the time base distortion is presented.

This procedure has two main drawbacks:

1. the estimated accretion rate is biased at the borders, because of the Gibbs phenomenon; and
2. the size of the window used to separate the time base distortion is chosen by the investigator.

In order to solve these problems, a parametric model for the time base distortion and for the signal are introduced. In the engineering community, the first problem is usually solved by dropping the observations near the borders. However, because of the small number of observations, the time and money consuming measurements, this option cannot be applied on environmental data records.

Therefore, in chapter four the time frame in the model is enlarged, compared to the measurement window. As a result, the Gibbs phenomenon, which mainly occurs in these virtual borders, can be dropped. The main disadvantages of this approach are that the size of these extra borders has to be chosen by the user and that the bandwidth of the time base distortion decreases with an increasing window size. A simple procedure is given, which can estimate the optimal size of the borders, while the bandwidth of the time base distortion can be enlarged by increasing the number of time base distortion parameters.

The second drawback is also related to the small number of observations. On such data-records, classical model selection criteria, like AIC or MDL fail. In chapter five, we have shown that the residual cost function decreases to zero if the number of parameters becomes equal to the number of observations. Therefore, these model selection criteria have the tendency to select too complex models. This bias is remedied by adapting the model selection criteria. A comparative study of the different model selection criteria has been made, based on simulations. Herein, we concentrated on short data records. Under this condition, the adapted rules outperform. The MDL rules seems to perform better than the AIC rules if the signal-to-noise-ratio is sufficiently high; otherwise the AIC rules are preferred. The numerical value of this critical signal-to-noise-ratio can be estimated by a Monte-Carlo simulation. Note finally that if the number of observations increases, the refined rules converges toward the classical ones.

The idea of a time base distortion, the reduction of the Gibbs phenomenon and a model selection criterion have been implemented in a software program. In chapter six, the outline of this algorithm is discussed and several smaller and more technical problems are handled.

In the last 25 years several alternative methods have been proposed. These methods are compared with both approaches proposed here, i.e. the non-parametric phase demodulation and the parametric method combined with a model selection criterion. First we have compared the basic ideas, the assumptions, tuning parameters and limitations of all these methods. Next, the algorithms have been implemented in order to test the sensitivity to stochastic noise and to model errors. Evidently, not all possible model errors could be discussed. Therefore, we have tested the presence of unknown overtones, unknown amplitude modulations, unknown trends and hiatuses added to a periodic signal model. In general, parametric models seems to perform better and due to the model selection criterion, the proposed parametric model outperformed.

To illustrate that this approach is not limited to periodic signals models only, we have finally adapted the algorithm, so that non-harmonically related periodic signals can be processed too. As a first example the stable isotope record from a coral was processed. Herein, a small El Niño effect has been identified, besides the large annual variations.

Next, and finally the temperature profile of the Vostok ice core record was processed. This example was chosen, because (i) it is a famous records and (ii) Petit *et al.* came to the conclusion that the 100 kyear was dominant and that the 41 kyear cycle had a strong imprint. This result is based on a Lomb-Scargle spectra for missing data. However, Broersen *et al.* came to the opposite conclusion. According to an AR-model, which was combined with a model selection criterion, none of both periodicities were significant. Both assumption were based on the GT4 time base. As a next step in this discussion, the GT4 time base was optimized, based on the assumption that both cycles are periodic. The optimization algorithm was combined with a model selection criterion in order to select the model with the appropriate complexity. Surprisingly, this led to the conclusion that only the 100 kyear periodicity is significant. In addition, five overtones of this frequency could be identified as well. However, from an identification point of view, we cannot make the distinction between the distortion of the time base and a frequency modulation of the 100 kyear periodicity.

Anyway, in both scenarios, the time base distortion and the frequency modulation, the 41 kyear cycle is insignificant.

9.2 Future perspectives

This paragraph enumerates some ideas and dreams for the future. In this scope, first some technical improvements of the algorithms are proposed (9.2.1 and 9.2.2). Where are the limits of a parametric deterministic time base distortion? Stochastic disturbances in the time base are discussed in (9.2.3).

Very recent, we have realized that observations are not points in space or time, but integrated intervals. These intervals can be equal in space, but do not necessarily spans equal time intervals, due to variations in accretion rate. So, these intervals connect the proxy-signal with the accretion rate. How this connection can be taken into account is explained in (9.2.4).

In the future, some extensions can be proposed to combine multiple proxies in the same data record (9.2.5) and one proxy in multiple cores (9.2.6). This opens the way to combine different ice core records and subtract the significant information from this class of records, and not from individual cores.

The simple signal models, based on periodicity can be replaced by time series analysis (9.2.7) or by physical models. In (9.2.8) we propose to exploit the dynamics in these physical models to construct the time base distortion.

9.2.1 Non-equidistant observations

So far, we have assumed that the observations were taken equidistant in space. If this was not the case an interpolation technique was used. This interpolation can alter the spectral contents of the record, by suppressing the higher frequency components. The algorithms can be extended so that non-equidistant records can be processed too, without the use of an interpolation.

9.2.2 Alternative bases for the time base distortion

So far, we have limited ourselves to orthogonal polynomials or Fourier series as a basis for the time base distortion. Under particular circumstances other basis functions can perform

better. If, for example, negative accretion rates occur in the model, which is impossible, a better solution than the one proposed here, is to use positive polynomials. In this basis negative accretion rates cannot occur.

If hiatuses can very well be present in the record, splines can be used, which will better match on the time base distortion. Herein, an additional parameter has to be optimized, which defines the position of the hiatus.

9.2.3 Time jitter

So far, we have proposed a deterministic time base distortion, which could be characterized by a limited number of parameters. However, imagine e.g. the start of spring each year: in this region, it starts mostly in the end of March, but sometimes it can already be April. Such stochastic variations in the time base can be modeled by an additive perturbation term, called time jitter [Schoukens *et al.*, 1997]. The order of magnitude of this stochastic disturbance is characterized by one parameter, i.e. the standard deviation, which can be estimated.

9.2.4 Accretion rates and the signal model

So far, it was implicitly assumed that an observation was a point in time or space domain. However, in reality the sampling procedure samples a certain volume of material along a line, on a surface or in a drill hole, depending on the measurement instrument. This results in observations which are actually an averaged value over this volume. If the instrument settings remain unchanged during the measurement procedure, the distance intervals covered are mostly the same. However, if the accretion rate differs over the sample, the corresponding time intervals are not equal anymore. If, for instance, the accretion rate was high, the time interval over which is averaged will be small and vice versa. This effect is simulated in Figure 1. The horizontal lines represent the averaged signal, sampled from the 'true' signal (continues line). Notice that the amplitude decreases with decreasing accretion rate (in fact also the phase changes, but that cannot be noticed in this Figure). So, if we simply use a signal model to estimate the accretion rate, an error will be made, because the accretion rate influences the signal model. The integrated signal, $\tilde{S}(t_n, \theta_{signal})$ can formally be written as

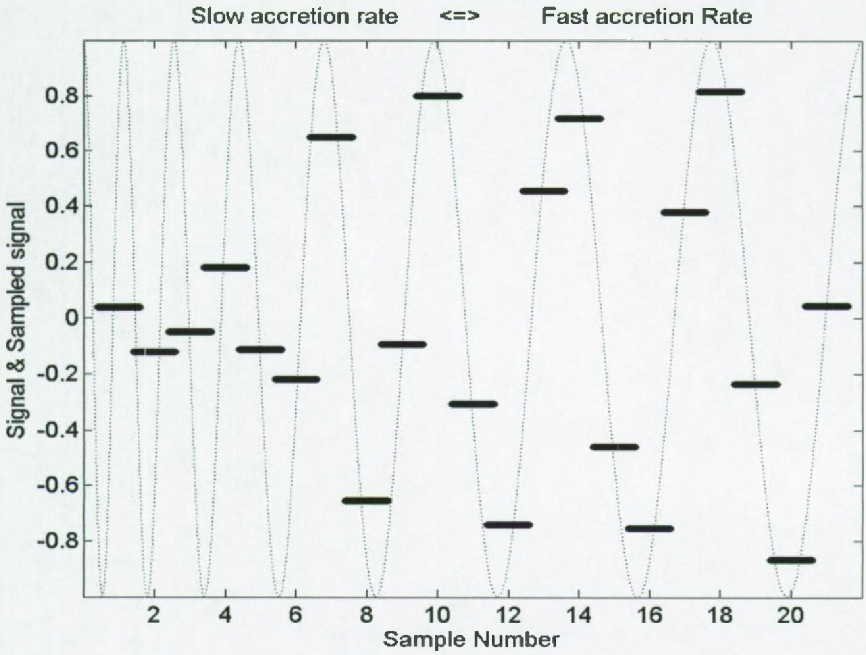

Figure 1: Simulation of sampling and averaging. The sines represents the 'true' signal, while the horizontal lines represents the measured signal.

$$\tilde{S}(t_n, \theta_{\text{signal}}) = \int_{x(1)}^{x(2)} S(t_n, \theta_{\text{signal}}) dx . \tag{9-1}$$

where θ_{signal} are the unknown signal parameters, $x(1)$ and $x(2)$ are the borders of the distance interval and S is the actual signal model. In order to solve this integral, t_n in the right hand side must be expressed as function of the distance, x, which can be done by a mapping function, $t_n = f(x(n), \theta_{\text{TBD}})$, which is function of the time base distortion parameters, θ_{TBD}. Inserting this mapping function into equation (9-1) shows that the integrated signal model, $\tilde{S} = \tilde{S}(t_n, \theta_{\text{signal}}, \theta_{\text{TBD}})$, is now a function of the time base distortion parameters, θ_{TBD} too. To conclude, we have derived the accretion rate from a signal model in this project. However, the signal is changed too, according to the accretion rate. Both can thus not be solved separately, but have to be solved in an iterative way, where the interaction between both is taken into account.

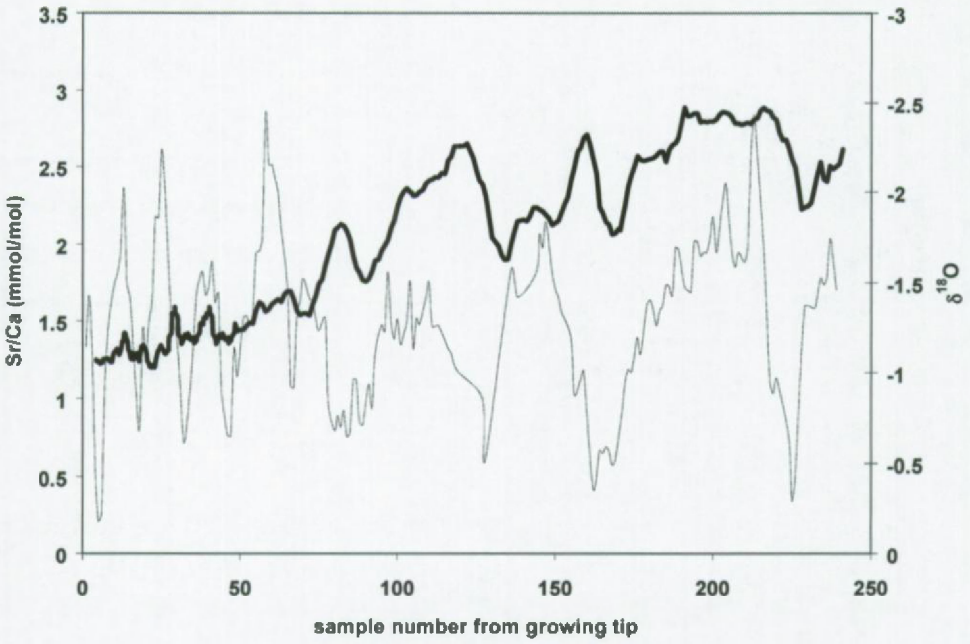

Figure 2: Raw data of Sr (thick line) and of $\delta^{18}O$ (thin line) measured in a shell (reprinted from David Gillikin *et al.* [submitted]).

To illustrate that this time averaging effect can have an influence on the measured signal, the Sr-signal in one of the *Saxidomus giganteus* clams from Washington state is shown in Figure 2. The accretion rate is visualized in the $\delta^{18}O$-signal (accretion rate increases from left to right). For the Sr-signal, which is measured on a different instrument, the signal seems to vanish if the accretion rate is too low.

This time averaging effect is extremely important when the same proxy measured in different species is compared, like e.g. the $\delta^{18}O$-signal measured in a shell and a coral, because these different species have a different accretion rate and the measured amplitude of the signal can differ significantly. Another important aspect is the transfer function, which relates the measured $\delta^{18}O$-signal to environmental parameters like temperature and salinity. If this transfer function is not calibrated for the accretion rate, it will only be valid for species which have the same accretion rate and are sampled with the same sampling strategy.

9.2.5 Multiple signal models within the same record

So far, we have focussed on single proxy records, where one time base and one signal model had to be optimized. However, in many records multiple proxies have been used, e.g. Vostok ice core record. For each of these proxies a signal model can be proposed and because they are all measured on the same substrate, the time base for all proxies must be the same. By combining multiple proxies, potential error, due to e.g. outliers, are minimized. Another potentially interesting combination consists of e.g. a periodic $\delta^{18}O$-record and radioactive dating measurements, where the calibration curve will be used as signal model too.

9.2.6 Multiple time bases and one signal model

Take for instance the three shells from Washington state. The $\delta^{18}O$-signals should have recorded the same environmental signal. These proxy-records can differ in two ways: the noise realization in each record differs and the accretion rate in each shell differs. If we would like to estimate which environmental signal is exactly recorded, a good estimate could be extracted simultaneously from the three shells. However, the time base for each shell should be estimated independently from the others. An alternative interesting combination are the temperature records, estimated from different ice cores. The time bases of the different records can be related by the signal model, they have in common. In addition, stochastic errors in the reconstructed temperature are averaged out.

9.2.7 The introduction of a time base distortion in time series

A first generalization of the time base distortion approach can be applied on time series analysis via AR-MA models [Broersen *et al.*, 2003]. These AR-MA can nowadays handle non-equidistant data. However, according to our knowledge, no efforts have been made to estimate this non-equidistancy. An AR-MA model is a generalization of the periodic models we have used so far. Therefore, we would like to try to implement the time base distortion in such an AR-MA model. Therefore, the dynamic behavior of the time series can be exploited to optimize the time base distortion. However, we will need some reasonably good idea about the AR-MA model parameters in order to initialize the time base distortion. In a next step, the model parameters can be improved. Iteratively, the best combination can be

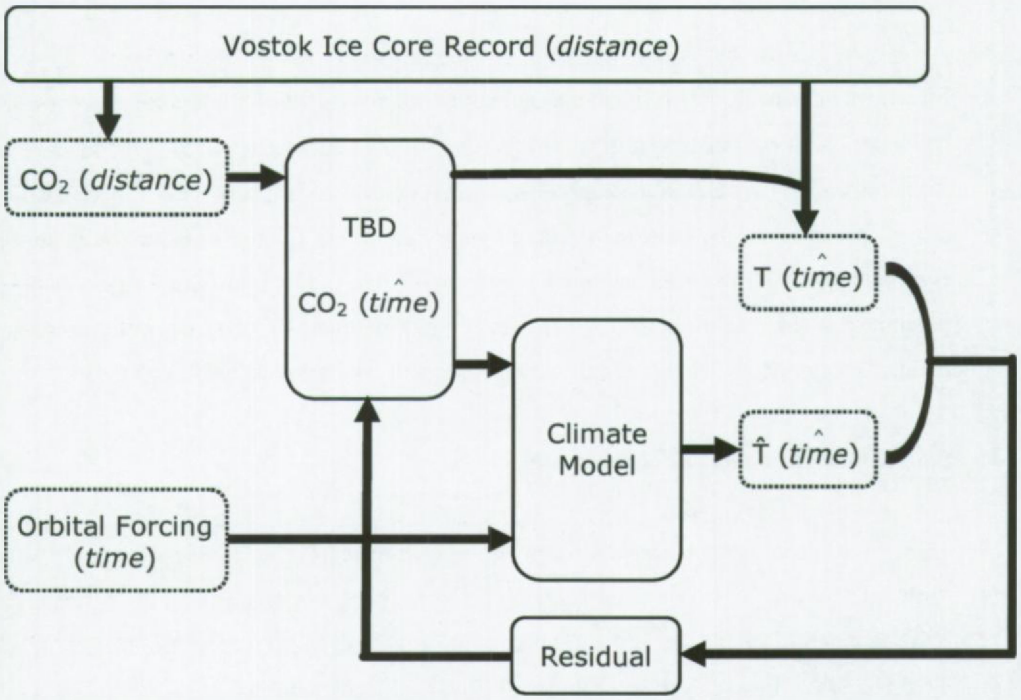

Figure 3: Structure which combines a time base distortion approach with a physical model: the Orbital forcing factor can be calculated directly, while carbon dioxide is measured in the ice core. First an arbitrary initial time base is used. These two forcing factors are used to estimate a temperature record, which can be compared with the measured temperature. The residual between both can be used to improve the estimate of the time base distortion. Iterating this process optimizes the time base.

reached. This approach can handle the optimization of multiple periodicities in the presence of colored noise, where the strategy followed so far will end up in too long computation times.

9.2.8 The introduction of a time base distortion in physical models

So far, we have described the signals by simple periodic signal models. However, it is also interesting to apply this time base approach to physical models. Therefore, we would like to start with relatively simple physical models for the temperature in Vostok (see Figure 3). One of the forcing factors in such a model is the carbon dioxide, which is also measured in the

Vostok record. It is now important to realize that the temperature responds dynamically to the carbon dioxide forcing. This dynamic behavior can be exploited to calibrate the time base. In a first approximation, it can be assumed that the physical model parameters are known. The time base distortion parameters can then be estimated in the same manner as described in this work. The main problem will consist of the initialization of the time base distortion parameters. Therefore, different time bases and different models for the distortion can be used or/and the optimization strategy can be adapted. More particular, the dynamical programming approach [Lisiecki and Lisiecki, 2002], can be useful, because of its speed and because it can avoid local minima.

Next, several pathways open:

1. the parameters of the physical model can be estimated together with the time base distortion parameters;
2. more complex models can be used to further refine the estimated time base for the Vostok record; and/or
3. the time base of other records, e.g. GRIP records, can be optimized by the same model. Here an additional parameter has to be estimated, i.e. the possible non-zero phase shift between the two proxy records. Although we are not dating the records in an absolute sense, both proxy records can then be compared on a coupled time base. So, if we know the date of one observation on one of the records, both time series or defined. If this algorithm is operational, it can easily be extended to more proxy records, like sediment records.

On the other hand, it is possible to give feedback to the modeler, if his model is not capable to describe the significant variations in both records. The philosophical idea behind this approach is that the time base cannot be separated from the model, that describes the paleo-record. Both must be optimized simultaneously. The missing link between both is the time base distortion.

REFERENCES

[1] Broersen P.M.T., de Waele S., Bos R., 2003. Estimation of Autoregressive Spectra with randomly missing data. IMTC 2003 - Instrumentation and Measurement Technology Conference, Vail, CO, USA, 20-22 May 2003, pp. 1154-1159.

[2] Gillikin, D. P., Ulens, H., Dehairs, F., Baeyens, W., Navez, J., Andre, L., Keppens, E., and the Calmars Group, 2004. Sr and Mg Profiles in Aragonitic Bivalves: Do They Record Temperature? Eos Trans. AGU, 84(52), Ocean Sci. Meet. Portland, OR, Suppl., Abstract OS42B-06.

[3] Lisiecki L. E. and Lisiecki P., 2002. Application of dynamical programming to the correlation of paleoclimate records. Paleoceanography 17(D4), 1049.

[4] Schoukens J., Pintelon R. and Vandersteen G., 1997. A Sinewave Fitting procedure for Characterizing Data Acquisition Channels in the Presence of Time Base Distortions and Time Jitter. IEEE Transactions on Instrumentation and Measurement 40 (4), 1005-1010.

GLOSSARY

All notations and symbols are locally defined in the different chapters. The following tables show general rules that were followed for the notation in this book. The symbols used in chapter 2 are listed separately from the others.

Symbols	Name	Dimension	Description
α, $\alpha(m)$	scale	counts per second	measure for the intensity of an internal standard
σ	uncertainty	-	measure for the stochastic uncertainty
σ_p	standard deviation on the drift pattern, p	scalar	measure for the uncertainty on the drift pattern
$\sigma_1, \sigma_2, \ldots, \sigma_m$	singular values	scalar	distribution of the variation
Σ	matrix with singular values	scalar	distribution of the variation
C	covariance matrix	(counts per second)^2	-
$e(n, m)$	error	counts per second	measure for the mismatch between model and measurement
$H(\alpha)$	scale factor matrix	counts per second	matrix containing the elements of α
i	label of isotope measured	scalar	$i \in \{1, \ldots, P\}$
I, $I(n, m)$	intensity	counts per second	raw signal measured by LA-ICP-MS
I_N, I_{NM}	unity matrix	scalar	-
K	cost function	scalar	measure for the residual variation

$K_{nwls}(\alpha)$	non-linear weighted least squares cost function	scalar	measure for the mismatch between model and measurement
m	number of internal standard	scalar	$m \in \{1, \ldots, M\}$
M	total number of internal standards	scalar	-
n	measurement number	scalar	$n \in \{1, \ldots, N\}$
N	total number of measurements	scalar	-
p, $p(n)$	drift pattern	scalar	measure for the drift pattern
P, $P(\alpha)$	scaling factor matrix	scalar	-
$Q(p)$	drift pattern matrix	scalar	-
S	intensity	counts per second	contains the columns of I stored on top of each other
S_w	weighted intensity vector	scalar	-
U, V	orthogonal matrices	scalar	-

Glossary

Symbols used in the remaining chapters.

Symbols	Name	Dimension	Description
β	linear coefficient in T.B.D.	scalar	-
θ	vector with model parameters	-	-
$\phi(m)$	phase	scalar	
σ	standard deviation	dimension of the respective record	-
ω	radial frequency	time^{-1}	parameters used in the signal model
A	harmonic amplitudes	dimension of the respective record	parameter used in the signal model
b	number of orthogonal bases	scalar	measure for the complexity of the time base
B	coefficients of T.B.D. model	scalar	coordinate of the T.B.D. w.r.t. the ortogonal basis
$e(\theta, Z)$	error vector of the (weighted) residuals	dimension of the respective record	-
$g(n)$	time base distortion	scalar	distortion of the time base
h	number of harmonics	scalar	measure for the complexity of the signal model
k	harmonic number	scalar	$k \in \{1, \ldots, h\}$
K	cost function	(dimension of the respective record)^2	measure for the variation not covered by the model
l	orthogonal basis number	scalar	$l \in \{1, \ldots, b\}$
m	periodicity number	scalar	$m \in \{1, \ldots, M\}$

M	number of periodicities	scalar	measure for the complexity of the signal
n	observation number	scalar	$n \in \{1, \ldots, N\}$
n_θ	dimension of θ	scalar	-
N	number of observations	scalar	measure for the length of the record
$p(n_\theta, N)$	penalty term	scalar	classical penalty term
$p_c(n_\theta, N)$	penalty term	scalar	corrected penalty term
$p_{\text{cic}}(R, N)$	penalty term	scalar	penalty term of the combined information criterion
$p_s(n_\theta, N)$	penalty term	scalar	penalty term corrected for Small number of observations
q	virtual border size	scalar	number of virtual observations at each border
$s(t_n)$	signal	dependent on the measurement	-
s_m	signal of the mth periodicity	dependent on the measurement	-
t_n	time instances	time	unknown discrete time instances of observations
T_s	sample period	time	time instance between two subsequent samples
Z	observations	dimension of the respective record	-