Breakpoint detection within the time series: modeling approach upon paleoclimatic proxy data

Lüder, Johannes; Brauer, Achim; Jurisch, Ronald

Breakpoint Detection within the Time Series. Modeling Approach Upon Paleoclimatic Proxy Data

Johannes Lüder, Achim Brauer & Ronald Jurisch *

Abstract: *»Bruchpunktdetektion im Prozess der Zeitreihenmodellierung anhand paläoklimatologischer Proxydaten«*. A large portion of research in time series analysis addresses questioning specific components like trend, cycle or seasonal behavior. Although there is a vast number of publications, only a small amount focuses on the research of irregularities, which are supposed to be within time series – especially in long-term data. Thus this paper focuses on detection of those irregularities and to illustrate the importance of breakpoint estimation. The underlying research theme is given by the discipline of Paleoclimatology. The investigation has been realized upon varved lake sediments as one of common proxydata in paleoclimatics. As the discipline provides insights in climate variability, questioning climate changes implies crucial information about mechanisms of rapid climate shifts. The paper shall also outline the importance of such information, since conducting time series modeling without interdisciplinary research constitutes an almost impossible task. Consequently time series analysis turns out in a procedure of modeling supposed components – like trend, cycles or irregularities – of the underlying datagenerating process and to image those in an appropriate degree.

Keywords: model and filterapproach, breakpoint, climate change, varved sediments.

1. Introduction

The researches in estimating breakpoints leads to mechanisms that will to define a suitable approximation on condition of strict assumptions of the datagenerating process.

Arguing characteristics of time series represents an unavoidable task to estimate time series components, however it does imply the necessity of a suitable modeling or filtering definition. Thus taking advance of a specific estima-

* Address all communications to: Johannes Lüder, Anhalt University of Applied Sciences, Department 3 – Architecture, Facility Management and Geoinformation, Schwabestraße 3, 06846 Dessau, Germany; e-mail: johannes.lueder@googlemail.com.
Achim Brauer, GFZ German Research Centre for Geosciences, Section 5.2 – Climate Dynamics and Landscape Evolution, Telegrafenberg C, 14473 Potsdam, Germany; e-mail: brau@gfz-potsdam.de.
Ronald Jurisch, Anhalt University of Applied Sciences, Department 3 – Architecture, Facility Management and Geoinformation, Schwabestraße 3, 06846 Dessau, Germany; e-mail: r.jurisch@afg.hs-anhalt.de.

tion, without questioning the underlying datagenerating process, can easily lead to misspecification of time series components – therefore irregularities like break points as well.

However, which kind of estimation is "suitable" or "appropriate" for a time series? A specific model or filter approach defines different assumptions and results in different estimations of the time series components. To find a suitable model, the datagenerating process needs to be described in an interdisciplinary research and define supposed components within the series. A suitable model or filter shall be able to consider those assumed components. Likewise even if a model has been chosen, this is supposed to be suitable, there is no guarantee, that the "true" model has been found (Metz 2010, 1056). Considering the estimation of breakpoints, the task gets more complex.

Without claiming breakpoints, the estimation of a model is impaired. A solution constitutes an iterative approach, which estimates models as well as breakpoints in a mutual procedure. Still, even this strategy cannot safely claim the suitable breakpoints of a specific series, since it does already presume an appropriate model. Therefore if the model does not consider all components, the estimation of breakpoints can also be impaired.

In order to give a possible solution to specify suitable breakpoints for paleoclimatic proxydata, the article is organized as follows: Section 2 outlines the problem of estimating components – including irregularities. Section 3 delves into the most important input – the characteristics of the varved sediments of Rehwiese Berlin and their supposed underlying components – focusing on the discussion of breakpoints and climate change. Section 4 illustrates the modeling and estimation strategy to detect breakpoints, suited upon the database. Section 5 leads to a discussion, questioning if a suitable estimation has been found – based upon interpretable content.

For further investigation in time series modeling (Metz, 2002; 2009; 2011) has provided comprehensive literature, occupying with breakpoints, long waves and trends in the time series modeling or filtering approach. Apart from arguing break points, his research outlines the mutual characteristics of estimated components – including irregularities.

2. Presentation of the Problem

Regarding the modeling procedure, it is necessary to point out the most common problems within the time series analysis approach. Tools to have insights in the datagenerating process are the descriptive statistics, which includes scatter plots, statistical parameters or transformations, the autocorrelation function, which provide insights in serial dependencies, or the periodogram, allowing the inspection of the frequency domain of a time series.

However those tools react significantly upon a form of non stationary process – the trend of a time series. The trend superimposes cyclic material, prevents insights in serial dependencies and is still one of the most common problems in time series analysis. Thus specific trend estimations result in different components because of their mutual influence. Additionally since the estimation of breakpoints presume a suitable model, an appropriate trend for a series is crucial.

In the result estimating breakpoints presume accomplishing a comprehensive discussion of the components of the series, their identification and finding modeling or filtering approaches, which are able to regard all argued and interpretable components.

However, why is the estimation of breakpoints that important? Finding breakpoints does also mean to change the trend definition. They are understood as events, where the data generating process may change completely. Ignoring those points lead to completely different results.

Additionally those points can be understood as one of the crucial components of a series. Thus in varved lake sediments, as an indicator of climate behavior, questioning breakpoints means questioning climate changes. In the result, estimating breakpoints lead to a more suitable model of a specific series and to an increase in the discussion and understanding of the investigated data.

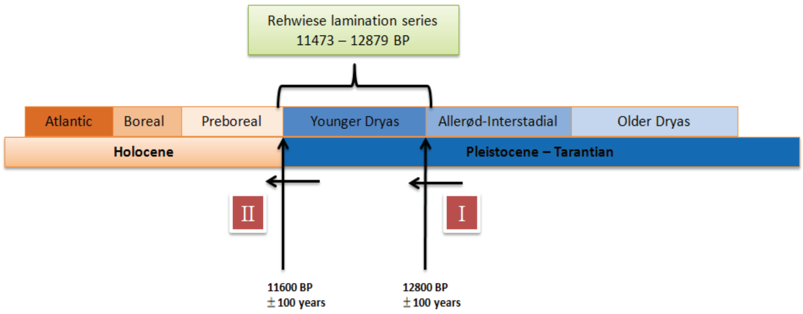
3. Strategy Modeling Varved Lake Sediments – Rehwiese Berlin

Before the actual modeling and filtering strategy, assumptions about components of the series have to be formulated. Therefore the investigated data need to be introduced. However the discussion will be shortened to set up the focus on estimating breakpoints – questioning climate change.

Varve Chronology Rehwiese Berlin

The investigated proxydata – the lake sediments of Rehwiese Berlin – encompass the climate phases "Allerød", "Younger Dryas" and the "Preboreal". The complete chronology covers 1407 varves and has been time referenced on 12837 until 11473 BP. As shown in the following image, the transitions between Allerød – Younger Dryas and Younger Dryas – Preboreal are dated on approximately 12800 BP $\pm$ 100 years and 11600 BP $\pm$ 100 years. Therefore it can be assumed that the present proxydata have imaged those changes. The research focuses on, how these transitions can be characterized and how their influences take effect within the proxydata.

Figure 1: Temporal Classification



Two laminations – the organic and second calcite lamination of the envisaged varve chronology have been investigated. However, only the breakpoint estimation of the SoO lamination will be illustrated. The research is based upon which kind of climate effects they are sensitive to and how they may change due to climate changes. However this information is not apparent, especially since the laminations used to be in seasonal dependency. Therefore changes within a specific lamination might easily cause changes in others as well. Nevertheless this kind of property might help to differentiate between single climate factors, which have been changed crucial and in result signalize changes in every lamination. Still it cannot be ruled out, that single factors, like temperature, may change in a degree that shows changes only in specific laminations and leaves others more or less unchanged. The following description, assumptions and disposition have been made on the basis of the diploma thesis of Neugebauer (Neugebauer 2010).

According to Neugebauer the chronology can be structured into five segments. The following table and the figure illustrate the ageing between every single part. Additionally the argumentation for this ageing has been made on the basis of specific characteristics at those transitions. These characteristics help to figure out changes, factors or influences, which partially can be referenced on single laminations.

Transition between segments I and II:

- Rising merge laminations
- Crossover from relative humid, windless character to the dry and cold Younger Dryas phase
- Descending temperature
- Descending lamination rates

Transition between segments II and III:

- Reduction of the entire lamination (III)
- Reduction of nutrient supply

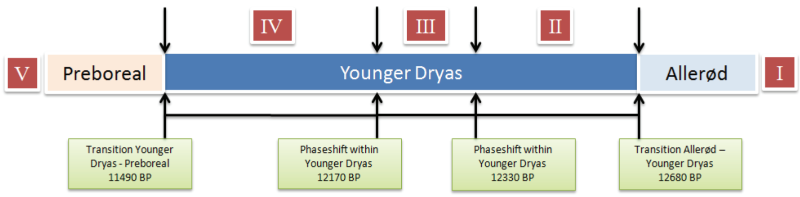
Transition between segments III and IV:
- Assumption about a trendbreak towards segment IV
- Increase of windy character – results in rising wave activity, which entail a stronger rearrangement and raised organic and merge laminations
- Crossover to even colder and dryer character
- Segment IV has been appointed to be very instable – speaking about stormy character, which may cause wind induced sluicing of additional material from the shore
- Descending calcite laminations, regular occurrence assumed (may be caused by solar cycles)

Transition between segments IV and V:
- Raising temperatures as well as calcite laminations
- Decline of merged laminations
- Lower lamination fluctuations
- Reduction of the instable character, less storms, less windy character

Table 1: Argued Climate Phases

Segment	Chronozone	Years in BP	Duration (years)
I	Allerød	12880-12680	200
II	Younger Dryas	12680-12330	350
III	Younger Dryas	12330-12170	160
IV	Younger Dryas	12170-11590	580
V	Preboreal	11590-11490	100

Figure 2: Climate Phases in Context of Varve Chronology



Upon this argumentation, the paleoclimatology indicates five possible shifts within the trend level of this lamination. Besides the assumption of irregularities, the following components need to be considered for an appropriate estimation and are supposed to be in these sediments:
- Trend as a long term component, representing the systematic change within the average level of the series
- Cyclic material, superimposed by the trend
    - 11 (9-14) years – Schwabe cycle
    - 22 years – Hale cycle
    - 44 years – (42-50) Schove cycle
    - 80-90 years – Gleissberg cycle

    - 180-210 – Suess cycle
    - Long wave cycles
- Single outlier

When these components are present in the series, they are crucial for the estimation of breakpoints and therefore need to be considered in the modeling process. However, cyclic or trend behavior cannot be estimated without having regard to these irregularities. Through this standoff it is necessary to discuss the announced shifts within the time series. Considering the climate as the essential element of the datagenerating process, the breakpoints are understood as climate changes and therefore need to be characterized.

Burroughs (Burroughs 2007) has found a suitable description:

> […], that changes in the climate constitute shifts in meteorological conditions lasting a few years or longer. These changes may involve a single parameter, such as temperature or rainfall, but usually accompany more general shifts in weather patterns that might result in a shift to, say, colder, wetter, cloudier and windier conditions.

Following this definition a significant breakpoint implies changes in several factors and therefore shows changes in every lamination, caused by seasonal dependency or sensitive behavior through a climate parameter. In addition Burroughs also announces:

> Due to the connection with global weather patterns these changes can result in compensating shifts in different parts of the world. More often they are, however, part of an overall warming or cooling of the global climate, but in terms of considering the implications of changes in the climate, it is the regional variations that provide the most interesting material, as long as they are properly set in the context of global change.

Following this assumption, differentiate spatial distinctions result in the possibility that supposed or argued shifts may not be present in the investigated time series. Thus a significant climate change may prop an already argued shift and therefore extend the spatial boundaries of a "global change". Finally Burroughs introduces the difference between climate variability and climate change: "This leads into the question of defining the difference between climate variability and climate change. Given that we will be considering a continuum of variations across the timescale from a few years to a billion years, […]".

In result, it is crucial to decide if a fluctuation of the series is interpretable as part of the long-term trend and therefore representing a breakpoint, or a short-term fluctuation, which may have been caused by climate variability. In fact, this possibility opens up the opportunity to deny an argued breakpoint, due to short term fluctuations. However the introduced breakpoints represent the most advisable possibilities of climate change, but need to be reviewed under investigation of additional influencing variables.

As a last input for detecting breakpoints, the descriptive analysis is able to provide hints about long-term fluctuations. However the resulting breakpoints will only be listed after they have already been compared to those, which were argued from paleoclimatic research:

SoO – summer organic lamination

- 12170 $\pm$ 50 – Transition Younger Dryas Segment III – IV, rising stormy character, instable
- 11700 BP $\pm$ 100 – Transition Younger Dryas – Preboreal, decreasing wind-induced intake of organic material

Still these assumed breakpoints have been found in the investigation of descriptive statistics and paleoclimatic input and do not need to be significant in terms of a low frequency, long – term trend comprehension. They represent the best interpretable solutions, how breakpoints may have influenced the level of the series. In the result, if the estimation process detects breakpoints, which on the one hand are interpretable in terms of paleoclimatics and on the other hand do not show contradictions in statistical findings, they represent the likeliest solution.

4. Breakpoint Estimation Strategy

This chapter focuses on a possible solution about estimating the trend, possible cycles and the discussion about resulting breakpoints. The model estimation has been made with structured time series models, which represent a special case of ARIMA models and facilitate the simultaneous estimation of time series components. This model has been chosen to consider multiple cycles as well as an estimation of "volatile" trends.

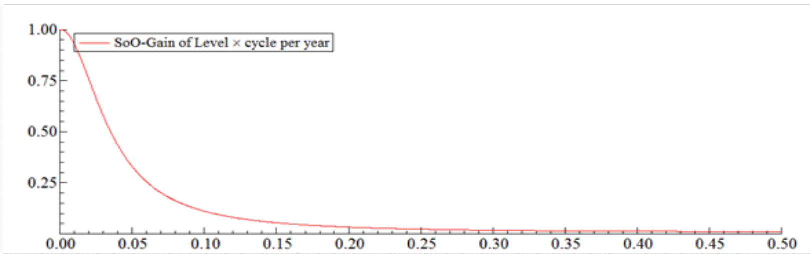
Volatile trends in structured time series models regard a trend evolving as a random walk having a slope also evolving as a random walk. The trend model is defined as follows:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \qquad \eta_t \sim NID(0,\ \sigma^2_\eta)$$
$$\beta_t = \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim NID(0,\ \sigma^2_\zeta)$$

In addition the disturbance variables η_t and ζ_t are modeled as independent white noise processes. However since the researcher does not know the significant components, the formulated model will not represent the actual datagenerating process. Therefore the model is not expected to estimate the significant breakpoints.

However the trend, being volatile, considers more short-term fluctuations within the estimation procedure (illustrated in the following picture). In terms of the frequency domain the trend model cycles and estimates breakpoints due to cycle fluctuations or climate variability, especially if separate cycles superimpose each other. A volatile trend model estimates a variety of breakpoints and tries to not distinguish between climate change and climate variability.

Figure 3: Gain Function of Volatile Trend Estimation



To validate the results – to decide whether a breakpoint has been estimated due to climate variability, cycles or climate change, the Cristiano-Fitzgerald (CF) filter has been chosen to investigate growth rates, frequencies and breakpoints. Therefore it is necessary to decide whether an estimation is suitable for the series. Those breakpoints are most likely that are assumable in paleoclimatic research, do not show any regularity behavior and several time series components change their structure at a specific point.

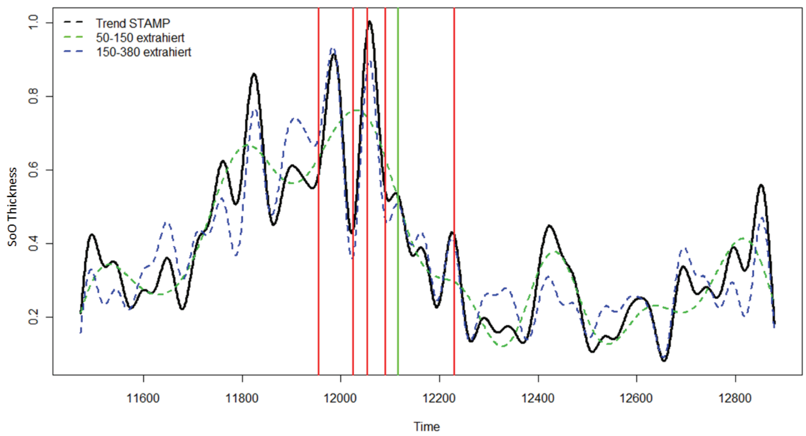
In result, the SoO lamination volatile trend estimation contains sixteen breakpoints. The most common climate changes have been mentioned, based upon descriptive analysis and paleoclimatic interpretations. However this does not lead to an exclusion of the remaining breakpoints. To investigate both series the CF filter has been formulated in several trend variations to distinguish whether the slope of every filtering approach does contain cyclic behavior.

SoO

- Trend > 60 years, frequency 0,01667
- Trend > 120 years, frequency 0,08333
- Trend > 225 years, frequency 0,00444
- Trend > 350 years, frequency 0,00285714

These filter definitions provide insights in possible cyclic movements and these cycles have been part of the volatile trend, at least partly. Therefore cycles might have caused the estimation of the variety breakpoints. That's why after investigating the slopes, the behavior has been formulated with the CF filter to have an insight in breakpoints in consideration of cyclic behavior. The results for the SoO lamination are shown in the following picture.

Figure 4: Extraction of Cyclic Behavior with Two CF-Pass Filter

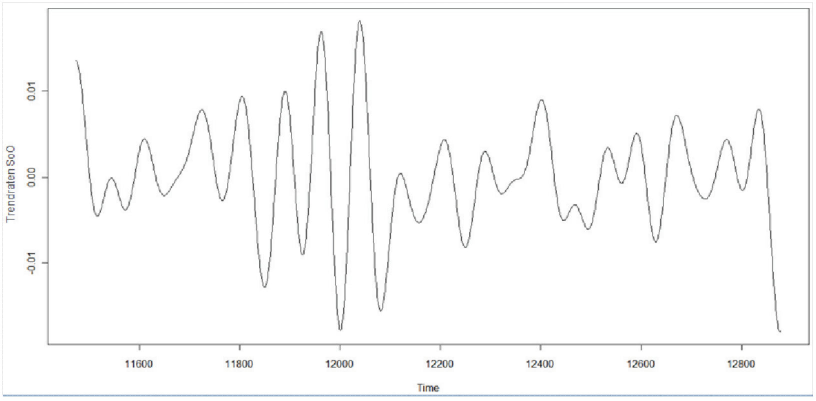


Regarding these information, most of the breakpoints have been estimated due to cyclic behavior and have been rejected as being significant. To improve the breakpoint estimation, cycles were added to the model definition and the procedure was repeated. The amount of possible breakpoints has been reduced significantly to four. Although, the model is still supposed to not consider all significant components of the datagenerating process, the remaining points can be distinguished with paleoclimatic argumentation and the information of frequency extraction.

In the result, one breakpoint dated at 12116 BP has been accepted as a significant change for the SoO lamination. It is time referenced upon the transition between Segment III and IV and does show increased thickness values in the following years. The change has been characterized as becoming stormier, windier and does show increasing instability. In addition, as breakpoints are understood as significant changes, they have also shown differences in cyclic patterns. An example, how those patterns look like, the following picture shows the growth rates of the SoO lamination for trend estimation with a frequency of > 1/60.

However the transition between segment IV and V is also debatable, but was not estimated neither within the volatile trend approximation nor a model regarding an additional cyclic component. Even if the transition is characterized as becoming less instable, the time series seems to evolve in a long-term behavior. Therefore declining an additional breakpoint, as the climate trend character is a long-lasting component, is practicable.

Figure 5: Trendrates of SoO Lamination



5. Continuous Procedure of the Time Series Modeling Approach

As breakpoint estimation has been considered and implemented, the modeling approach has its statistical and paleoclimatic input, so it can be realized. Therefore the modeling has been repeated with STAMP regarding the breakpoint at 12116 BP. Its influence is estimated and the time series becomes adjusted, regarding the estimated power. Within the assumed boundaries of the argumentation, the time series components are not damaged or injured by additional breakpoints. Therefore modeling as well as filter procedures are used to work as expected.

However, since the argued breakpoint is only one of a variety of possible solutions, the estimation procedure may show contradictions, which lead to a repetition of the identification procedure and to a readjustment of the modeling components.

6. Conclusion

The paper is intended to outline the necessity regarding breakpoints for time series modeling or filtering approaches. Basically it is part of a comprehensive time series analysis and therefore the illustrated procedure shows a huge and important base of the actual model approach, but does not show the results of estimating trend and cyclic behavior as well as its validation. In addition the strategy using structured time series models and the Cristinao-Fitzgerald filter is just one of many possible solutions of estimating breakpoints. In addition, it is not necessarily the best solution at all, but has been found in consideration of the data. In result, the estimation procedure does always need to be chosen in regard to a comprehensive investigation of the data.

However ignoring possible breakpoints can easily result in misspecifications and impair the constitutive analysis, like the time series procedure itself or statistical methods like regression, correlation or even multivariate time series modeling. In paleoclimatic data, it has been shown that estimating breakpoints can increase the actual understanding of the data and introduce another possibility of investigation. In addition, the procedure has shown that climate changes in climate proxy need to be judged in its sensitive behavior in consideration of differences in a topical (specific proxy), temporal, spatial and methodical way.

References

Burroughs, W. J. 2007. *Climate Change A Multidisciplinary Approach.* Cambridge: University Press.

Metz, R. 2002. *Trend, Zyklus und Zufall – Bestimmungsgründe und Verlaufsformen langfristiger Wachstumsschwankungen.* Köln: Franz Steiner Verlag.

Metz, R. 2009. Filter-design and model-based analysis of trends and cycles in the presence of outliers and structural breaks. *Cliometrica* 4 (1): 51-73.

Metz, R. 2010. Zeitreihenanalyse. In *Handbuch der sozialwissenschaftlichen Datenanalyse*, ed. Christof Wolf and Henning Best, 1053-90. Wiesbaden: VS Verlag für Sozialwissenschaften.

Metz, R. 2011. Do Kondratieff waves exist? How time series techniques can help to solve the problem. *Cliometrica* 5 (3): 205-38.

Neugebauer, I. 2010. *Mikrofazielle und geochemische Untersuchungen laminierter Seeablagerungen des Spätglazials der Rehwiese Berlin*: Diploma Thesis.