DFG-Abschlußbericht im Schwerpunktprogramm 1097

„ERDMAGNETISCHE VARIATIONEN: RAUM-ZEITLICHE VARIATIONEN, PROZESSE UND WIRKUNGEN AUF DAS SYSTEM ERDE“

“GEOMAGNETIC VARIATIONS: SPATIO-TEMPORAL VARIATIONS, PROCESSES AND IMPACTS ON THE SYSTEM EARTH”

Short title:

Nonlinear Phase and Correlation Analysis of Palaeomagnetic and Palaeoclimatic Records

(Az.: KU 837/15-2)

May 2003

Jürgen Kurths & Udo Schwarz
Interdisziplinäres Zentrum für Nichtlineare Dynamik
Universität Potsdam
Am Neuen Palais
14469 Potsdam

1 Aim

The aim of this project was to apply and develop modern techniques of nonlinear data analysis to study complex dynamical relationships between the variations of the Earth's magnetic field and of the climate during the last 100 000 years. We mainly intended to study phase stability and nonlinear correlations and dependences of these multivariate data based on new approaches of synchronization analysis, cross recurrence plots, nonlinear correlation and graphical models. Problems of special interests were to test for a possible phase delay between orbital/ climate signals and the relative palaeointensity, to study the influence of long-term variations, especially in the sub-Milankovich range, and to analyse excursions of relative palaeointensity in detail. The nonlinear analysis of sediment data has been mainly based on the sediment investigations performed in the group of Prof. Negendank/ Dr. N. Nowaczyk (GFZ Potsdam).

2 Research Period

1. Period: April 1, 2000 – March 31, 2002
2. Period: May 1, 2002 – April 30, 2003

3 Evolution of the Work, Problems

The theoretical preparation of the research through a substantial development of the method of *cross recurrence plots (CRP)* has occupied the first stage of the project. The work with CRPs has opened broader insights and developments as initially expected. Therefore, the also planned method of *graphical models* was not concerend to apply.

In the second stage we have investigated the available palaeo- and rock-magnetic data sets. At the beginning of the project we have planned to investigate palaeo- and rock-magnetic measurements of lake-sediments (Italian lakes, data provided by the GFZ Potsdam) and marine sediments (data should be provided by University of Bremen). Unfortunately, until the end of the project the data of the marine sediments were not yet available. Therefore we have concentrated on the data from two the Italian lakes, which were provided by the group of Prof. Negendank (GFZ Potsdam).

The available data have shown a huge amount of instationarity, gaps and interrelations between themself. This has lead to a comprehensive pre-analysis of the data, which inludes an analysis of linear correlation and the nonlinear technique of maximal correlation.

The strong interrelations between all single data sets have required the application of the *independent component analysis (ICA)*. The ICA could be applied by using the FastICA toolbox from the Laboratory of Computer and Information Science of the Helsinki University of Technology. The ICA provided an adjusted data set of palaeo-intensity which does not include a climatic influence anymore. This data was used in a further CRP analysis for the search of linkages between the palaeo-intensity and climatic/ orbital forcings.

This neccessary data preparation and pre-analysis have occupied more time as scheduled. This extended work on pre-analysis techniques should be of much importance for various other analysis in the SPP 1097 and we have made them available in our toolbox (cf. Sec. 6). However, a further evaluation of the results could not be done during the period of support. This is a challenging problem for the future activities in this SPP.

4 Results

4.1 Methodical Developments

In the first stage of the research, we have mainly focused on the further development of the *cross recurrence plots (CRP)*. The CRP is a bivariate extension of *recurrence plots (RP)* and visualizes similar time evolution of two processes $\vec{x}_i$ and $\vec{y}_i$:

$$\mathbf{CR}_{i,j} = \Theta\left(\varepsilon - \left\|\vec{x}_i - \vec{y}_j\right\|\right), \quad \vec{x}_i, \vec{y}_j \in \mathbb{R}^m, \\ i = 1 \dots N_x,\ j = 1 \dots N_y. \tag{1}$$

At the beginning of the project, this method was a completely new approach, nothing was available in the literature and we have worked out the first applications. In this first stage we have introduced new measures of complexity, which can assess the similarity of two dynamical processes; these include the *recurrence rate*, the *determinism* and the *mean diagonal line length*. Furthermore we have studied the reliability of these measures. The applications to prototypical model systems have shown that this method is able to find linear interrelations as well as nonlinear interrelations between different systems (Marwan and Kurths, 2002). A first application to geological data could reveal a clear interrelation between a climatological process (ENSO) and a geological process (sedimentation in a lake in northwestern Argentina). Therein we have used the method of CRPs for the comparison of palaeoclimatic data and present-day climatic data (Marwan et al., 2003).

The methodical development includes also an improvement of RPs and its related quantification analysis. New measures of complexity were introduced, which especially enable to find chaos-chaos transitions, which is an important problem in nonlinear data analysis (Marwan et al., 2002b). Note that the classical quantification analysis of RPs is only able to find chaos-order transitions.

The further investigation of CRPs has yet revealed another interesting applicability. The local slope of lines in a CRP corresponds to the transformation of the time axes of the two considered time series. This feature can be used for the aligning of different data series which represent the same process but have underlied different time-scale transformations, as compession or stretching. This is, e.g., a problem in drilling campaigns, where different cores have to be aligned to the same time (or depth) scale (Marwan et al., 2002a).

This technique can also be used in order to find the closest matching segments in two data series. For example, in the geological framework there could be a long reference data series which has a time scale and a second but much shorter profile with the same geophysical measurement. The task lies in finding the section in the reference data which matches to the second profile in order to yield the corresponding time

Table 1: Data sequences used for the analysis.

Data	Time (yr)
LGMJ	19 300–100 050
LMZC1	208–3 974
LMZC2	14 500–30 820
LMZG1	776–3 689
LMZG2	15 160–29 800

scale for the profile. Such a section can be determined by using CRPs (examples can be found in Marwan, 2003).

4.2 Data Analysis of Palaeomagnetic Measurements

For the investigation of relationships between the variations of the Earth's magnetic field and climatic or orbital variations, only three data sets from two Italian lakes were available. One data set is gained from the lake *Lago Grande di Monticchio (LGM)* and covers about 20 000 years to 100 000 years. The other two data sets are from *Lago di Mezzano (LMZ)* and cover the last 30 000 years.

The data analysis of these data sets was rather problematic because the data contains numerous gaps and nonstationarities. Moreover, some measurements have a rather non-gaussian distribution. Moreover, the measurements of rock-magnetic properties contain a clear climatic signal. This fact does not allow to compare these measurements directly with other climatic data series, because a significant correlation between them would be obvious. The up to now used way to reduce this climatic influence in the data is to use a quotient of rock-magnetic measurements, as $\frac{NRM}{ARM}$.

To enhance the reliability in the cores of the lake LMZ, their data series were splitted into an upper and a lower section (Tab. 1). In the medium section the amount of magnetic minerals is too small, leading to a rather less reliability for palaeo-magnetic interpretations.

4.2.1 Preanalysis

In a first step, we have used the quotients $\frac{NRM}{ARM}$ and $\frac{NRM}{SIRM}$ as a proxy for the palaeo-intensity of the Erath's magnetic field and have applied a linear correlation test and a test based on the ACE algorithm (maximal correlation, cf. **?**).

As climatic and orbital variation proxies we have used pollen data (herbs and trees pollen) from the Lago Grande di Monticchio, ^{18}O isotopes data (Gisp2) and insolation data for the latitude 44°N (July and December insolation).

Linear analysis The data gained from the two locations of Lago Grande di Monticchio (LGM) and the Lago di Mezzano (LMZ) show different dependences between the climate and the rock magnetic properties.

In the LGM core, the influence of the climate (pollen proxy) cannot be eliminated for the quotients. The data series of the quotients are stronger correlated with the climate proxies as the original data.

In the LMZ data, the correlations do not vanish by the application of the quotients. The correlations can even be amplified (as in LMZG).

Maximal correlation The results gained by the application of the maximal correlation analysis are similar to that of the linear analysis. The correlation can be undesirably amplified by the computation of the quotient.

However, the maximal correlation is less and never exceeds 0.9, which means, that this correlation may be neglected. But the optimal transformations show a distinguishable shape, which means that there may be some interrelations between the data outside the direct climatic ones (cp. Fig. 1).

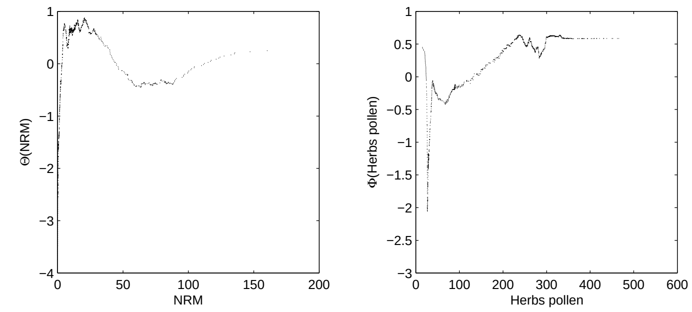

Figure 1: Optimal transformations of a rock-magnetic measurement (NRM) and a climate proxy (herbs pollen) obtained by using the ACE algorithm. The NRM data are from the LMZC core.

Bivariate AR-coefficient analysis The coefficients for bivariate auto-regressive (AR) models contain information about the strength of correlation between the two subsystems. With a stepwise, least squares estimation of multivariate AR-models these coefficients were estimated.

$$\begin{pmatrix} \vec{x}_i \\ \vec{y}_i \end{pmatrix} = \sum_{j=1}^{p} A_j \begin{pmatrix} \vec{x}_{i-j} \\ \vec{y}_{i-j} \end{pmatrix} + C \begin{pmatrix} \xi_i \\ \zeta_i \end{pmatrix}$$

where $A_j = \begin{pmatrix} a_{x,x} & a_{x,y} \\ a_{y,x} & a_{y,y} \end{pmatrix}_j$ are the AR-coefficients for the time lag j and ξ and ζ are Gaussian distributed white noise.

The analysis was applied on LGMJ and the lower sections of the cores LMZC and LMZG. The data were normalized and shifted above the zero-axis ($\min(x) > 0$) and logarithmized. With this data preparation the extreme non-gaussian distribution of the data can be transformed to a more Gaussian distribution. The coefficients were estimated for AR-processes of the maximal order of three.

The AR-coefficients and thus the correlation between the rock-magnetic and the climatic/ orbital data differ significantly between the cores (Tab. 2). This may be a sign for the rather high degree of nonstationarity of the rock-magnetic data.

The herbs pollen mostly are not correlated with the rockmagnetic parameters; however, if they are correlated (LGMJ and LMZG), then positive. Similar the trees pollen,

Table 2: Pairwise correlation – estimated from the AR-coefficients – between rock-magnetic measurements and climate/ orbital proxies, exemplary shown for the lake LMZC. 0 – no correlation, $+$ – correlation, $-$ – anti-correlation, boxed – strong correlation.

	NRM_{20}	ARM_{20}	$SIRM$	$\frac{NRM}{ARM}$	$\frac{NRM}{SIRM}$
Herbs	x	x	x	x	x
Trees	$-$	x	$\boxed{-}$	x^2	$-$
Gisp2	x	x	x	x	x
Inso7	$+$	$+$	$\boxed{+}$	$\boxed{-}$	$\boxed{+}$
Inso12	x	$-$	$-$	$\boxed{+}$	$-$

which are mostly not correlated; however, if they are correlated (LMZC), then negative. Surprisingly, the Gisp2 data are not correllated. The insolation for July is strongly correlated with all rockmagnetic parameters. The insolation for December shows more different corellation signs.

From these results we have inferred, that the single rock-magnetic parameters as well as their quotients contain too strong climatic signals and are not suitable as a proxy for the palaeo-intensity of the Earth's magnetic field. Therefore, any simple investigation of their interrelation with climatic or orbital variations is not possible. First of all, a signal which is exempt from the climatic influence is to reconstruct. We have chosen the ICA technique, which decomposes signals in independent components (in contrast to the PCA, which decomposes signals only in uncorrelated components).

4.2.2 ICA and CRP Analysis

ICA The task is to separate a signal from the measurements which contains only the intensity of the Earth's magnetic field. For this objective we have chosen these three measurements: the natural remanent magnetization (NRM), the anhysteretic remanent magnetization (ARM) and the susceptibility (κ). The natural remanent magnetization is the magnetization of the sample as it comes directly from the drilling. Therefore it contains a signal of the intensity of the Earth's magnetic field in the past. The other two measures are determined after demagnetization of the sample in the laboratory and, therefore, do not contain any information about the Earth's magnetic field. Each of the three measures depend on the grain size and the concentration of the magnetic minerals, but each depends on them in a different way. The anhysteretic remanent magnetization is impacted only by small magnetic minerals, whereas the susceptibility is impacted by only large magnetic minerals. Finally, the grain size and the concentration depend on the climate variation, therefore all three measures depend on the climate and correlate with some climate proxy data sets, as oxygene isotops or pollen data.

For the ICA we use the following model:

$$\begin{aligned} NRM &= f_1(F) + f_2(c) + f_3(s), \quad c, s = f(\mathrm{C}) \\ ARM &= g_1(c) + g_2(s_{small}) \\ \kappa &= h_1(c) + h_2(s_{large}) \end{aligned}$$

where NRM – natural remanent magnetization; ARM – anhysteretic remanent magnetization; κ – susceptibility; F – Earth' magnetic field; C – climate; c – concentration and s – grain size of magnetic minerals

The separation of the factors F, c and s with the ICA reveals three independent components (ICs) s_i ($\vec{x} = \mathbf{A}\vec{s}$); for example, for LGMJ the corresponding mixing matrix obtained from the ICA is:

$$\mathbf{A} = \begin{pmatrix} 16 & -4 & -12 \\ 205 & -897 & -931 \\ 16 & -36 & -136 \end{pmatrix}.$$

These ICs contain a magnetic field signal (s_1) and a climate signal (s_2 and s_3).

In order to check whether the ICs contain a climatic signal, we have applied correlation test between the independent components and some climate proxies[1]. The first component does not correlate with the climate signals, whereas the other two correlate and, hence, contain a climate signal (Tab. 3). The first IC contains much less climate impact as the usualy used quotients $\frac{NRM}{ARM}$ and $\frac{NRM}{SIRM}$ (Tab. 4). We have compared this independent component and the other measures with the reference data set SINT800 of the palaeo-intensity of the Earth magnetic field (Tab. 5).

Table 3: Correlation coefficients between ICs of LGMJ and the underlying signals as well as proxy data for the climate, reveals a clear distinction of these signals. Q – Quercus pollen; $CLIM$ – proxy for global temperature

	NRM	κ	ARM	$\frac{NRM}{\kappa}$	$\frac{NRM}{ARM}$	Q	$CLIM$
s_1	0.80	0.16	0.11	0.51	0.49	−0.07	0.02
s_2	−0.18	−0.69	−0.26	0.41	−0.03	0.19	0.15
s_3	−0.58	−0.71	−0.96	0.08	0.16	0.21	0.19

Table 4: Furthermore, the first IC s_1 of the LGMJ contains much less climate impact as the usually used ratios of NRM with ARM and κ, respectively. Q – Quercus, P – Pinus pollen; $CLIM$ – proxy for global temperature

	P	Q	$CLIM$
s_1	−0.03	−0.07	0.02
NRM/κ	−0.15	0.15	0.21
NRM/ARM	−0.09	0.06	0.10

The ICA was applied to the data of the three cores LGMJ, LMZC and LMZG and revealed for each core a signal, which contains the intensity signal of the Earth's magnetid field but does not include a climate signal. These results were used for a subsequent CRP analysis.

Furthermore, this analysis has shown that by using the ICA the signals can be separated into independent signals which are better proxy data sets for the Earth's magnetic field than the usually used proxies.

[1]Besides the pollen data we have used a ^{18}O isotopes data series which we have gained as an IC from the Gisp2 and Vostok ice cores, which would be a proxy for the global temperature.

Table 5: The comparison with the SINT800 reference data set shows also an improved magnetic field component obtained by the ICA.

	SINT800
s_1	0.21
NRM	0.19
NRM/κ	0.10
NRM/ARM	0.11

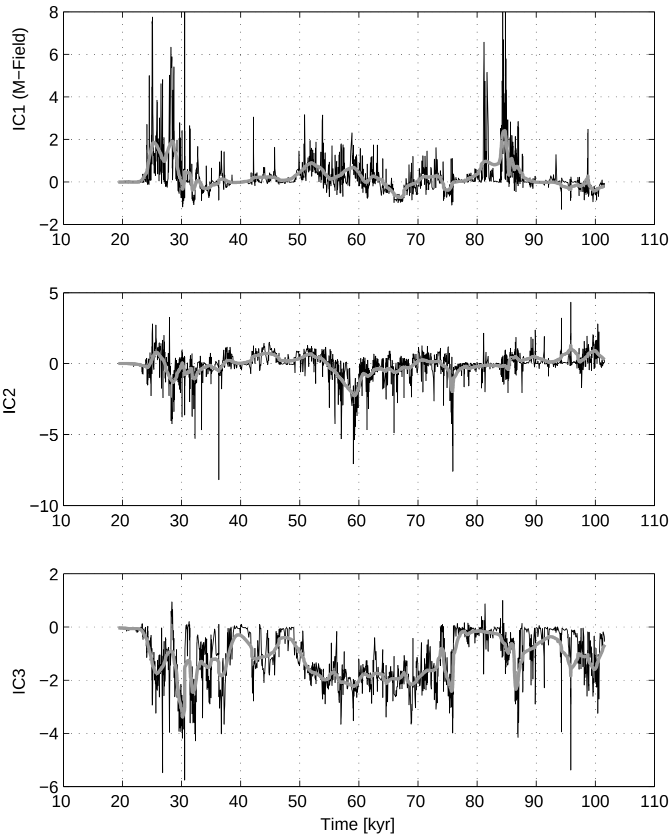

Figure 2: The independent components of the ICA of the LGMJ core (in gray the smoothed signal is shown).

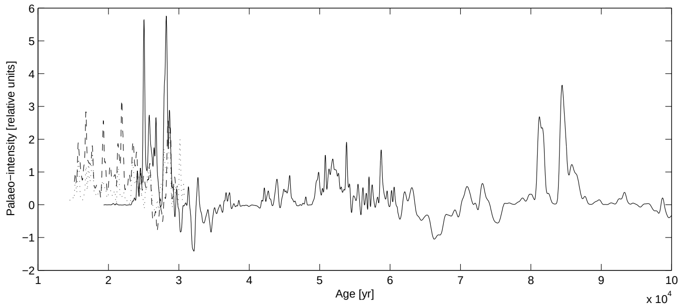

Figure 3: Proxies for the palaeo-intensity of the Earth's magnetic field gained from the ICA analysis (dotted – LMZC2, dashed – LMZG2, line – LGMJ).

CRP analysis For the CRP analysis we have used that IC from the ICA, which has the smallest correlation with climate proxies but the highest correlation with the NRM. We have analysed such palaeo-intensity proxies from the entire LGMJ core and from the lower sections of the LMZC and LMZG cores. We have to note, that linear correlation tests would fail for this analysis, because the data are rather nonstationary and non-gaussian.

For all data sets we have found an appropriate embedding of $m = 3$ and $\tau = 6$. The CRPs were computed by using a fixed amount of nearest neighbours (10% of all possible recurrences, for details see Marwan, 2003).

The quantitative analysis of the corresponding CRPs has revealed a slight interrelation between the intensity of the Earth's magnetic field and the climate, which were represented by pollen data. However, we could not find a clear influence of the global climate (^{18}O) on the Earth's magnetic field. Furthermore, a slight and delayed interrelation between the intensity and the inclination could be observed.

In more detail, the main results are (cp. Tab. 6):

- negative interrelation with Pinus pollen; positive and negative interrelation with herbs pollen; slight negative or none interrelation with Quercus pollen
- interrelation with insolation (the test could not be done for the LMZ cores because the data covers a to short time span regarding to the insolation variation)
- positive interrelation with the standard reference SINT800 for LGMJ, but negative interrelation for LMZC and LMZG
- delayed interrelation for the inclination (delay about 100 years)

From these results we infer that the intensity of the used proxies for the intensity of the Earth's magnetic field underly some linkages with local climate and orbital variations. However, the global climate variation does not influence these intensity proxies. The positive and negative interrelations with the herbs pollen suggest a nonlinear mechanism behind these interrelations. It seems, that there is also a nonlinear interrelation between the inclination and the intensity. An interesting result is that of the negative interrelation to the palaeo-intensity reference SINT800 for the LMZ cores. This may be an effect of the less resolution of the SINT800 time series. Thus, the composition of the palaeo-intensity proxies of the LMZ and LGM cores could be a suitable addition to the SINT800 reference (Fig. 3).

Table 6: Interrelations between climate/ orbital proxies and proxies for the palaeo-intensity of the Earth's magnetic field gained from the quantitative analysis of CRPs. $+$ – positive interrelation; $-$ – negative interrelation; $*$ – no analysis because of to small variation of the insolation in the considered time; braced signs – less interrelation.

	Inclin.	SINT800	Inso12	Inso7	Pinus	Quercus	Herbs	^{18}O
LGMJ	$(\pm)$	$+$	$+$	$+$	$(-)$	$-$	$\pm$	0
LMZC2	$(+)$	$-$	$*$	$*$	$-$	0	$+$	$(+)$
LMZG2	$+$	$-$	$*$	$*$	$-$	$(+)$	$\pm$	0

5 Perspectives

On the basis of the results of our investigation, further studies are necessary.

Our new proxy for the intensity of the Earth's magnetic field should be compared with further climate and orbital data. Especially the application of another new extension of RPs, which is called XRP (extended cross recurrence plot or intersected recurrence plot) promise an improvement. XRPs are currently under development in our group and extent the method of recurrence plots to a real multivariate test and makes it to a tool for the study of phase-synchronization.

The applied methods (ICA and CRPs) enable a new approach to data from geology and climate, and especially under the focus on data from the Earth's magnetic field. Therefore, the evaluation of the results with other data sets from lake and marine environment is necessary. The expansion of the study to spatio-temporal measurements is recommended.

6 Utilisable Results

The application of the ICA to the rock-magnetic data from two Italian lakes is one of the first applications to geological data. It delivers a data set of the palaeo-intensity of the Earth's magnetic field, which does not include a climatic signal. This data set could be used for further invastigations of the Earth's magnetic field in the last 100 000 years.

We have created a Matlab® toolbox of modern data analysis techniques, especially in respect of the purposes considered within the SPP. It includes mainly methods of RP and CRP analysis. Moreover it containes methods for estimating the maximal correlation, mutual information and for data preparation. This toolbox is available through the WorldWideWeb. The current address is http://tocsy.agnld.uni-potsdam.de (follow the link CRP toolbox).

7 Qualifications

During the project time the Ph.D. student Norbert Marwan has prepared and finished a Ph.D. thesis about recurrence/ cross recurrence plots with the title "Encounters With Neighbours – Current Developments Of Concepts Based On Recurrence Plots And Their Applications". A large amount of work on this project is flown into the Ph.D. thesis.

Publications Resulting From the Project

MARWAN, N., Encounters With Neighbours – Current Developments Of Concepts Based On Recurrence Plots And Their Applications. Ph.D. thesis, University of Potsdam, 2003.

MARWAN, N., KURTHS, J., Nonlinear analysis of bivariate data with cross recurrence plots. Physics Letters A 302 (5–6), 2002, 299–307.
DOI: 10.1016/S0375-9601(02)01170-2

MARWAN, N., MEINKE, A., Extended recurrence plot analysis and its application to ERP data. International Journal of Bifurcation and Chaos "Cognition and Complex Brain Dynamics" 14 (2), 2004.

MARWAN, N., THIEL, M., NOWACZYK, N. R., Cross Recurrence Plot Based Synchronization of Time Series. Nonlinear Processes in Geophysics 9 (3/4), 2002a, 325–331.

MARWAN, N., TRAUTH, M. H., VUILLE, M., KURTHS, J., Comparing modern and Pleistocene ENSO-like influences in NW Argentina using nonlinear time series analysis methods. Climate Dynamics 21 (3–4), 2003, 317–326.
DOI: 10.1007/s00382-003-0335-3

MARWAN, N., WESSEL, N., KURTHS, J., Recurrence Plot Based Measures of Complexity and its Application to Heart Rate Variability Data. Physical Review E 66 (2), 2002b, 026702.
DOI: 10.1103/PhysRevE.66.026702