Independent Component Analysis of Rock Magnetic Measurements

Norbert Marwan

March 18, 2003

Title

Today I will not talk about recurrence plots. Marco and Mamen will talk about them later. Moreover, for the analysis of rock magnetic data I could not successfully use them. However, I've got to know another method, which I could successfully apply, and I would like to talk about this method. It is the independent component analysis. I have startet with this method just in the last month, thus I'm not yet a really expert in that. First I will introduce the problem of rock magnetic measurement, then I will give an overview about the independent component analysis and finally I will show you the first results.

Motivation

The main aim of all the work with the Earth' magnetic field is the understanding of the geodynamo, because the variations of the Earth' magnetic field, for example the reversal of the magnetic poles in the most extreme case, are of major importance for the mankind.

Palaeo Intensity

One step on the way for this understanding is the study of the past variations of the Earth' magnetic field. These past variations can be found in various geological archives. One possibility are lake sediments, in which the magnetic minerals can store information about the direction and the intensity of the Earth' magnetic field. However, when we measure the magnetic properties of these sediments, these measurements will contain also a large climatic impact. For example, the concentration of the magnetic minerals depends on the weathering processes, which on the other side depends on the climate. Or the kind of the magnetic minerals as well as their concentrations can depend on the biological activity in the lake. Finally, we always measure a mixture of cli-

Norbert Marwan, Talk PhD-day, Nonlinear Dynamics Group, University of Potsdam, 2003-03-18

mate and real Earth' magnetic field signals when we measure rock magnetic parameters. Since we are interested only in the information about the Earth' magnetic field, we need a method, which can separate these mixtures.

Basic Model

We will consider a linear model. Some independent source signals are linearly mixed and only these mixtures can be measured. I will denote the source signals by „s", the observations by „x" and the mixing matrix by „A". One well known method for the separation of the source signals is the principal component analysis. This method separates the components by using the criterion, which says, that these components will be linearly uncorrelated. Another method, which is rather unknown, is the independent component analysis. The difference is, that this method will separate components, which are not only linearly uncorrelated, but also nonlinearly uncorrelated; with other words, which are independent.

Independence/Uncorrelatedness

Independent variables are always uncorrelated, but linearly uncorrelated variables are not independent in general. This can be seen on this example of two uniformly distributed random variables. When we mix these variables and uncorrelate them by the PCA, we will get the right joint distribution, those components are obviously not independent. If we go to the maximal value of the one component, the other component will have some restricted values, which means that it depends on the first. Here we can see the task for the ICA: it has to rotate the distribution in that way, that the components become independent.

Gaussian Distribution

The gaussian distribution is a special case, because uncorrelated random variables with such a distribution are always independent. This is clear when we consider the joint distribution: it can be rotated in any way, which corresponds with orthogonal transformations and the preservation of the uncorrelatedness, but we will not find any direction, where the variables will have any dependence. Therefore, the ICA will not obtain any other result than the PCA for gaussian distributed variables.

Estimation Principles for ICA

For the estimation of the independent components, two basic concepts are used. The first uses algorithms, which determine components which are linearly and nonlinearly uncorrelated. The second concept uses an algorithm, which uses higher moments of the distributions in order to maximize the nongaussianity of linear combinations of the mixed signals. Here I will stop the explanation about the background, because I will not have enough time to explain. This alone would require an entire presentation.

Ambiguities

I have to mention that the ICA has some ambiguities which are, however, not really dramatic. The ICA cannot determine the number of the independent components. Some knowledge about the studied process is therefore useful. The variances and the signs of the independent components cannot be determined, but this is not of a great importance for us, because data will usually be normalized by their standard deviations and absolut variances play a minor role.

Illustration

Now I will present you an illustrating example, which will show how the ICA works and what the differences to the PCA are. This example consists of three oscillating sources, where the first source was transformed to be uniformly distributed. I have chosen such a mixing, that only the first two components are mixtures. Furthermore, I have added uniformly noise to these components.

Data Source and Mixed Signals

Left are the original sources and right are the mixtures.

Data PCA and ICA Components

Now we apply the PCA and the ICA to these mixed signals. First I show you the obtained components. As you can see, the first two components of the PCA do not match with the original source signals, whereas they match with ICA very well.

PCA

The result of the PCA are three uncorrelated signals, which can be obtained by the separation matrix V. The inverse of the separation matrix should be the mixing matrix A. But if we compare this matrix with the original mixing matrix, we see a large difference, especially if we consider the ratios of the interesting mixing coefficients.

ICA

The ICA, however, decomposes the mixtures in three independent components. The obtained mixing matrix seems to be quite similar to the original, except, that the signs and the absolut values do not match. However, if we focus to the ratios of the interesting mixing coefficients, they match rather well. I have not yet studied the statistically distribution or significance of the results. Therefore, I cannot say anything about the variance of these results.

Joint Distributions

When we look on the distributions of the source, mixing and separated signals, we can see what the PCA and the ICA have done with the data. The mixing has rotated and distorted the joint distribution; the PCA has rectified the joint distribution and the ICA has rectified and rotated the joint distribution. This joint distribution corresponds to the original.

Application to Rock Magnetic Data

Now let's come back to the main problem: the palaeo intensity of the Earth' magnetic field. Finally, I have some data sets from two Italian lakes, which contain various rock magnetic measurements and span a time range up to 100,000 years before present. The aim is to separate a signal which contains only the intensity of the Earth' magnetic field. For this task I have chosen these three measurements: the natural remanent magnetization, the anhysteretic remanent magnetization and the susceptibility. 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' 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' magnetic field. Each of the three measures depend on the grain size and the con-

Norbert Marwan, Talk PhD-day, Nonlinear Dynamics Group, University of Potsdam, 2003-03-18

centration 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.

Data Rock Magnetic Measurements

Here are the time series of these three measurements. The red line is some smoothing of the data.

Application to ... (back) & Data ICs

I have used the ICA in order to separate three independent signals from these mixtures, where one of them should contain the information about the Earth' magnetic field and the other information about the magnetic minerals and the climate, respectively. The result is this mixing matrix and these independent components.

Results Rock Magnetic Data

However, have we got a suitable result? In order to get an answer to this question, I have computed correlations between the independent components and some climate proxies. As you can see, the first component does not correlate with the climate signals, whereas the other two correlate and, hence, contain a climate signal. On this place I have to mention that in the palaeo magnetic community the ratios of natural remanent magnetization and susceptibility and natural remanent magnetization and anhysteretic remanent magnetization are usually used as proxies for the intensity of the Earth' magnetic field. These both ratios correlates well with the first component, however, each of them correlates also with one of the second and third component. This is a clear sign that these ratios still contain some rock properties, especially the grain size, and are, therefore, impacted by a climate signal. This can be seen in the second table, where I have computed the correlations for these measures with some climate proxies. Furthermore, as you can see, the independent component contains much less climate impact as the other both measures. Finally, I have compared this independent component and the other measures with a reference data set of the palaeo intensity of the Earth magnetic field. This reference data is stacked data from 33 records around the world and covers the last 800,000 years. These correlation coefficients show also the improvement by using the ICA. The independent component correlates better with the reference data than the other measures, especially the ratios.

Quercus pollen are a proxy for the local temperature and Pinus pollen for temperature and rainfall. The proxy for the global temperature was computed as a principal component from the arctic and antarctic oxygene isotopic records of the Vostok ice core in Antarctica and the GISP2 ice core in Greenland.

Conclusion

Let's come to the end and to the conclusions. The first conclusion is that the ICA is more general than the PCA, the second, that the ICA separates mixed signals. Furthermore, we can state that the application to rock magnetic measurements reveals an intensity signal of the Earth' magnetic field, which is a better representation of the palaeo intensity signal than the so far used ratios of rock magnetic measurements. This is a good starting point for the further work, where I will apply this method to two further data sets from italian lakes and compare their results.

Independent Component Analysis of Rock Magnetic Measurements

Norbert Marwan

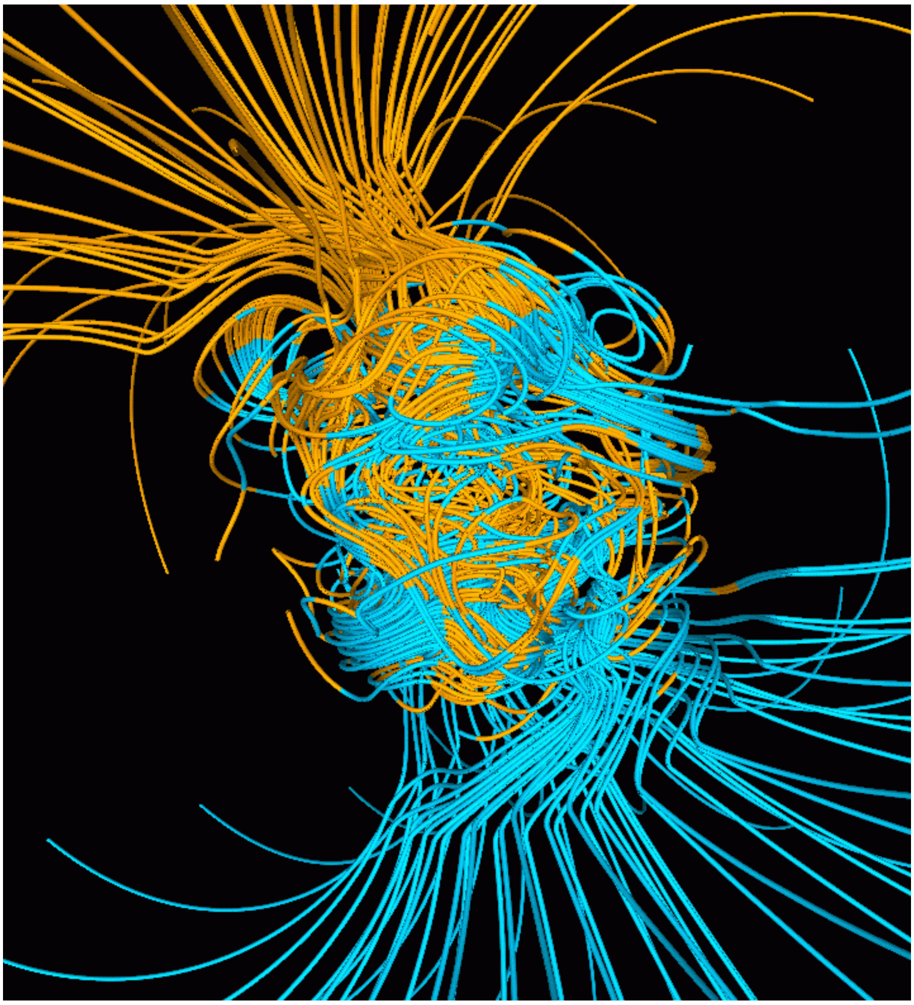
1. Basic Problem

2. Independent component analysis

3. Application to rock magnetic measurements

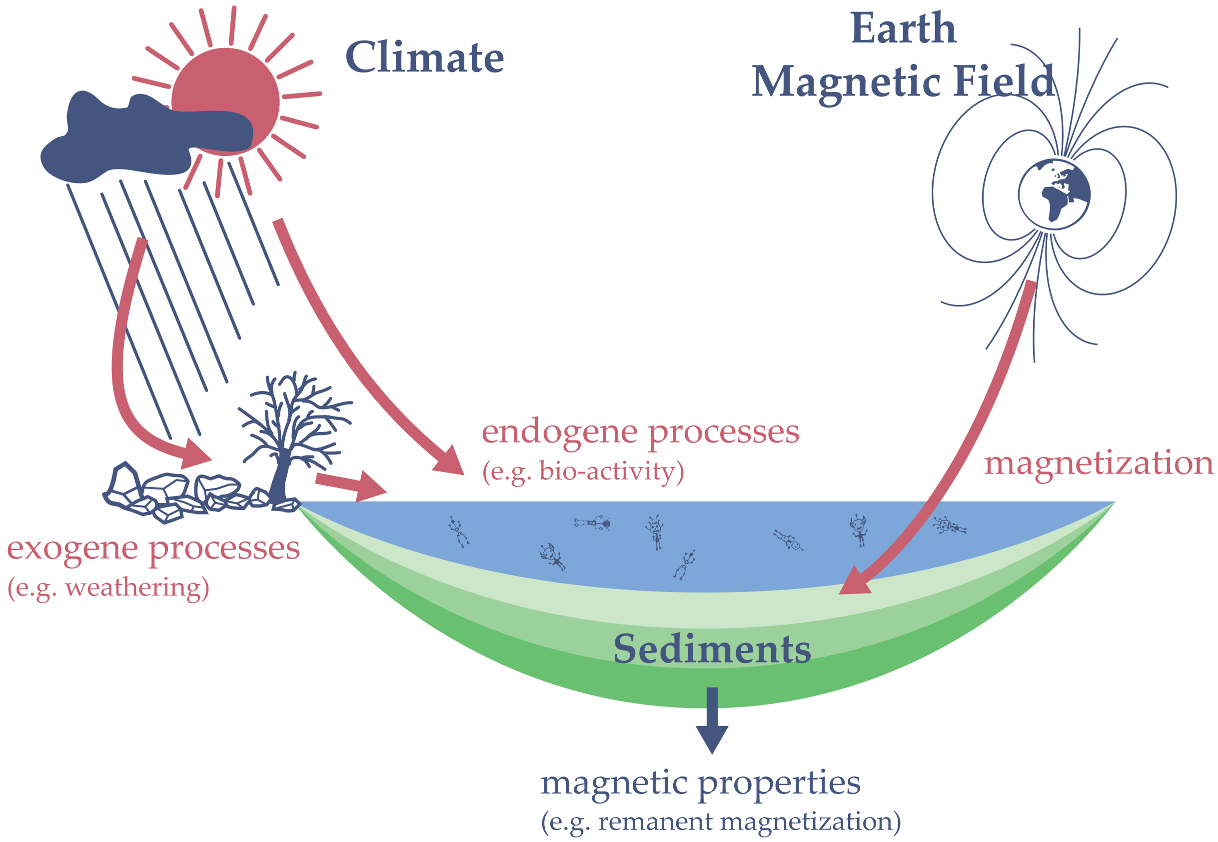
Motivation

Understanding the Earth′ geodynamo (variation of the Earth′ magnetic field)



Palaeo Intensity

Rock magnetic measurements contain mixtures of local and regional climatic as well as Earth' magnetic field intensity signals of the past.



Can we extract the different source signals from these mixtures?

Basic Model

n independent source signals $s_i(t)$,
m observations $x_j(t)$ – linear mixings of s_i

$$\vec{x}(t) = \mathbf{A}\,\vec{s}(t)$$

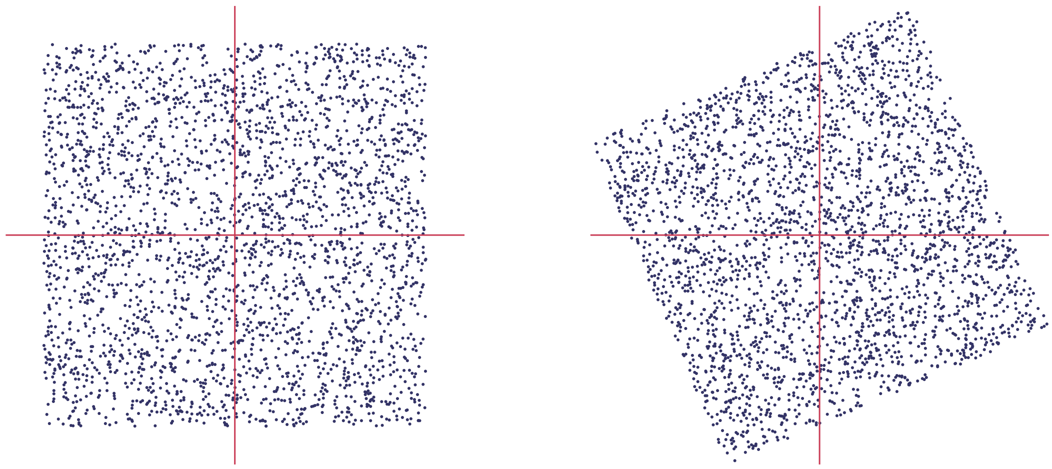
Problem:
Separation of the sources s_i from the observations, i. e. estimation of the mixing matrix $\mathbf{A}$.

Methods:
1. Principle component analysis (PCA) – uncorrelated components
2. Independent component analysis (ICA) – statistically independent components

independence $\Rightarrow$ uncorrelatedness

$$
\begin{aligned}
\operatorname{cov}(g(x), h(y)) &= \mathrm{E}\{g(x)\, h(y)\} \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x)\, h(y)\, p_{x,y}(x,y)\, dx\, dy \\
&= \int_{-\infty}^{+\infty} g(x)\, p_x(x)\, dx \int_{-\infty}^{+\infty} h(y)\, p_y(y)\, dy \\
&= \mathrm{E}\{g(x)\}\; \mathrm{E}\{h(y)\},
\end{aligned}
$$

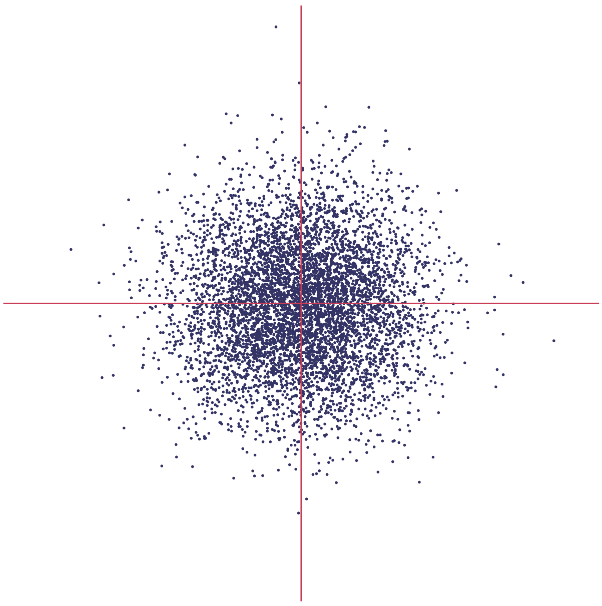
uncorrelatedness $\not\Rightarrow$ independence



Joint distributions of independent random variables with uniform distribution (left) and their uncorrelated – but not independent – mixtures (right).

special case: uncorrelated variables with gaussian joint distribution

$$
\begin{aligned}
p_{x,y}(x,y) &= \frac{1}{2\pi}\,\mathrm{e}^{-\frac{x^2+y^2}{2}} \\
&= \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{x^2}{2}}\,\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{y^2}{2}} = p_x(x)\,p_y(y)
\end{aligned}
$$



Joint distribution of two independent Gaussian variables.

Estimation Principles for ICA

Nonlinear decorrelation:
Find such components y_i which are uncorrelated *and* whose transformed components $f_i(y_i)$ are uncorrelated, (f_i are some suitable nonlinear functions).

Maximum nongaussianity:
Find local maxima of nongaussianity of linear combinations $y = \sum b_i x_i$; each local maximum gives one independent component.

Motivation for maximum nongaussianity

The sum of independent random variables tends closer toward a Gaussian distribution than the original random variables (central limit theorem).

$\vec{x} = \mathbf{A}\,\vec{s}$, with observations $\vec{x}$ and source signals $\vec{s}$.

We consider the linear combination $y = \sum_i b_i x_i$ (corresponding to $y = \vec{b}^T \mathbf{A} \vec{s}$). Such a vector $\vec{b}$, so that $\vec{b}^T \mathbf{A}$ has only one nonzero component, reveals actually one of the independent components.

According to the central limit theorem, $\vec{b}$ has to be estimated in that way that the linear combination y is maximal nongaussian.

Measures for nongaussianity are e. g. kurtosis

$$kurt = E\{x^4\} - 3[E\{x^2\}]^2$$

or negentropy

$$J(\vec{x}) = H(\vec{x}_{gauss}) - H(\vec{x}).$$

Ambiguities

The ICA cannot determine

1. the number of the independent components;

2. the variances of the independent components;

3. the sign of the independent components.

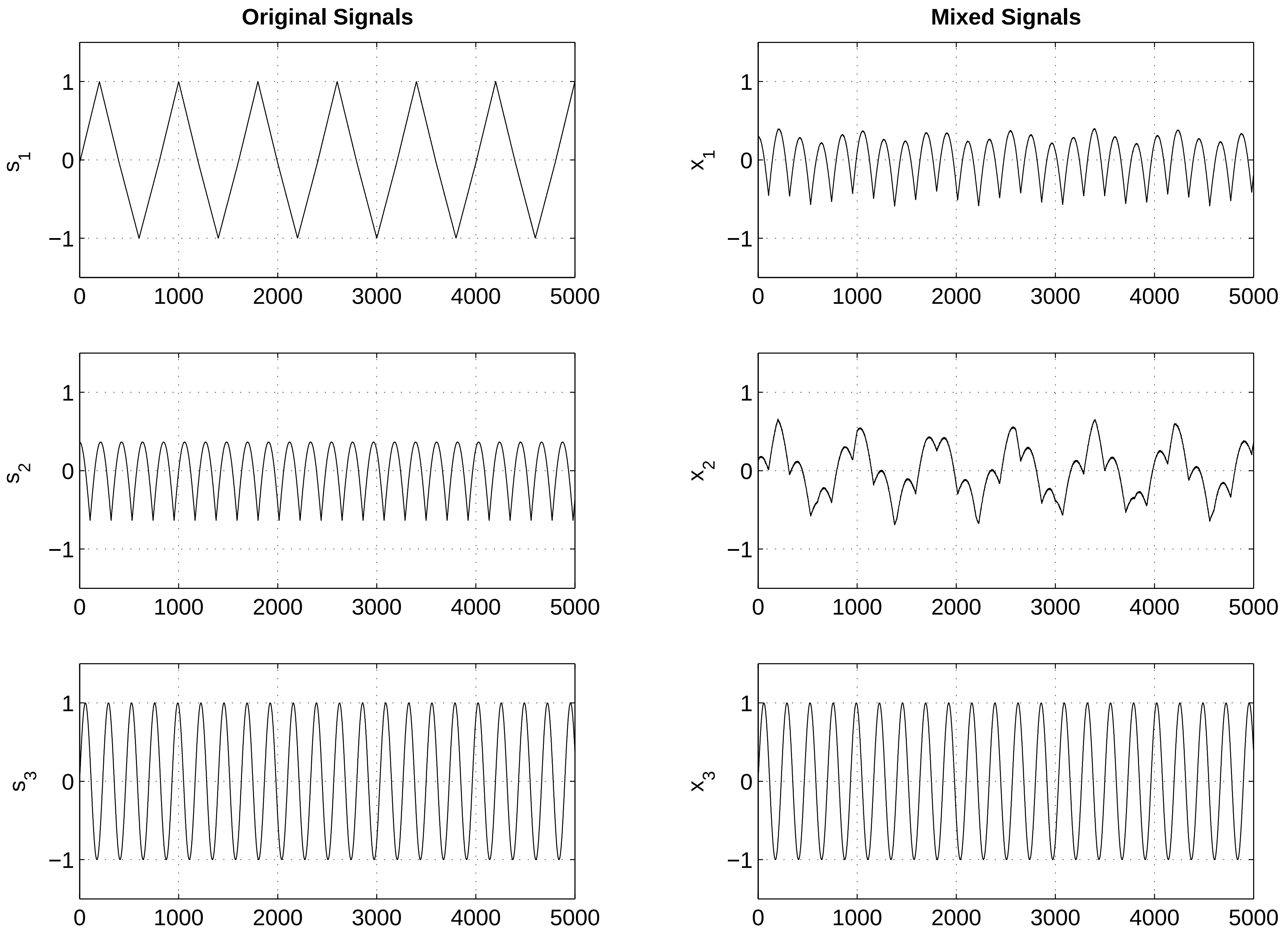
Illustration

$$
\begin{aligned}
s_1(t) &= \sin\left(\frac{2\pi}{800}t\right) \quad \text{(IID transf.)} \\
s_2(t) &= \left|\cos\left(\frac{2\pi}{424}t\right)\right| \\
s_3(t) &= \sin\left(\frac{2\pi}{233}t\right)
\end{aligned}
$$

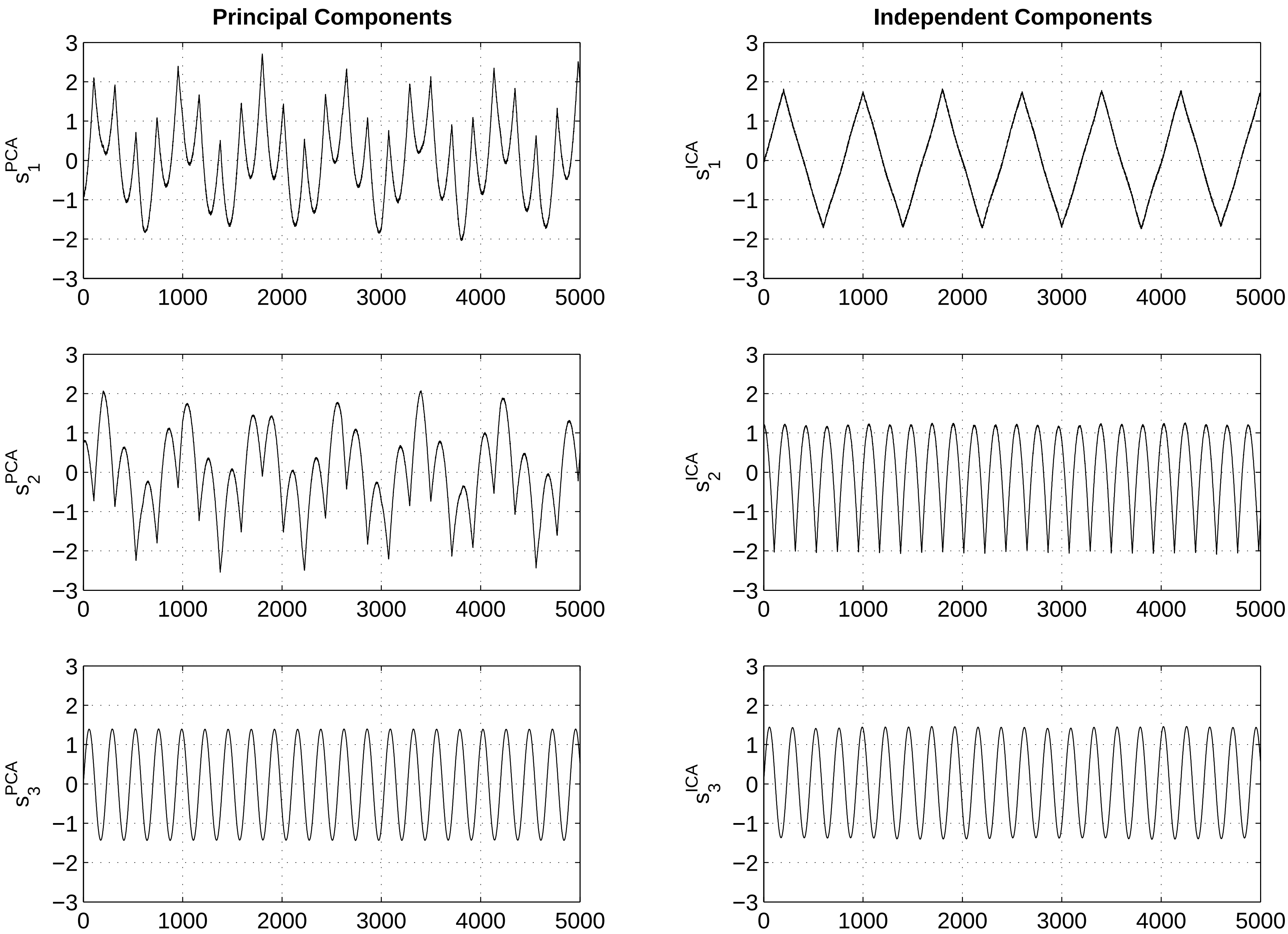
with its mixing signals

$$
\begin{aligned}
x_1(t) &= 0.1\, s_1(t) + 0.8\, s_2(t) + 0.01\, \xi_1 \\
x_2(t) &= 0.5\, s_1(t) + 0.4\, s_2(t) + 0.02\, \xi_2 \\
x_3(t) &= s_3(t)
\end{aligned}
$$

$s_1(t)$ is transformed to the uniform distribution; ξ_i is uniformly distributed noise



Original Signals
Mixed Signals
s1
s2
s3
x1
x2
x3
1
0
−1
0
1000
2000
3000
4000
5000



Principal Components
Independent Components
s_1^{PCA}
s_2^{PCA}
s_3^{PCA}
s_1^{ICA}
s_2^{ICA}
s_3^{ICA}
3
2
1
0
−1
−2
−3
0
1000
2000
3000
4000
5000

Original:

$$\mathbf{A} = \begin{pmatrix} 0.1 & 0.8 & 0.00 \\ 0.5 & 0.4 & 0.00 \\ 0.0 & 0.0 & 1.00 \end{pmatrix}$$

$$\left|\frac{a_{11}}{a_{12}}\right| = 0.125, \quad \left|\frac{a_{21}}{a_{22}}\right| = 1.25$$

PCA

The PCA decomposes the observations in three uncorrelated signals

$$\vec{s} = \mathbf{V}\,\vec{x}$$

e. g. by eigenvalue decomposition of the covariance matrix

$$\mathbf{C} = \mathbf{E}\mathbf{D}\mathbf{E}^T, \quad \mathbf{V} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T$$

with

$$\mathbf{V} = \begin{pmatrix} 1.32 & -1.32 & 0.04 \\ 0.54 & 0.54 & -0.02 \\ 0.00 & -0.04 & -1.00 \end{pmatrix}$$

$$\mathbf{A}^{\mathbf{PCA}} = \mathbf{V}^{-\mathbf{1}} = \begin{pmatrix} 0.38 & 0.93 & 0.00 \\ -0.38 & 0.92 & -0.04 \\ 0.01 & -0.04 & -1.00 \end{pmatrix}$$

$$\left|\frac{a_{11}}{a_{12}}\right| = 0.41, \quad \left|\frac{a_{21}}{a_{22}}\right| = 0.41$$

ICA

The ICA decomposes the observations in three independent signals

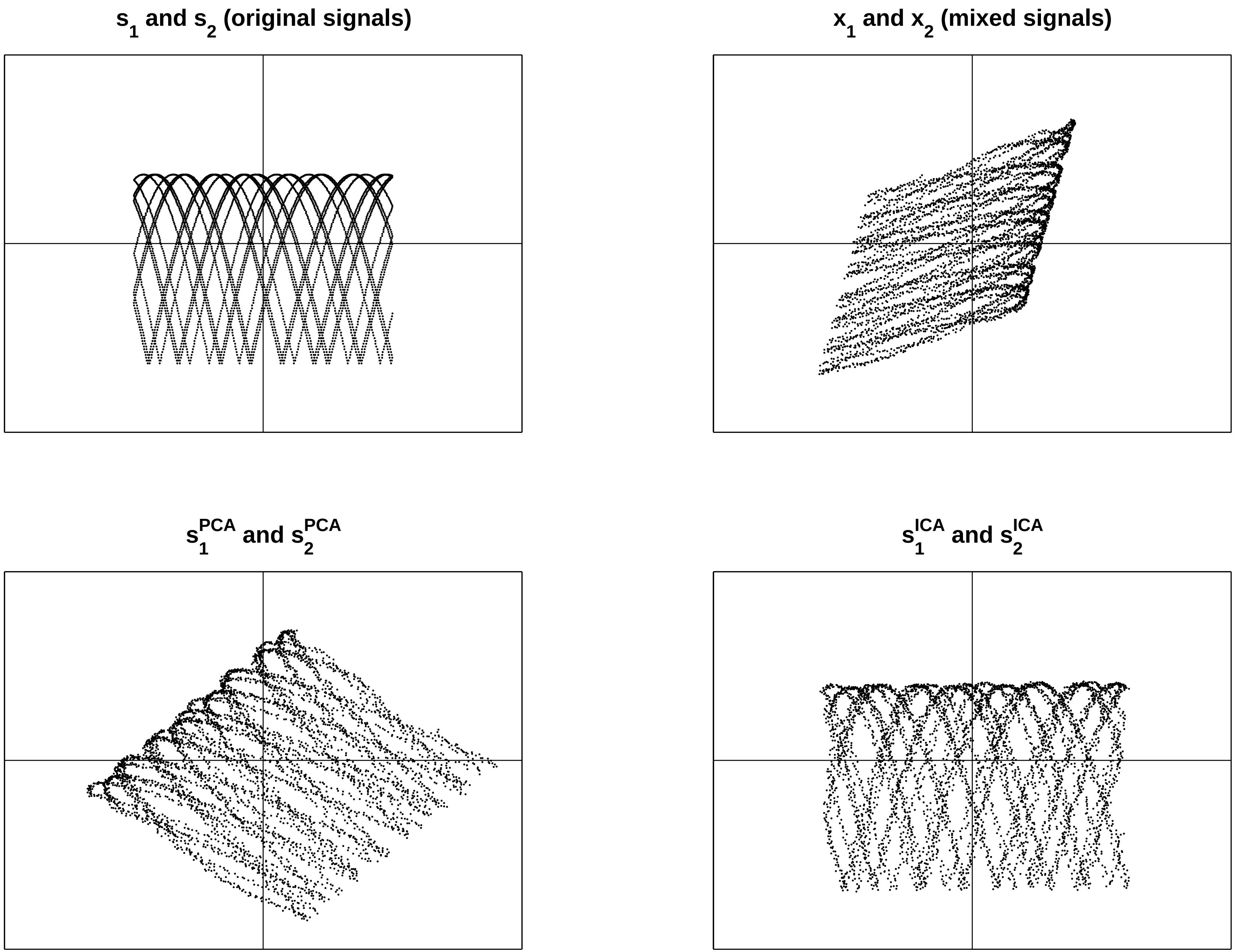
$$\vec{s} = \mathbf{W}\,\vec{x}$$

with

$$\mathbf{W} = \begin{pmatrix} 0.90 & -1.42 & 0.02 \\ 1.11 & -0.16 & 0.01 \\ 0.01 & 0.00 & -1.00 \end{pmatrix}$$

$$\mathbf{A}^{\mathbf{ICA}} = \mathbf{W}^{-1} = \begin{pmatrix} -0.11 & 0.99 & -0.01 \\ -0.78 & 0.63 & 0.01 \\ 0.00 & -0.01 & 1.00 \end{pmatrix}$$

$$\left|\frac{a_{11}}{a_{12}}\right| = 0.11, \quad \left|\frac{a_{21}}{a_{22}}\right| = 1.24$$


s1 and s2 (original signals)
x1 and x2 (mixed signals)
s1PCA and s2PCA
s1ICA and s2ICA

Application to Rock Magnetic Data

Rock magnetic measurements (Lake Lago Grande di Monticchio in Italy)

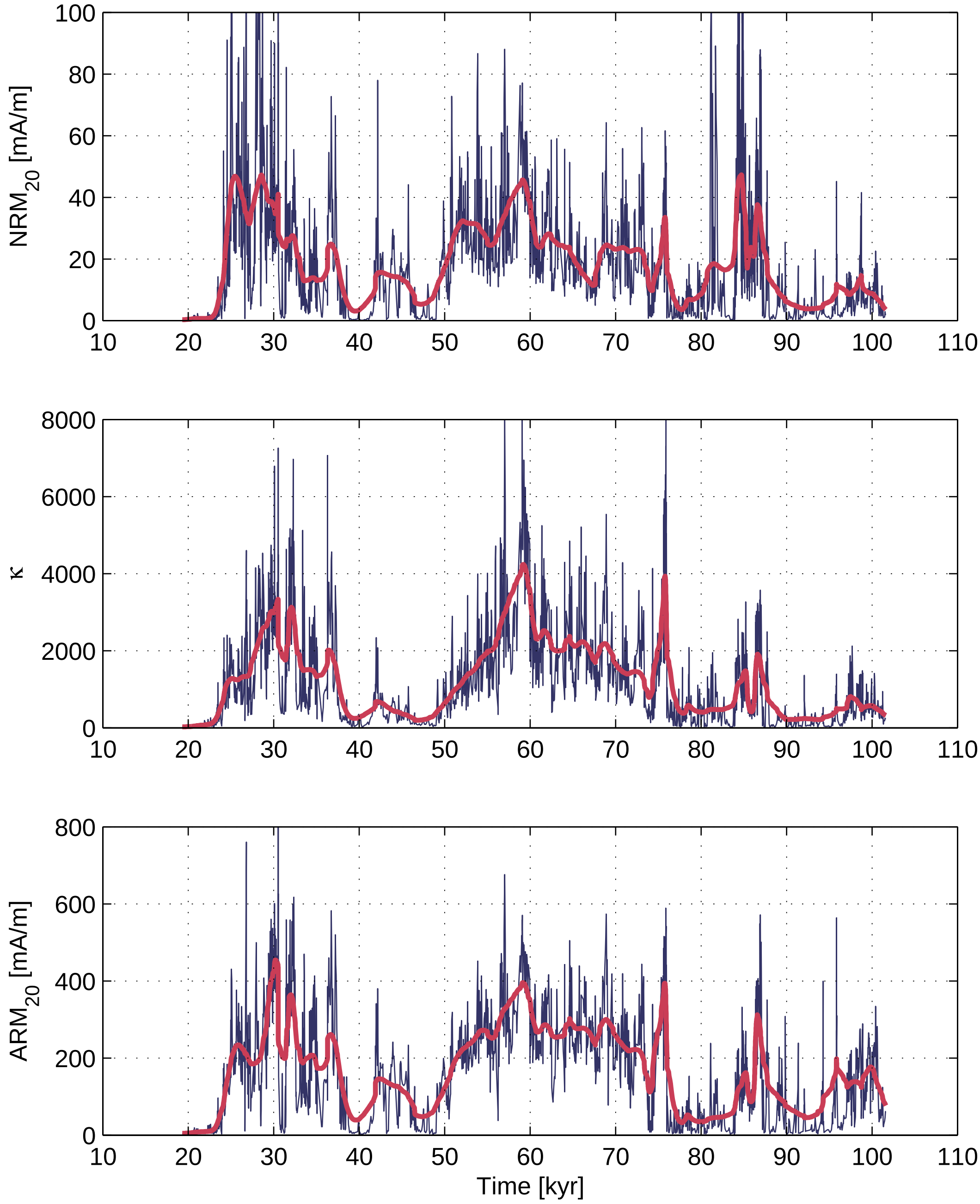
$$
\begin{aligned}
NRM &= f_1(F) + f_2(c) + f_3(s), \quad c, s = f(C) \\
ARM &= g_1(c) + g_2(s_{small}) \\
\kappa &= h_1(c) + h_2(s_{large})
\end{aligned}
$$

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

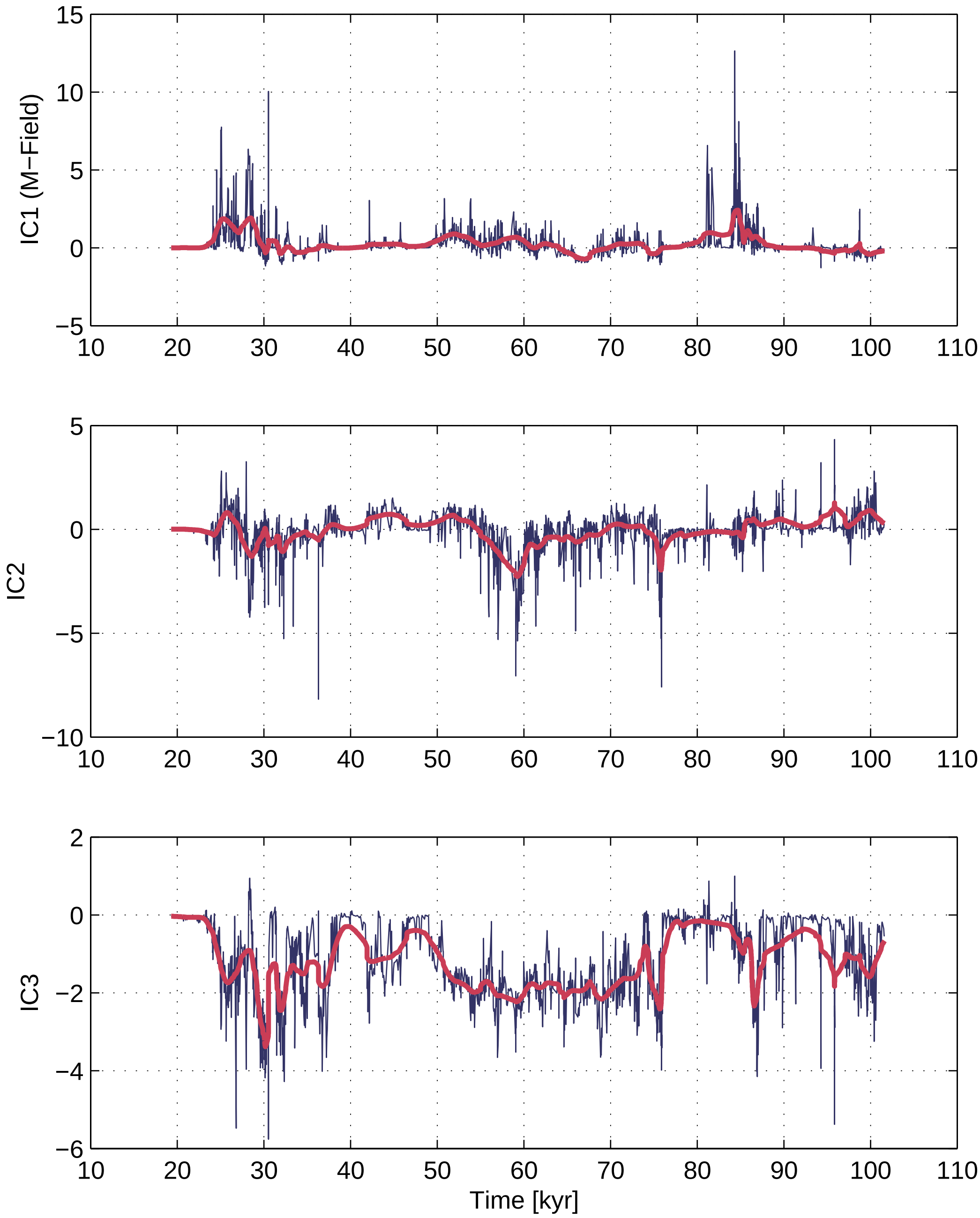
Separation of the factors F, c and s with the ICA. The analysis reveals three ICs s_i ($\vec{x} = \mathbf{A}\vec{s}$) with the mixing matrix:

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

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



100
80
60
40
20
0
NRM20 [mA/m]
8000
6000
4000
2000
0
κ
800
600
400
200
0
ARM20 [mA/m]
10
20
30
40
50
60
70
80
90
100
110
Time [kyr]


IC1 (M−Field)
IC2
IC3
Time [kyr]

Results Rock Magnetic Data

Correlation coefficients between these ICs and the underlying signals as well as the proxy data for the climate, reveals a clear distinction of these signals.

	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

Q – Quercus pollen; $CLIM$ – proxy for global temperature

Furthermore, the first IC s_1 contains much less climate impact as the usually used ratios of NRM with ARM and κ, respectively.

	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

Q – Quercus, P – Pinus pollen; $CLIM$ – proxy for global temperature

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

Conclusion

1. The ICA is more general than PCA.

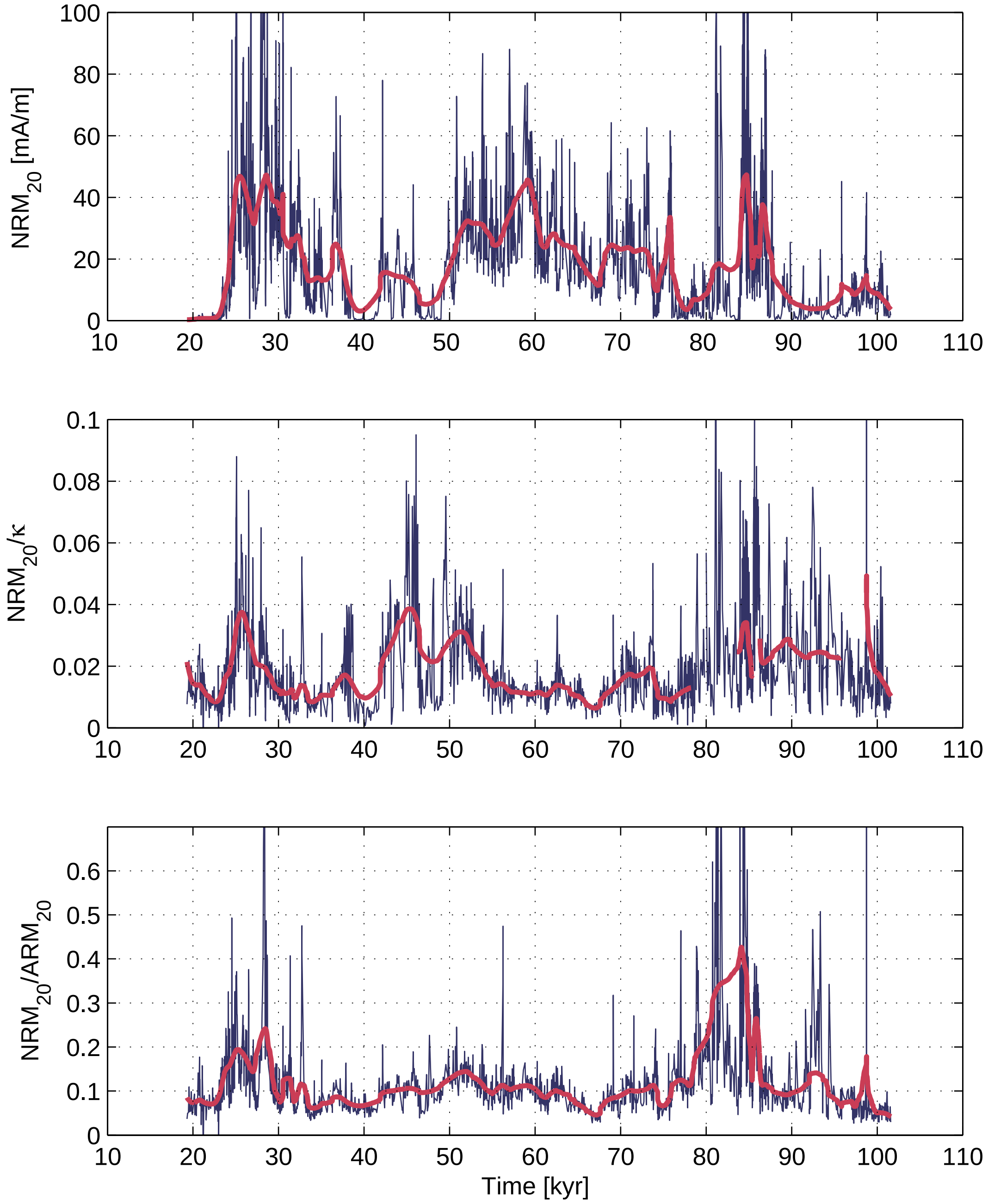
2. The ICA separates mixed signals.

3. The application to rock magnetic measurements reveals an intensity signal of the Earth' magnetic field.

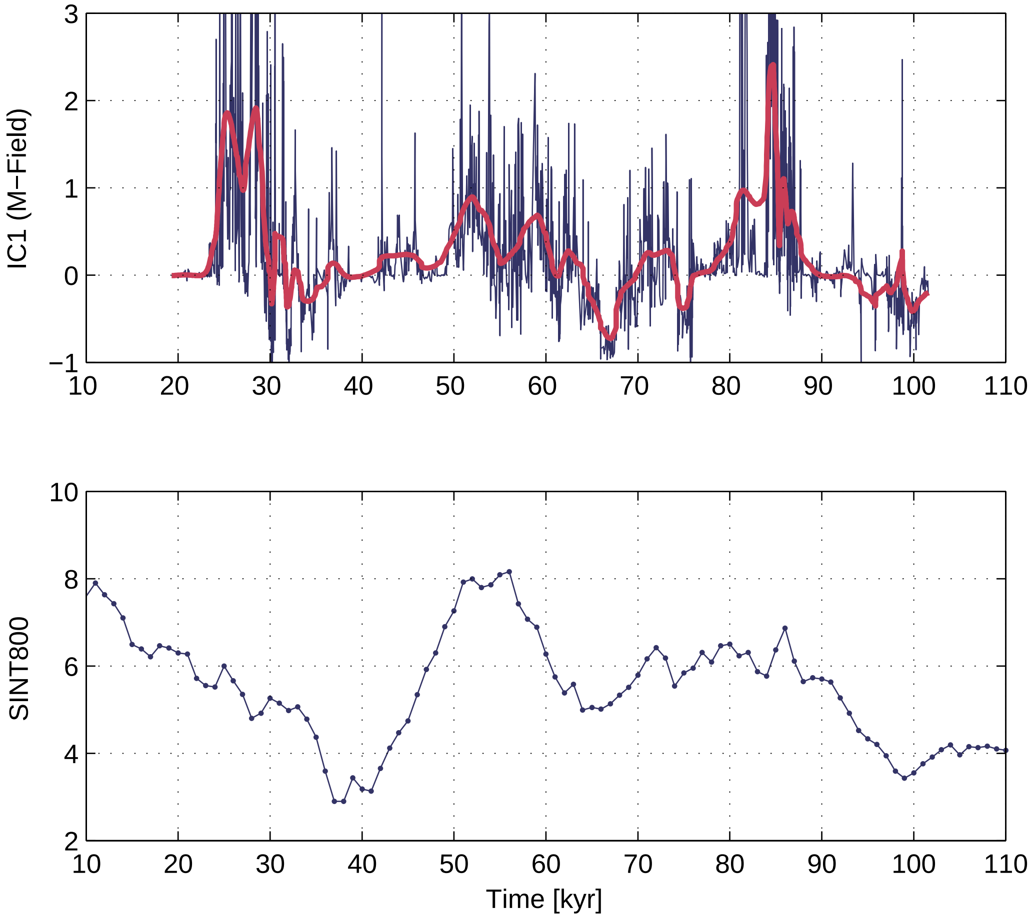
4. The obtained component is a better representation of the Earth' magnetic field than usually used ratios of rock magnetic measurements.

To do

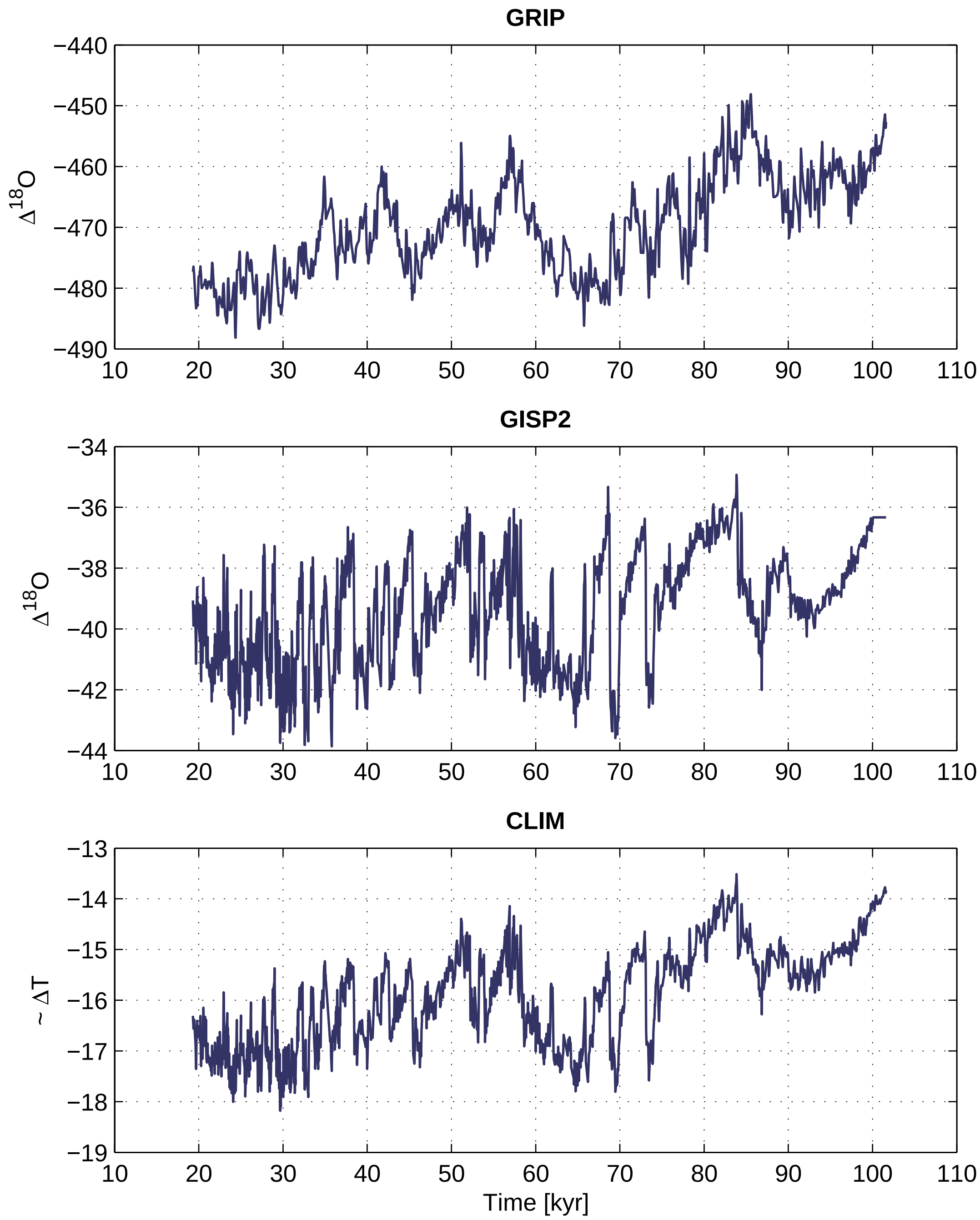
Application of the ICA to two further available data sets from italian lakes and comparison of the results.



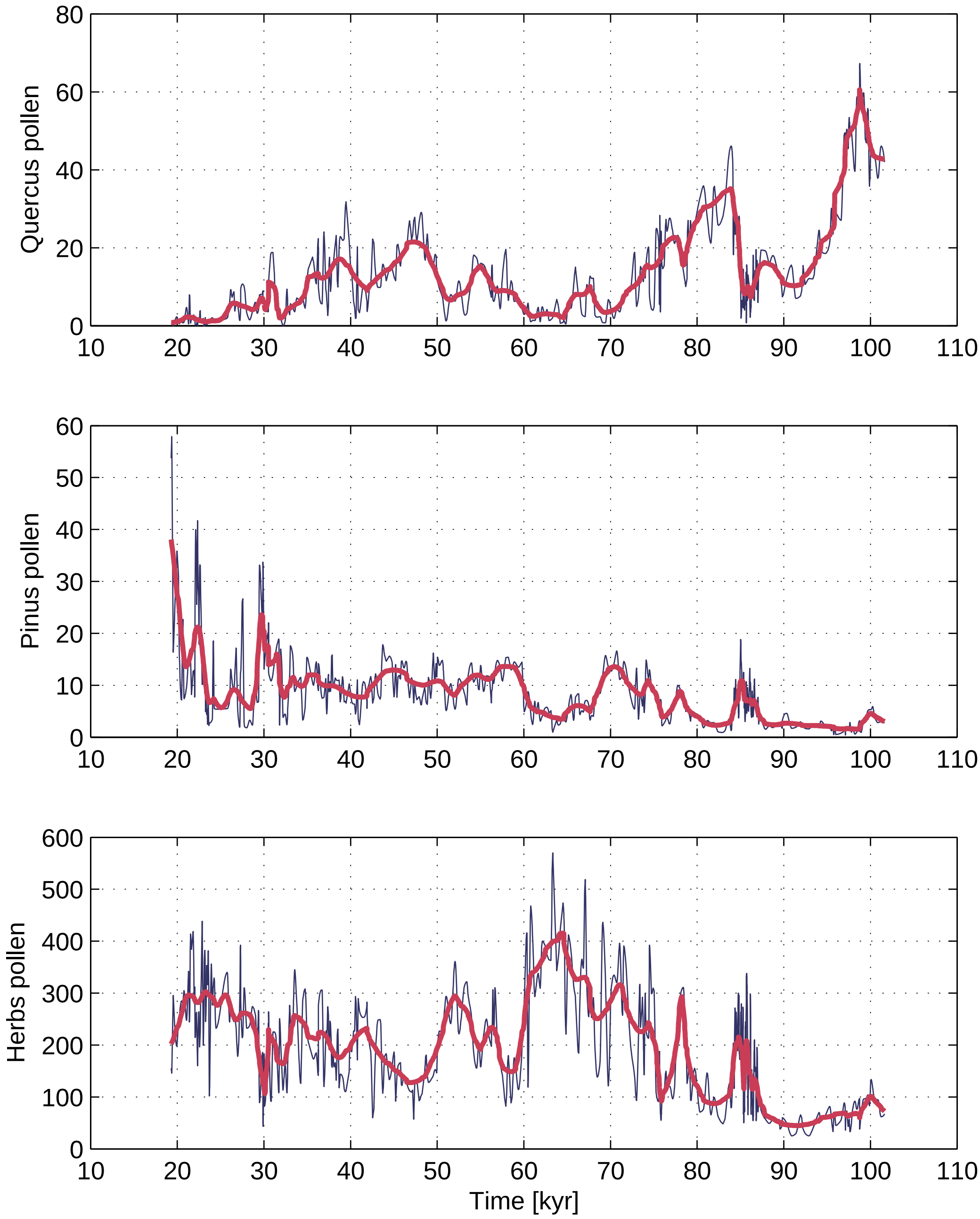
NRM20 [mA/m]
100
80
60
40
20
0
10
20
30
40
50
60
70
80
90
100
110
NRM20/κ
0.1
0.08
0.06
0.04
0.02
0
NRM20/ARM20
0.6
0.5
0.4
0.3
0.2
0.1
0
Time [kyr]



3
2
1
0
−1
IC1 (M−Field)
10
20
30
40
50
60
70
80
90
100
110
10
8
6
4
2
SINT800
Time [kyr]



GRIP
−440
−450
−460
−470
−480
−490
Δ18O
10
20
30
40
50
60
70
80
90
100
110
GISP2
−34
−36
−38
−40
−42
−44
Δ18O
CLIM
−13
−14
−15
−16
−17
−18
−19
~ ΔT
Time [kyr]



Quercus pollen
Pinus pollen
Herbs pollen
Time [kyr]