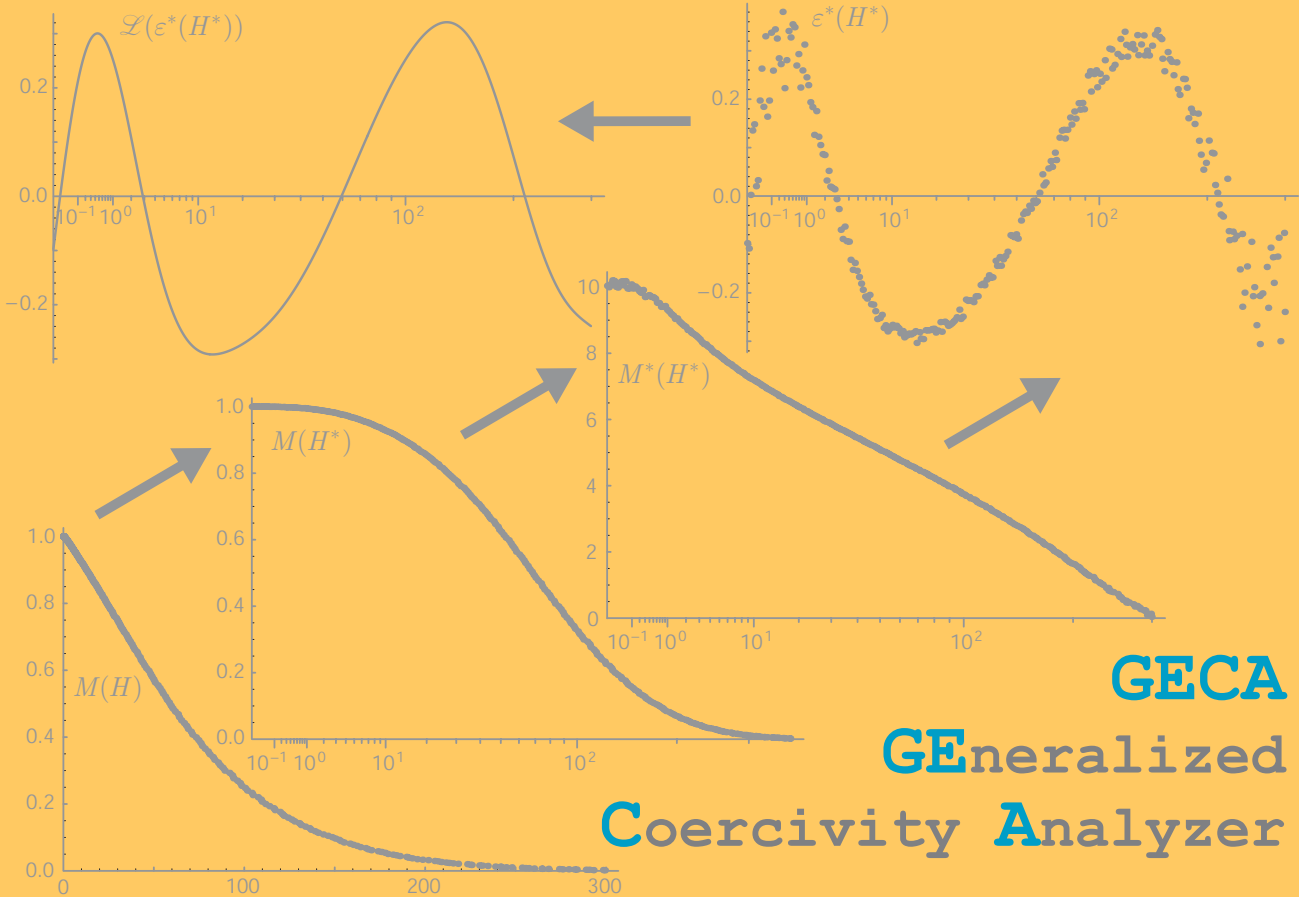
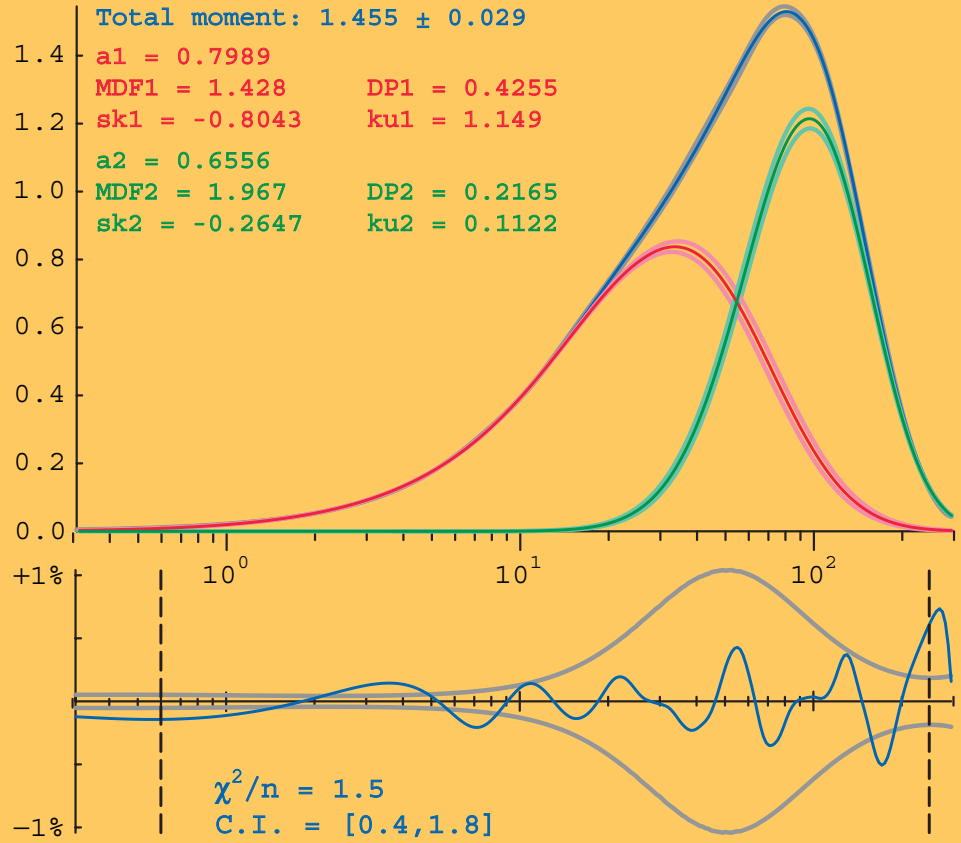


GECA 2.0

© Ramon Egli 2004



Introduction

The program GECA (Generalized Coercivity Analyzer) is part of the package CODICA. GECA performs a component analysis based on special generalized functions which can fit the coercivity distributions of natural and artificial magnetic components particularly well. It also performs a Pearson's χ^2 goodness of fit test to evaluate the number of functions required to model a coercivity distribution. Finally, GECA performs an error estimation and calculate the confidence limits for each model parameter.

Read carefully this manual to learn about GECA and take full advantage from the different possibilities offered by the program to perform a component analysis and verify its significance. This manual contains a theoretical part, which gives you the background to understand the basic ideas of GECA, and a practical part, which guides you through each step of the program. You can practice with the examples delivered together with this program. CODICA is designed to work optimally on coercivity distributions calculated with CODICA and stored in files with extension `.slog`.

Click on the following topics to see the contents of this manual:

<u>Theoretical background: coercivity distributions</u>	2
• <u>Finite mixture models</u>	2
• <u>logarithmic Gaussian functions</u>	2
• <u>Skewed Generalized Gaussian functions (SGG)</u>	3
• <u>distribution parameters</u>	3
<u>Some aspects of component analysis</u>	5
<u>Performing and testing a component analysis</u>	8
• <u>merit function</u>	8
• <u>mean squared residuals</u>	8
• <u>Chi-square estimator</u>	8
• <u>local and global minima of the merit function</u>	9
• <u>Pearson's Chi-square goodness of fit test</u>	11
<u>A program example</u>	14
<u>Cautionary note</u>	35

Theoretical background: coercivity distributions

A group of magnetic grains with similar chemical and physical properties, distributed around characteristic values, is called a *magnetic component*. Examples of magnetic components are pedogenic magnetite (nanometric magnetite particles with a wide grain size distribution), and magnetosomes (prismatic magnetite with a very narrow grain size distribution between 40 nm and 80 nm). Magnetic components have a simple-shaped, unimodal distribution of coercivities. Commonly, the coercivity distribution of a single magnetic component is modelled with a logarithmic Gaussian function:

$$G(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{\log^2(x/\mu)}{2\sigma^2}\right] \quad (1)$$

In the literature, x is identified with the magnetic field H , μ is the *median destructive or acquisition field* $H_{1/2}$, called also MDF and MAD respectively, and σ the *dispersion parameter* DP. However, not all coercivity distributions can be modelled appropriately with (1). Experimental and theoretical coercivity distributions of single components are better described by [distribution functions with four parameters](#). The two additional parameters control the [skewness](#) and the [squareness](#) of the distribution.

The coercivity distribution $f(H)$ of a mixture of different magnetic components may be considered as a linear combination of the coercivity distributions of the single components:

$$f(H) = \sum_{i=1}^n c_i M_{ri} f_i(H | \theta_i) \quad (2)$$

where c_i and M_{ri} are the concentration and the saturated magnetization of the i -th component respectively, and $f_i(H | \theta_i)$ is the corresponding coercivity distribution with the parameters $\theta_i = (\theta_{i1}, \dots, \theta_{ik})$. Equation (2) is called a *finite mixture model*, and $f_i(H | \theta_i)$ are the so-called *end members*. Equation (2) assumes that the magnetization of all components adds linearly (*linear additivity*). This assumption does not hold in case of magnetic interactions between the magnetic grains of different components. However, magnetic interactions between different components are not likely to occur in natural samples, since each component is expected to have a different origin and to hold different places within a nonmagnetic matrix. On the other hand, magnetic interactions within the same component are possible, but they do not affect the linear additivity law.

Coercivity distributions of single magnetic components are described by *probability density functions* (PDF). The shape of a PDF is controlled by a set of *distribution centers* μ_n with related *dispersion parameters* σ_n , with $n \in \mathbb{N}$. Special cases are given when $n = 1$ (μ_1 is the median, σ_1 the mean deviation), $n = 2$ (μ_2 is the mean, σ_2 the standard deviation), and $n \rightarrow \infty$ (μ_∞ is the mid-range and σ_∞ the half-range). The *dispersion parameter* DP corresponds to σ_2 on a logarithmic field scale. The symmetry of a PDF is described by the *coefficient of skewness* s , where

$s = \sigma_3^3/\sigma_2^3$. Symmetric distributions are characterized by $s = 0$ and $\mu_n = \mu_2$. The curvature of a PDF is described by the *coefficient of excess kurtosis* k , where $k = \sigma_4^4/\sigma_2^4 - 3$. The Gaussian PDF is characterized by $k = 0$.

The description of non-Gaussian PDF involves the use of functions with more than two independent parameters. It is of great advantage if such functions maintain the general properties of a Gauss PDF: the n -th derivative should exist over \mathbb{R} and $\sigma_n < \infty$ for all values of $n \in \mathbb{N}$. Furthermore, the Gaussian PDF should be a particular case of such functions. A good candidate is the *generalized Gaussian distribution* GG , known also as the *general error distribution*. The Gaussian PDF is a special case of GG distributions. Other special cases are the Laplace distribution and the box distribution. The GG distribution is symmetric: $s = 0$. In GECA, a particular set of *skewed generalized Gaussian distributions*, called SGG , is used to model single components. A SGG function is given by:

$$SGG(x, \mu, \sigma, q, p) = \frac{1}{2^{1+1/p} \sigma \Gamma(1 + 1/p)} \frac{|qe^{qx^*} + q^{-1}e^{x^*/q}|}{e^{qx^*} + e^{x^*/q}} \exp \left[-\frac{1}{2} \left| \ln \left(\frac{e^{qx^*} + e^{x^*/q}}{2} \right) \right|^p \right] \quad (3)$$

with $x^* = (x - \mu)/\sigma$, $x = \log H$, and $0 < |q| \leq 1$. The GG distribution is a special case of (3) for $q = 1$, and the Gauss distribution is a special case of (3) with $q = 1$ and $p = 2$. The relation between the distribution parameters μ, σ, q, p and some statistical properties is given in Table 1.

Distribution properties	Definition	Relation with the distribution parameters	Comments
Median $x_{x/2}$	$\int_{-\infty}^{x_{1/2}} f(x) dx = \frac{1}{2}$	$x_{1/2} = \mu$	$x_{x/2}$ is also called MDF or MAF
Mean μ_2	$\mu_2 = \int_{-\infty}^{+\infty} f(x)x dx$	$\mu_2 \approx \mu + \frac{s}{6}(1 + 0.856k)$ for $q \rightarrow 1, p \rightarrow 2$	generally not used in the literature
Standard deviation σ_2	$\sigma_2^2 = \int_{-\infty}^{+\infty} f(x)(x - \mu_2)^2 dx$	$\sigma_2^2 = \sigma^2(1 + 0.856k)(1 - s /3)$ for $q \rightarrow 1, p \rightarrow 2$	σ_2 is also called DP
Skewness s	$s = \sigma_3^3/\sigma_2^3$ $\sigma_3^3 = \int_{-\infty}^{+\infty} f(x)(x - \mu_2)^3 dx$	$s \approx -6 \operatorname{sgn} q (1 - q)^2 (1 + 1.856k)$	$q > 0$: left skewed $q < 0$: right skewed
Kurtosis k	$k = \sigma_4^4/\sigma_2^4 - 3$ $\sigma_4^4 = \int_{-\infty}^{+\infty} f(x)(x - \mu_2)^4 dx$	$k \approx 2 - p$	$p > 2$: box-shaped $p < 2$: tip-shaped

Table 1: Relation between statistical distribution properties and distribution parameters for a SGG function. Except for the median, the relations are not analytical; approximations are given in the case of small deviations from a Gaussian distribution.

Examples of SGG functions with different parameters are given in Fig. 1. The parameters of the coercivity distribution of some calculated and measured coercivity distributions are plotted in Fig. 2.

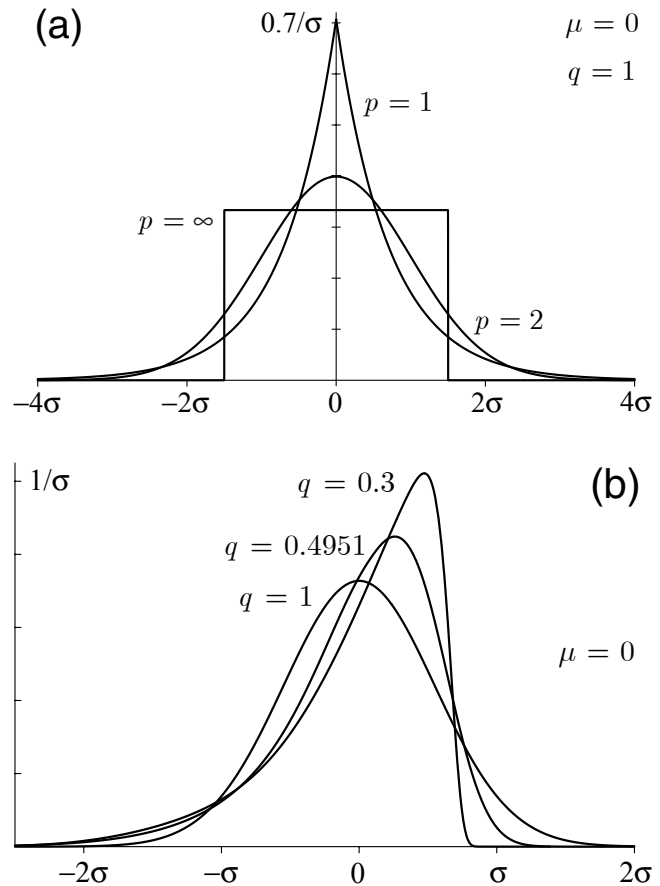


Fig. 1: Examples of SGG distributions. (a) Some particular cases with $\mu = 0$, $\sigma_2 = 1$ and $q = 1$ are plotted. The skewness of all curves is zero. Furthermore, $p = 1$ for a Laplace distribution, $p = 2$ for a Gauss distribution and $p = \infty$ defines a box distribution. (b) Some left-skewed SGG distributions with $\mu = 0$ and $\sigma_2 = 0.5484$ are plotted. The SGG distribution with $q = 0.4951$ is an excellent approximation of the logarithmic plot of a negative exponential distribution.

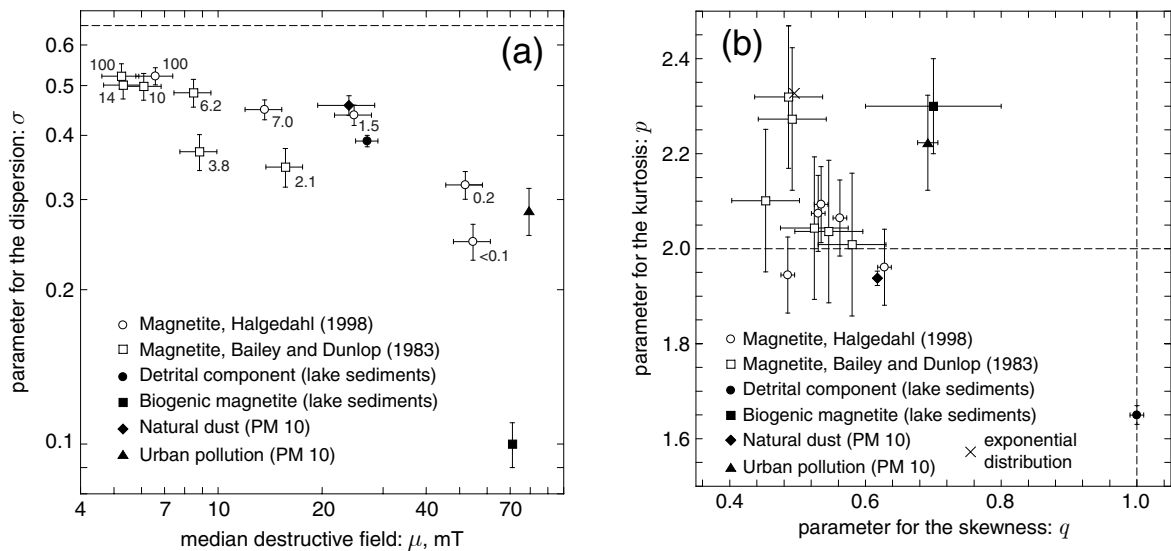


Fig. 2: Coercivity distribution parameters μ , σ , q and p for the AF demagnetization of IRM in various synthetic and natural samples. Numbers beside the points indicate the grain size in μm . (a) Scatter plot of μ and σ . The dashed line indicates the value of σ for a negative exponential distribution. (b) Scatter plot of q and p . The cross point of the dashed lines corresponds to the values of q and p for a logarithmic Gaussian distribution. All samples show significant deviation from a logarithmic Gaussian distribution. All parameters of sized magnetite are intermediate between those of a logarithmic Gaussian distribution and those of an exponential distribution.

Some aspects of component analysis

The result of a component analysis depends upon the PDF chosen to model the end-member coercivity distributions, and particularly on the number of parameters assigned to each PDF. Strong differences exist between the results obtained with a linear combination of [Gaussian distributions](#) on the one hand, and a linear combination of [SGG distributions](#) on the other. Since finite mixture models with non-Gaussian coercivity distributions have not been reported in the literature, it is not possible to decide from a-priori informations which kind of PDF should be used as a basis for a [finite mixture model](#). From the mathematical point of view, all PDFs are equivalent, since the [goodness of fit](#) which can be reached with a particular model depends only upon the total number of parameters assumed, regardless of how they are assigned to individual components. Generally, the use of few PDFs with more distribution parameters, instead of a large number of distributions with fewer distribution parameters leads to results of the fitting model which are more stable against measurement errors. The stable behavior of a fitting with SGG distributions can be explained by the fact that small deviations from an ideal coercivity distribution, which arise from measurement errors, are taken into account by variations in skewness and kurtosis, rather than by variations in the contributions of the single components. Obviously, the values obtained for skewness and kurtosis may not be significant at all. A similar stability can be obtained with Gaussian functions if some of them are grouped as if they were one component. However, it is not always evident which distributions group together, and multiple solutions are often possible. The aspects discussed above are illustrated with the examples of [Fig. 3](#) and [Fig. 4](#). Both figures show the results of a component analysis performed with GECA on the coercivity distribution of a sample of urban particulate matter. In [Fig. 3](#), the component analysis is performed with logarithmic Gaussian functions. Four logarithmic Gaussian functions are needed to fit the measured data so that the misfit between model and data is compatible the measurement errors. However, it is impossible to identify these four distributions with an equivalent number of magnetic components. In [Fig. 4](#), the component analysis is performed with SGG functions. The measurements are already well fitted with one SGG function, however, the measured and the modelled coercivity distributions differ significantly. This model could be adequate to describe low-precision measurements of the same sample. Two SGG functions fit the data within the margins given by the measurement errors. However, multiple solutions are possible, but only one solution minimize the difference between model and measurements. The other solutions imply rather uncommon shapes for the coercivity distribution of the individual components, which are not likely occur in natural samples.

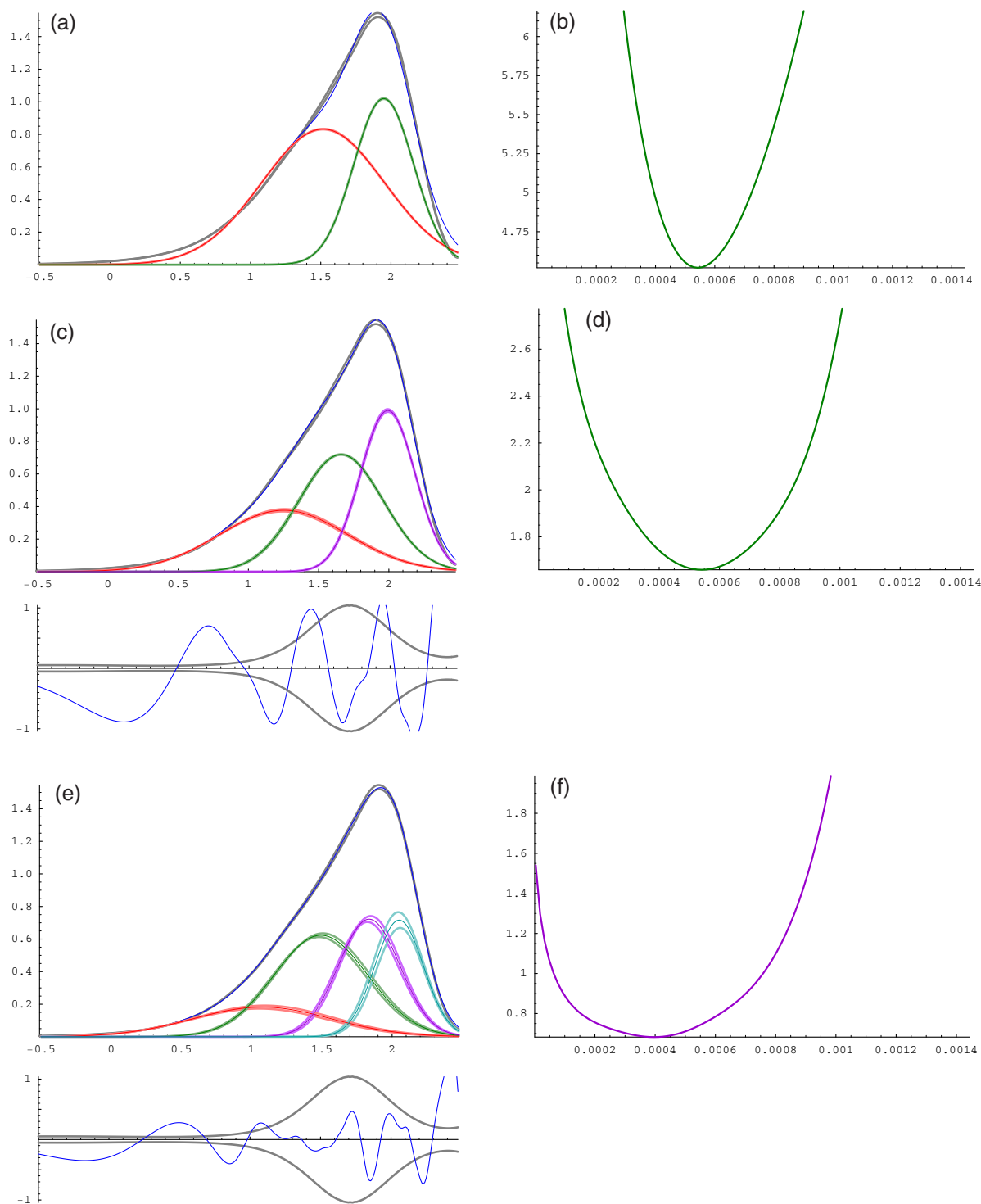


Fig. 3: Component analysis on a sample of urban atmospheric dust collected in Zürich, Switzerland. The component analysis is performed with logarithmic Gaussian functions. Results of the component analysis are shown in (a), (c) and (e). The gray pair of line indicates the confidence limits of the measured coercivity distribution. The blue line is the modelled coercivity distribution, expressed as the sum of the logarithmic Gaussian distributions (red, green, violet and light blue). Confidence limits are plotted around each function. Below each plot, the difference between measured and modelled curve is drawn in blue; the gray pair of curves indicates the amplitude of the measurement errors. The mean quadratic residuals of each model are plotted in (b), (d) and (f) as a function of the amplitude of the logarithmic Gaussian function labelled with the same color. The solutions plotted in (a), (c) and (e) represents the absolute minima of (b), (d) and (f).

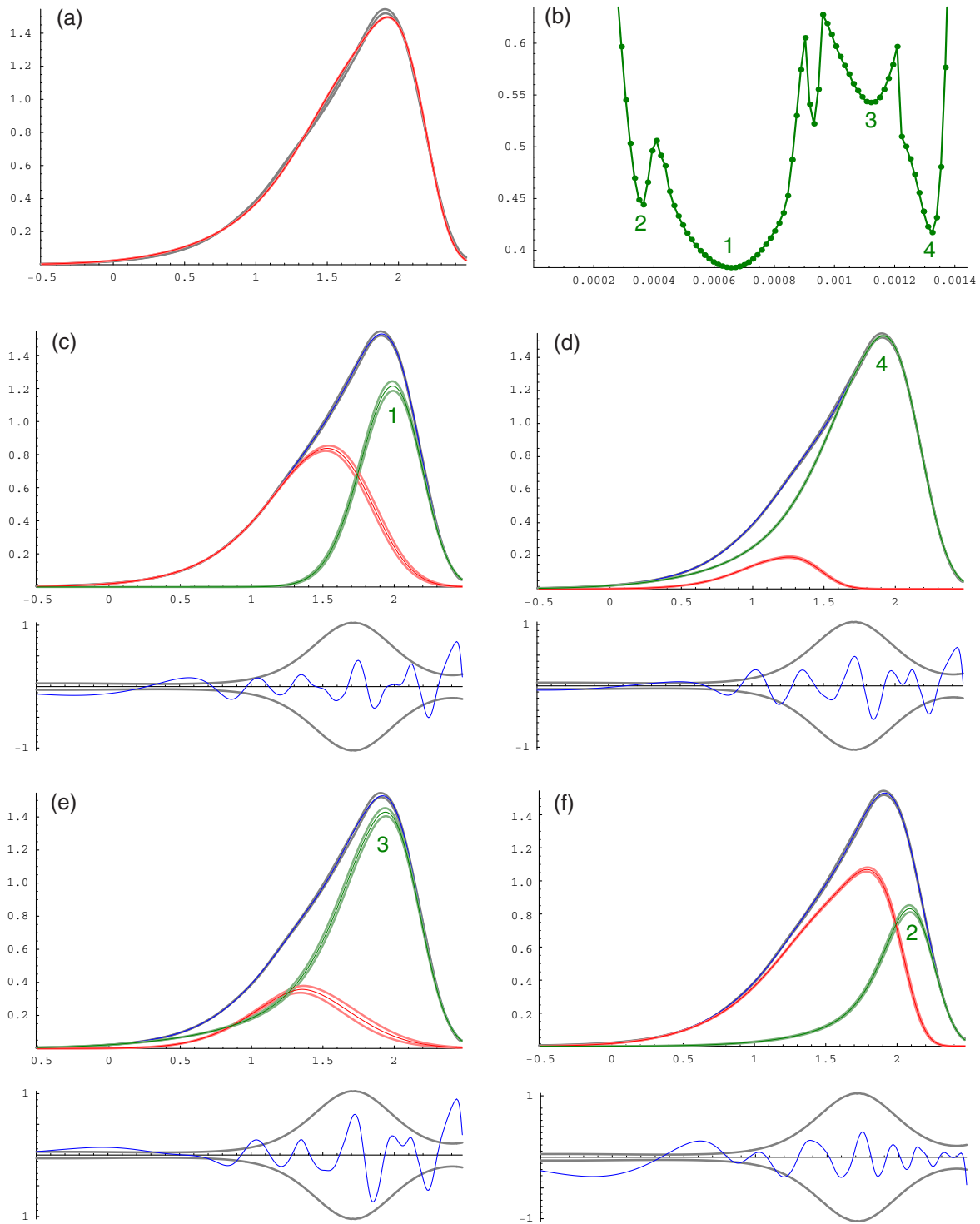


Fig. 4: Component analysis of the same sample as in Fig. 3. The component analysis is performed with SGG functions. The same notation as in Fig. 3 is used for the plots. (a) Component analysis with one SGG function. The modelled coercivity distributions is significantly different from the measured distribution. (b) Mean quadratic residuals of a model with two SGG functions, plotted as a function of the amplitude of one function. Different local minima which correspond to stable solutions of the component analysis are labelled with numbers. The corresponding solutions are plotted in (c), (d), (e) and (f). The solution plotted in (c) corresponds to the global minimum of (b) and the resulting components are compatible with the coercivity distributions of natural dust (red), and combustion products of motor vehicles (green). The solutions corresponding to the other local minima of (b) are not realistic.

The fundamental questions related to component analysis are:

- How many components are needed to fit a given coercivity distribution?
- Are multiple solutions possible? If yes, which solution is correct?

The answer to these questions is not simple. In the example of Fig. 4 the number of components and the identification of the correct solution among multiple solutions is evident. However, this is not always possible, especially if good measurements are not available, or if the coercivity distributions of individual components are too widely overlapped. In this case, some additional information is needed to put appropriate constraints to the number of end-members and to their distribution parameters.

Performing and testing a component analysis

When component analysis is performed, a modelled coercivity distribution $f(x | \boldsymbol{\theta})$ with parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ is compared with the measured coercivity distribution, given by a set of numerical values $(x_i, f_i \pm \delta f_i)$ with measurement errors δf_i . A solution of the component analysis is represented by a set of values of $\boldsymbol{\theta}$ which minimizes a so-called *merit function* $\varepsilon(\boldsymbol{\theta})$. The merit function is an estimation of the difference between the modelled and the measured curve: $\varepsilon = 0$ if the model is identical with the measurements. Examples of $\varepsilon(\boldsymbol{\theta})$ are the mean squared residual:

$$d^2(\boldsymbol{\theta}) = \sum_{i=1}^N [f(x_i | \boldsymbol{\theta}) - f_i]^2 \quad (4)$$

used for a least-squares fitting, and the χ^2 estimator:

$$\chi^2(\boldsymbol{\theta}) = \sum_{i=1}^N \left[\frac{f(x_i | \boldsymbol{\theta}) - f_i}{\delta f_i} \right]^2 \quad (5)$$

used for a minimum χ^2 fitting. GECA uses following weighted version of the χ^2 estimator:

$$w^2(\boldsymbol{\theta}) = \sum_{i=1}^N r_i^{-2} \left[\frac{f(x_i | \boldsymbol{\theta}) - f_i}{\delta f_i} \right]^2 \quad (6)$$

where $r_i = \delta f_i / f_i$ are the relative errors. In this case, measurement points affected by a large relative error are less considered for the component analysis. Equation (6) can be rewritten as:

$$w^2(\boldsymbol{\theta}) = \sum_{i=1}^N \left(\frac{\delta f_i}{\sqrt{f_i}} \right)^{-4} [f(x_i | \boldsymbol{\theta}) - f_i]^2 \quad (7)$$

If $f(x)$ originates from the sum of a finite number of elementary contributions, f_i is a Poisson distributed variable, and $(\delta f_i)^2 \propto f_i$. An experimental confirmation of this assumption is shown in Fig. 5. After these considerations, $w^2(\boldsymbol{\theta}) \propto d^2(\boldsymbol{\theta})$ and $d^2(\boldsymbol{\theta})$ is used by GECA as an improved merit function with respect to (6), since the randomizing effect of the measurement errors on the weighting factors r_i is removed.

Generally, the merit function $\varepsilon(\boldsymbol{\theta})$ has several local minima $\varepsilon_{\min} = \varepsilon(\boldsymbol{\theta}_{\min})$, which correspond to stable solutions $\boldsymbol{\theta}_{\min}$ of the component analysis. Among these minima, there is an absolute minimum $\varepsilon_{\text{MIN}} = \varepsilon(\boldsymbol{\theta}_{\text{MIN}})$. Depending on the starting values $\boldsymbol{\theta}_{\text{ini}}$ of $\boldsymbol{\theta}$, one of these solutions is attained by GECA.

If the model used for component analysis is adequate and if there are no measurement errors, $\varepsilon_{\text{MIN}} = 0$. Let n be the number of magnetic components and m the number of end-member functions used in the model. Then, $\varepsilon_{\text{MIN}} > 0$ for $m < n$ and $\varepsilon_{\text{MIN}} = 0$ for $m \geq n$, so that the number of components can be easily guessed (Fig. 6a). In case of an inadequate model, the end-member functions cannot reproduce exactly the coercivity distribution of all magnetic components, and $\varepsilon_{\text{MIN}} > 0$, even without measurement errors.

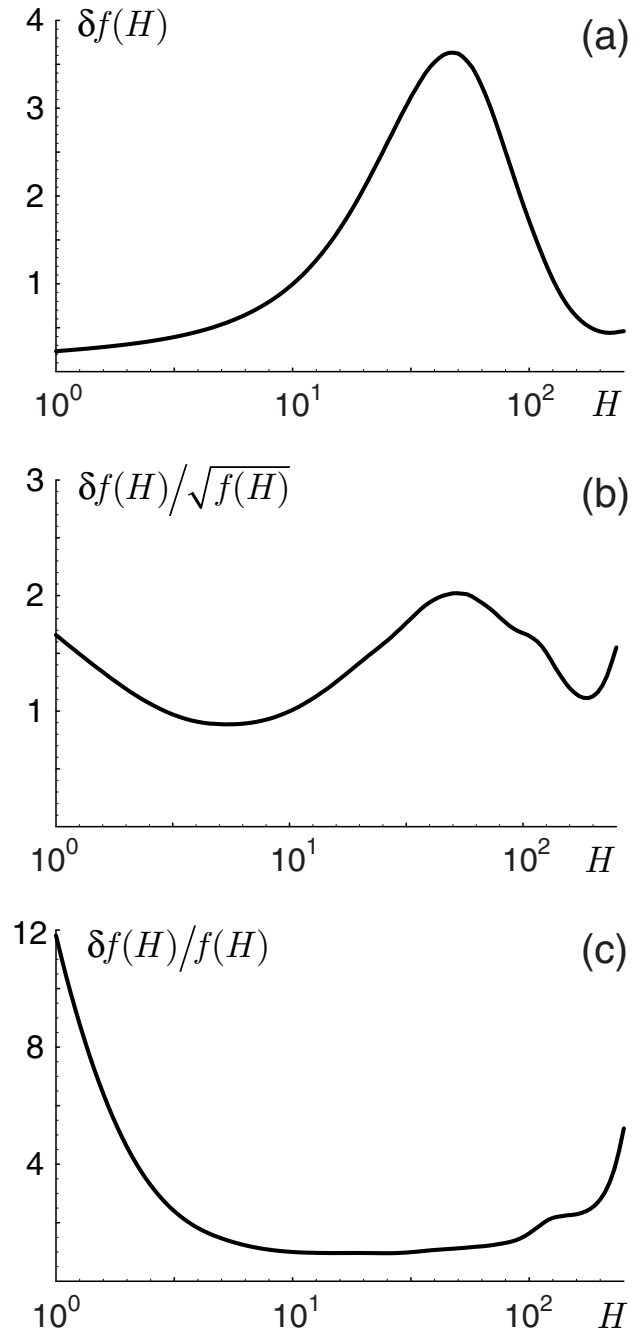


Fig. 5: Mean measurement error of the coercivity distribution of six samples of loess, soil, lake sediments, marine sediments and atmospheric particulate matter. The absolute error $\delta f(H)$ and the relative error $\delta f(H)/f(H)$ are plotted in (a) and (c), respectively. In (b), the absolute error is normalized by the square root of $f(H)$. The field unit is mT. All curves are normalized by their value at 10 mT.

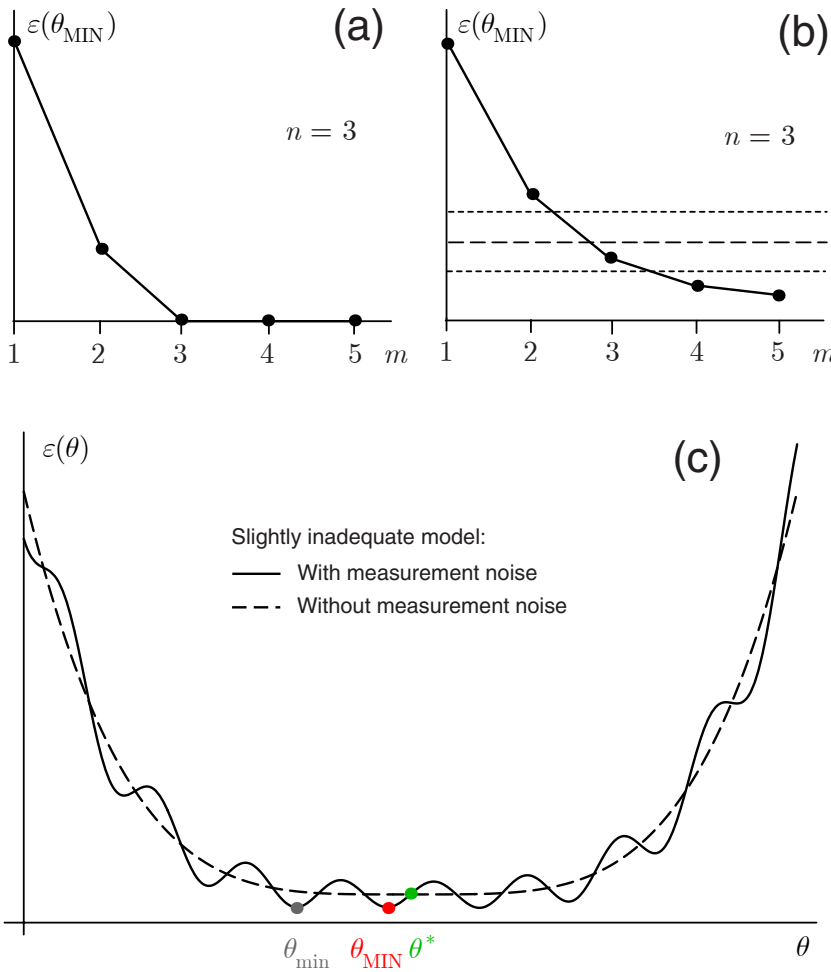


Fig. 6: Dependence of the merit function $\varepsilon(\theta)$ on the parameters of the model chosen for fitting a coercivity distribution. In (a) a noise-free coercivity distribution with $n = 3$ magnetic components is fitted with an adequate model with m end-member functions. The functions are assumed to reproduce exactly the coercivity distribution of each component. If $m < n$, some components cannot be considered into the model and $\varepsilon(\theta) > 0$; on the other hand $\varepsilon(\theta) = 0$ for a given combination $\theta = \theta_{\text{MIN}}$ of parameters when $m > n$. The number of components can be easily guessed. The situation becomes more complex in (b), where measurement errors are taken into account. In this case, there is always a misfit between model and measurements,

and $\varepsilon(\theta_{\text{MIN}})$ decreases monotonically as the number of end-members taken into account by the model is increased. In this case, the number of components is guessed with the help of a Pearson's χ^2 test. According to this test, $\varepsilon(\theta_{\text{MIN}})$ is compared with the expected value of ε (dashed line). If $\varepsilon(\theta_{\text{MIN}})$ is compatible with the expected value within given confidence limits (dotted lines), the model is accepted. If $\varepsilon(\theta_{\text{MIN}})$ is too large, the modelled coercivity distribution is significantly different from the measured coercivity distribution and more parameters should be included in the model. On the other hand, if $\varepsilon(\theta_{\text{MIN}})$ is too small, the model fits the measured data unrealistically well and random effects produced by the measurement errors are included in some parameters which are not significant. The model is accepted if $\varepsilon(\theta_{\text{MIN}})$ belongs to the range of values given by the confidence limits. The complex dependence of the merit function on the model parameters is illustrated in (c) for the case of a model with a fixed number of end-member functions which approximatively fit the coercivity distribution of all magnetic components. These end-member functions produce a small misfit between model and data, even if the measurement errors are not considered (dashed curve). Nevertheless, there is only one stable solution θ^* of the component analysis (green point), which corresponds to an absolute minimum of $\varepsilon(\theta)$. If the measurement errors are taken into account, the shape of $\varepsilon(\theta)$ becomes rather complex, with numerous local minima $\varepsilon(\theta_{\text{min}})$. Some of these local minima represent possible solutions which fits the measurements as good as the absolute minimum $\varepsilon(\theta_{\text{MIN}})$, even if they do not model the coercivity distribution of the real components. The absolute minimum (red point) represents a solution θ_{MIN} which is still close to the reality. With larger measurement errors, this could not be the case, and a realistic solution may be given by a local minimum of $\varepsilon(\theta)$.

If an adequate model is used to fit data affected by measurement errors, $\varepsilon_{\text{MIN}} > 0$, and $\varepsilon_{\text{MIN}} \rightarrow 0$ for $m \rightarrow \infty$ (Fig. 6b).

Two fundamental questions arise at this point:

- 1) How many end-member distributions should be considered for a component analysis?
- 2) Is a particular solution $\hat{\theta}$ close to the (unknown) real solution θ^* ?

These questions can be easily answered only if the model chosen for the component analysis is adequate and the measurement errors are sufficiently small. The first condition can be approximately attained by using a set of SGG functions to model the coercivity distributions of the magnetic components. SGG functions are able to reproduce all fundamental characteristics of the coercivity distribution of a single component ([median](#), [dispersion parameter](#), [skewness](#) and [kurtosis](#)).

If the measurement errors are small enough, the solution θ_{MIN} which corresponds to a global minimum of $\varepsilon(\theta)$ is close to the real solution θ^* (Fig. 6c). In case of large measurement errors, the real solution θ^* may be close to one or more a local minima of $\varepsilon(\theta)$. In this case, additional independent informations are needed to individuate the correct solution among all possible solutions θ_{min} .

The problem of the number of end-members to consider for a component analysis is evaluated with a *Pearson's χ^2 goodness of fit test*. To perform this test, the statistical distribution of the χ^2 estimator given in equation (5) is considered. The χ^2 estimator is a statistical variable which is distributed according to a χ^2 distribution with $N - k - 1$ degrees of freedom, being N the number of *independent* points to fit with a given model, and k the number of model parameters. The expected value of the χ^2 estimator is $N - k - 1$. The confidence limits at a confidence level α (generally $\alpha = 0.95$) are given by $\chi_{N-k-1;\alpha}^2$ and $\chi_{N-k-1;1-\alpha}^2$, with:

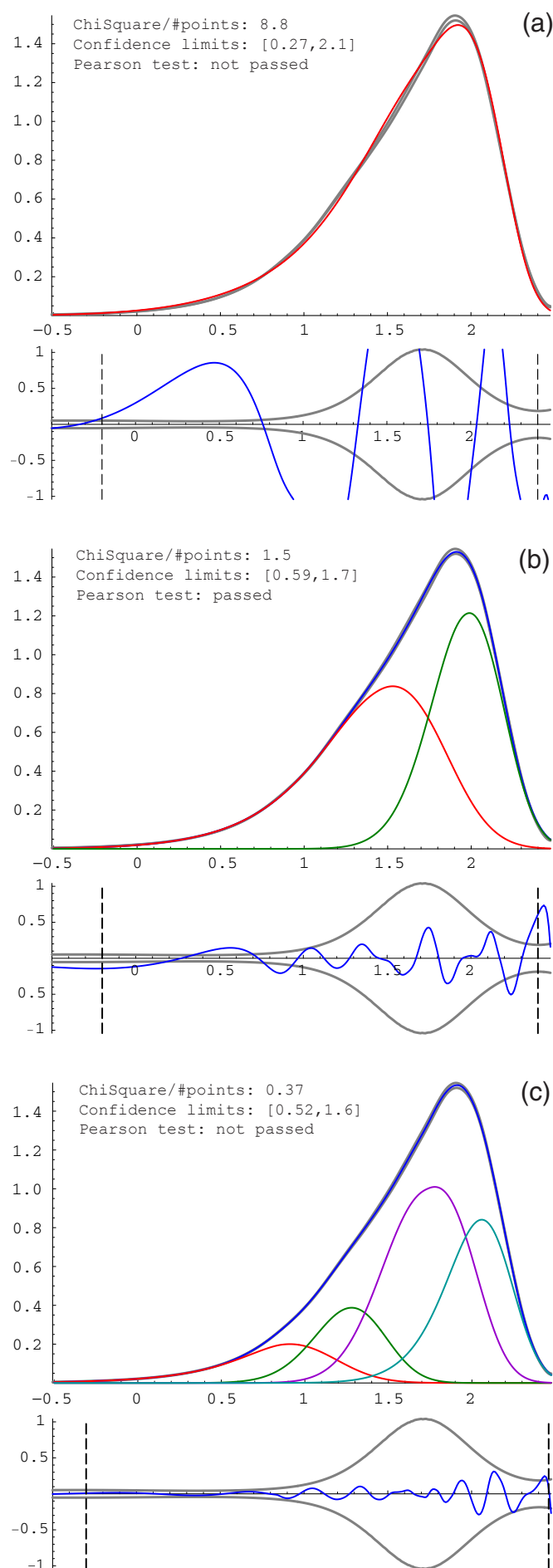
$$\int_{\chi_{N-k-1;p}^2}^{\infty} \chi_{N-k-1}^2(t) dt = p \quad (8)$$

If $\chi^2 > \chi_{N-k-1;1-\alpha}^2$, the model differs significantly from the measurements. The model should be refined by adding new parameters, eventually by considering an additional end-member function. If $\chi^2 < \chi_{N-k-1;\alpha}^2$ the differences between model and measurements are unrealistically small. An excessive number of parameters allow the model to include random effects of the measurement errors. Consequently, some of these parameters are not significant. The model should be revised to include a smaller number of parameters, eventually by reducing the number of end-members or by keeping some parameters fixed. If $\chi_{N-k-1;\alpha}^2 \leq \chi^2 \leq \chi_{N-k-1;1-\alpha}^2$ the model is acceptable.

To calculate the χ^2 estimator with equation (5) some knowledge about the measurement errors δf_i and the number of independent data points is necessary. The measurement errors are automatically estimated with CODICA, when a coercivity distribution is calculated from an acquisition/demagnetization curve. The number of independent data points is more difficult to estimate. It is identical with the number of measurements if the measurement errors are equivalent to an ergodic noise signal, that is, when the autocorrelation of the noise signal is equivalent to a Dirac δ -function.

This is often not the case with real measurements, where entire groups of measured points are affected by the same error. Furthermore, the coercivity distributions calculated by CODICA are low-pass filtered, and an autocorrelation of the remaining measurement errors is unavoidable. GECA estimates the degrees of freedom of the fitting model by evaluating the residuals curve which results from the difference between the model and the measurements. The residuals curve contains a certain number of random oscillations around a mean value of zero. To reproduce these oscillations a minimum number l of points is necessary, whose spacing defines the Nyquist frequency of the signal. GECA sets $l - 1$ equal to the number of zero crosses of the residuals. Obviously, the shape of the residuals curve depends on the model chosen for component analysis.

Fig. 7: Examples of Pearson's χ^2 test on the component analysis of a sample of urban atmospheric particulate matter. The gray and the blue curves are the measured and the modelled coercivity distributions, respectively. Curves labelled with other colors represent the coercivity distributions of individual end-members. Below each plot, the difference between model and measurements is plotted (blue line) together with the measurement errors (pair of gray lines). In (a), a model with one SGG function is evaluated. The differences between model and measurements are too large, and the model is rejected. In (c) the a model with four SGG functions is rejected for the opposite reason: the model fits the data unrealistically well for the given measurement errors. A model with two SGG functions is represented in (b). In this case, the χ^2 statistics is compatible with the expected value within at a 95% confidence level, and the model is accepted.



A model with a small number of parameters produces a residuals curve with few, large oscillations. The more parameters are included in the model, the more oscillations characterize the residuals and the confidence limits of the χ^2 estimator become closer to the expected value. Consequently, models with a too large number of parameters are rejected. An example of Pearson's χ^2 test is shown in [Figure 7](#) with the example of a sample of urban atmospheric particulate matter. In Fig. 7a, the coercivity distribution is fitted with one SGG function. The residuals curve has 5 zero crosses in the range considered for fitting, and GECA assumes $l = 6$ degrees of freedom for the χ^2 distribution. The confidence limits of χ^2/l are 0.27 and 2.1, while $\chi^2/l = 8.8$ for that model, which is rejected. In Fig. 7b, two SGG functions are used for the component analysis. Now, $l = 12$, and the confidence limits of χ^2/l are 0.44 and 1.8. With $\chi^2/l = 1.5$ this model is acceptable. With four SGG functions (Fig. 7c), $l = 18$ and the confidence limits of χ^2/l are 0.52 and 1.6, while $\chi^2/l = 0.37$ for that model, which is rejected.

A program example

`In[1]:= <<Utilities`Codica``

[Load the program](#)

Program package Codica v.2.3 for Mathematica 3.0 and later versions.
Copyright 2000-2003 by Ramon Egli. All rights reserved.

`In[2]:= Geca`

[Start the program](#)

Data from file

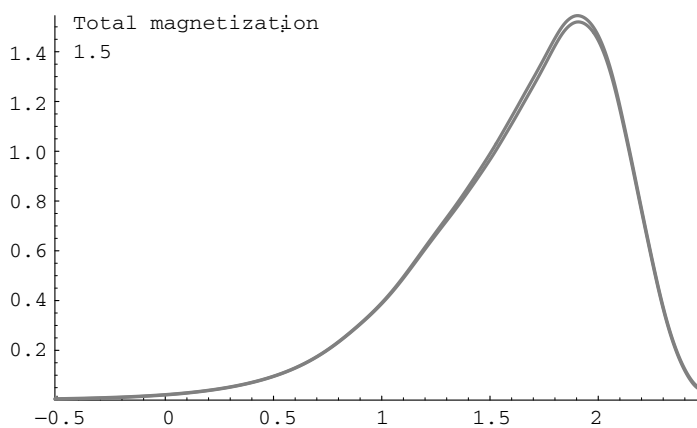
[Enter file name](#)

`C:/users/ramon/papers/fitting/WDKarm.slog`

Checking the coercivity distribution...

Confidence limits of the coercivity distribution:

[Plot the distribution](#)

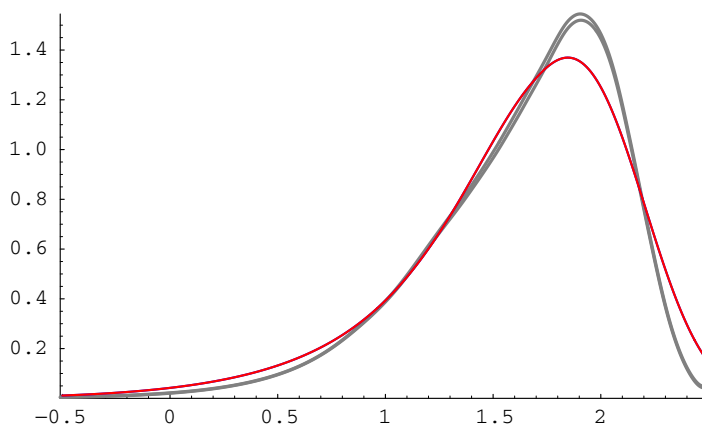


Coercivity distribution is significant between -0.5 and 2.474

Fitting is performed in the range between -0.1957 and 2.396

[Set the fitting range](#)

[Enter initial parameters \(1\)](#)



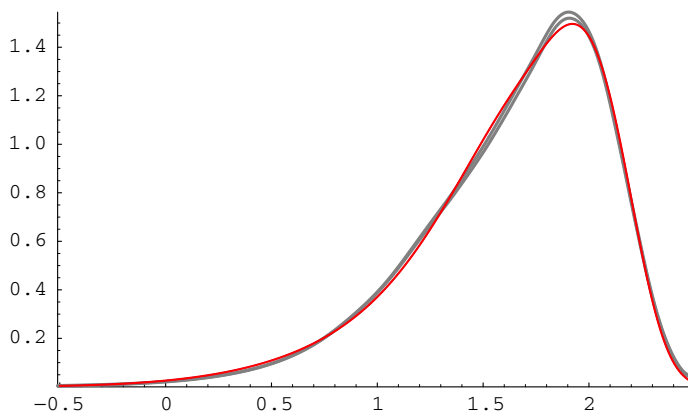
Initial distribution parameters:

a1 = 1.5
 m1 = 1.7 s1 = 0.6
 q1 = 0.5 p1 = 2.2

Optimizing the distribution parameters. Please wait...

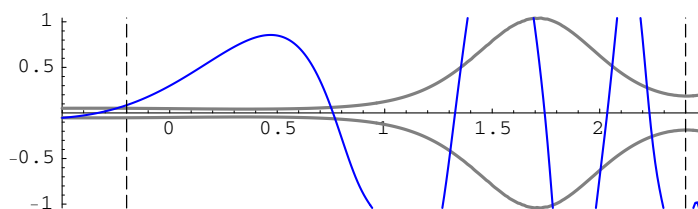
[Perform a component analysis \(1\)](#)

{a1=0.0015, m1=1.7, s1=0.6, q1=0.5, p1=2.2}



Optimized distribution parameters:

a1 = 1.454
 m1 = 1.714 s1 = 0.6875
 q1 = 0.4256 p1 = 3.076



Residuals (modelled - measured) and measurement errors

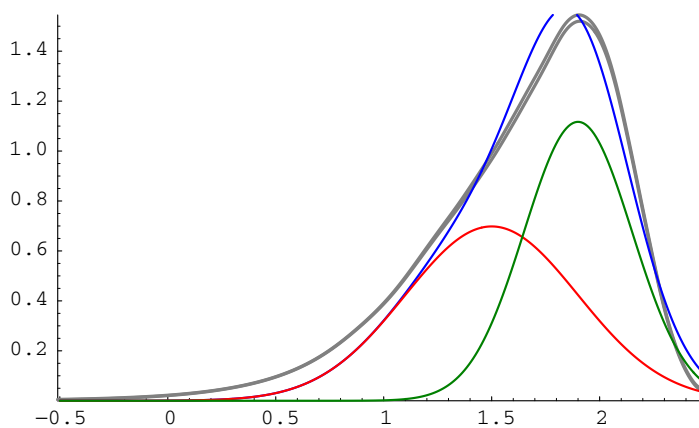
in % of the maximum value of the coercivity distribution.

Dashed lines delimitate the interval considered for the component analysis.

ChiSquare/#points: 8.8
 Confidence limits: [0.52,1.9]

Model and data are significantly different. Refine your model.

[Enter initial parameters \(2\)](#)



Initial distribution parameters:

```

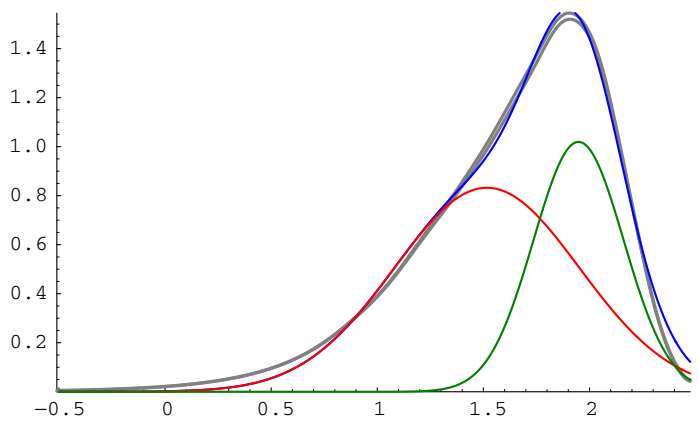
a1 = 0.7
m1 = 1.5      s1 = 0.4
q1 = 1.      p1 = 2.

a1 = 0.7
m1 = 1.9      s1 = 0.25
q1 = 1.      p1 = 2.
    
```

Optimizing the distribution parameters. Please wait...

[Perform a component analysis \(2\)](#)

```
{a1=0.0007,m1=1.5,s1=0.4,a2=0.0007,m2=1.9,s2=0.25}
```

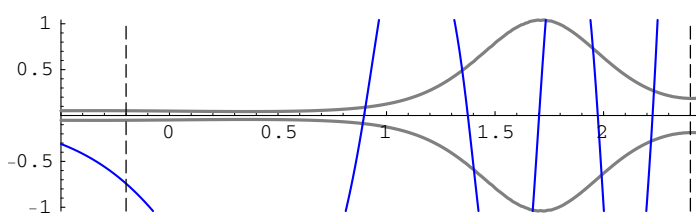


Optimized distribution parameters:

```

a1 = 0.9101
m1 = 1.517   s1 = 0.4363
q1 = 1.      p1 = 2.

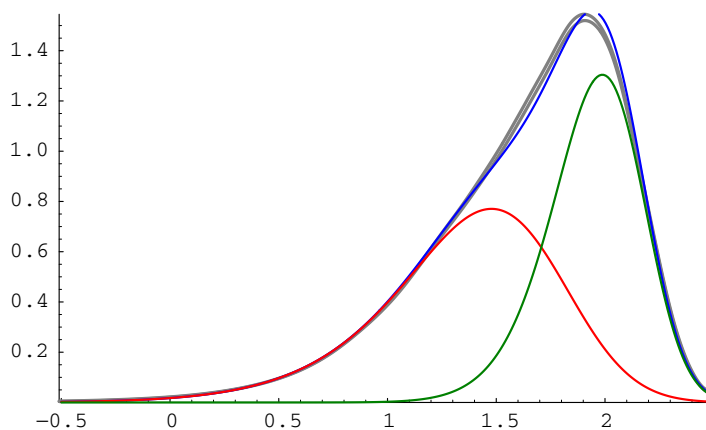
a1 = 0.5428
m1 = 1.948   s1 = 0.2124
q1 = 1.      p1 = 2.
    
```



Residuals (modelled - measured) and measurement errors in % of the maximum value of the coercivity distribution. Dashed lines delimitate the interval considered for the component analysis. ChiSquare/#points: 27. Confidence limits: [0.52,1.9]

Model and data are significantly different. Refine your model.

[Enter initial parameters \(3\)](#)



Initial distribution parameters:

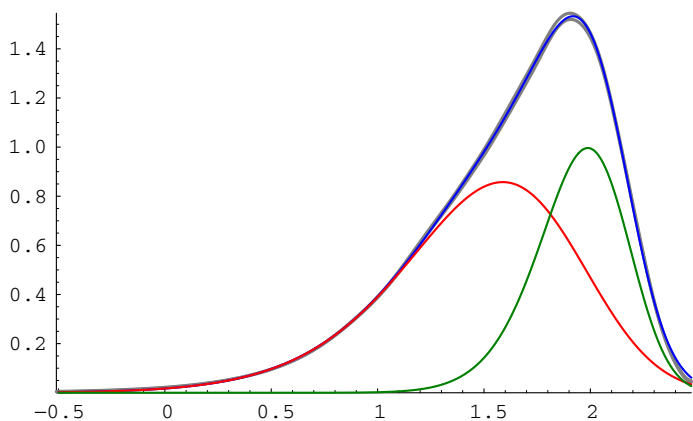
a1 = 0.75
 m1 = 1.4 s1 = 0.45
 q1 = 0.6 p1 = 2.

a1 = 0.7
 m1 = 1.957 s1 = 0.235
 q1 = 0.663 p1 = 2.

Optimizing the distribution parameters. Please wait...

[Perform a component analysis \(3\)](#)

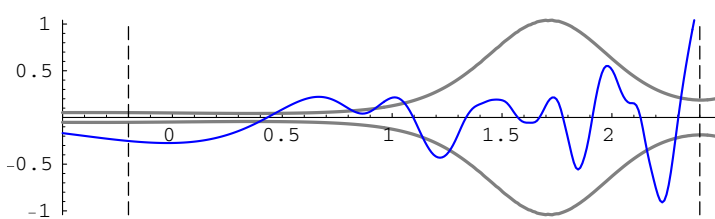
{a1=0.00075,m1=1.4,s1=0.45,q1=0.6,p1=2.,a2=0.0007}



Optimized distribution parameters:

a1 = 0.9217
 m1 = 1.509 s1 = 0.4882
 q1 = 0.6235 p1 = 2.023

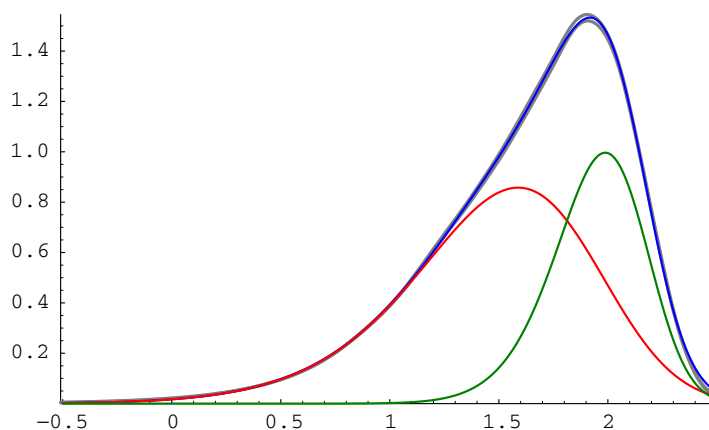
a1 = 0.5348
 m1 = 1.957 s1 = 0.235
 q1 = 0.663 p1 = 2.



. Residuals (modelled - measured) and measurement errors in % of the maximum value of the coercivity distribution. Dashed lines delimitate the interval considered for the component analysis. ChiSquare/#points: 2.8 Confidence limits: [0.63,1.6]

Model and data are significantly different. Refine your model.

[Enter initial parameters \(4\)](#)



Initial distribution parameters:

```

a1 = 0.9217
m1 = 1.509   s1 = 0.4882
q1 = 0.6235  p1 = 2.023

a1 = 0.5348
m1 = 1.957   s1 = 0.235
q1 = 0.663   p1 = 2.
    
```

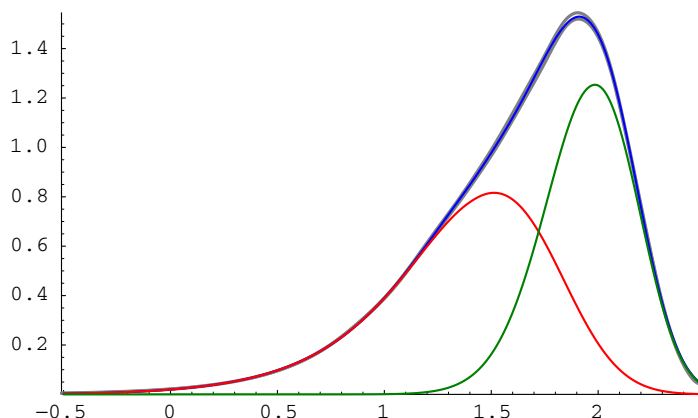
Optimizing the distribution parameters. Please wait...

[Perform a component analysis \(4\)](#)

```
{a1=0.0009217,m1=1.509,s1=0.4882,q1=0.6235,p1=2.023,
 a2=0.0005348,m2=1.957,s2=0.235,q2=0.663,p2=2.}
```

FindMinimum::fmlim: The minimum could not be bracketed in 50 iterations.

```
{a1=0.0006243,m1=1.311,s1=0.4483,q1=0.4878,p1=2.171,
 a2=0.0008625,m2=1.963,s2=0.2331,q2=0.7765,p2=2.107}
```

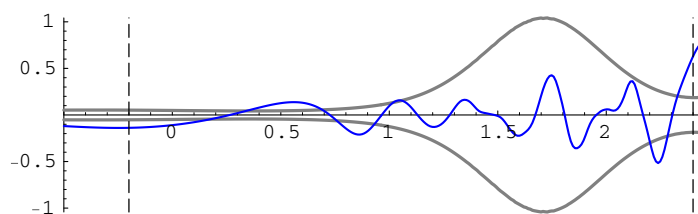


Optimized distribution parameters:

```

a1 = 0.7699
m1 = 1.412   s1 = 0.4699
q1 = 0.5537  p1 = 2.106

a1 = 0.6847
m1 = 1.962   s1 = 0.2329
q1 = 0.7221  p1 = 2.053
    
```



Residuals (modelled - measured) and measurement errors in % of the maximum value of the coercivity distribution. Dashed lines delimitate the interval considered for the component analysis.
 ChiSquare/#points: 1.5
 Confidence limits: [0.59,1.7]

Model and data are compatible. You may accept this component analysis.

[Systematic solution search](#)

Perform an automatic variation of the contribution of component #2:

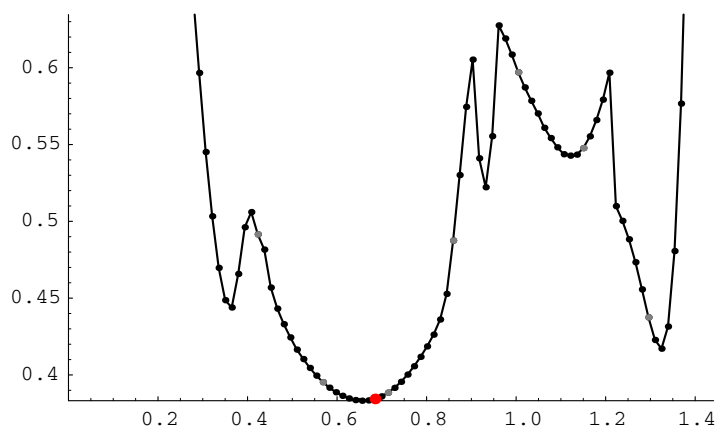
This process takes several minutes. Please wait...

Decreasing contribution of component #2...

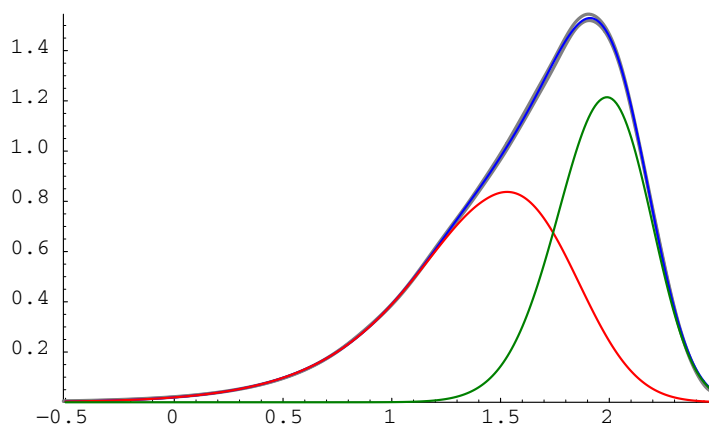
Increasing contribution of component #2...

Residuals as a function of the contribution of component #2

(Every 10th point in gray, first point is #21, red point is the starting solution)



[Choose initial parameters](#)



Initial distribution parameters:

```

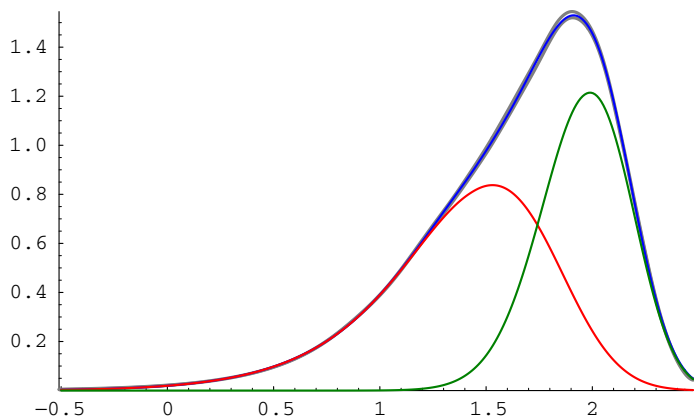
a1 = 0.7989
m1 = 1.428    s1 = 0.4739
q1 = 0.5562   p1 = 2.108

a1 = 0.6556
m1 = 1.967    s1 = 0.2297
q1 = 0.7265   p1 = 2.052
    
```

Optimizing the distribution parameters. Please wait...

[Perform a component analysis \(5\)](#)

```
{a1=0.0007989,m1=1.428,s1=0.4739,q1=0.5562,p1=2.108,a2=0.0006556,
m2=1.967,s2=0.2297,q2=0.7265,p2=2.052}
```

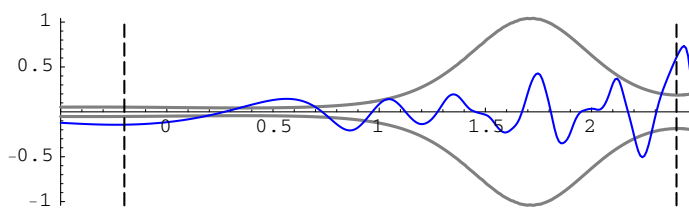


Optimized distribution parameters:

```

a1 = 0.7989
m1 = 1.428    s1 = 0.4739
q1 = 0.5562   p1 = 2.108

a1 = 0.6556
m1 = 1.967    s1 = 0.2297
q1 = 0.7265   p1 = 2.052
    
```



[Residuals \(modelled - measured\) and measurement errors](#)
in % of the maximum value of the coercivity distribution.
Dashed lines delimitate the interval considered for the component analysis.
ChiSquare/#points: 1.5
Confidence limits: [0.59,1.7]

Model and data are compatible. You may accept this component analysis.

Calculating statistical parameters of the distributions...

[Perform an error estimation](#)

Perform an error estimation of the distribution parameters with 64 error simulations.

Accuracy of the error estimation: 12.2%

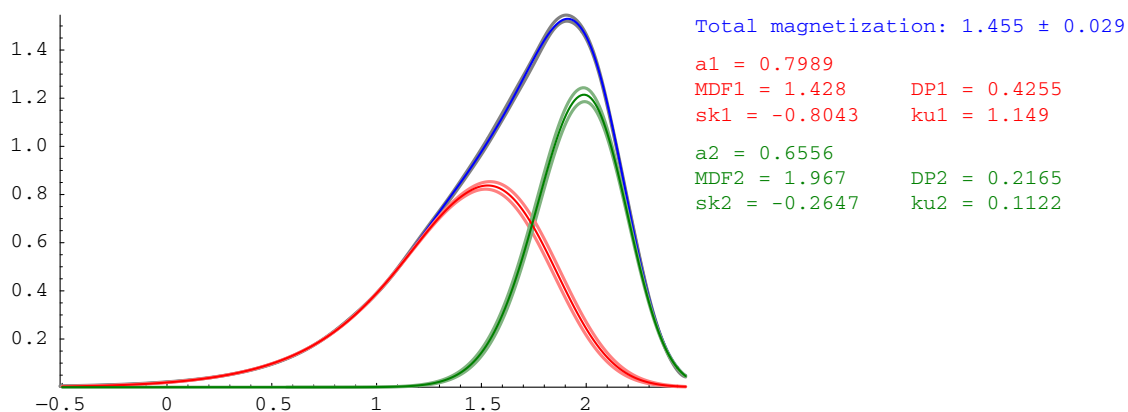
This process takes several minutes time. Please wait...

Error estimation of the statistical parameters. Please wait...

Calculating the confidence limits of the components. Please wait...

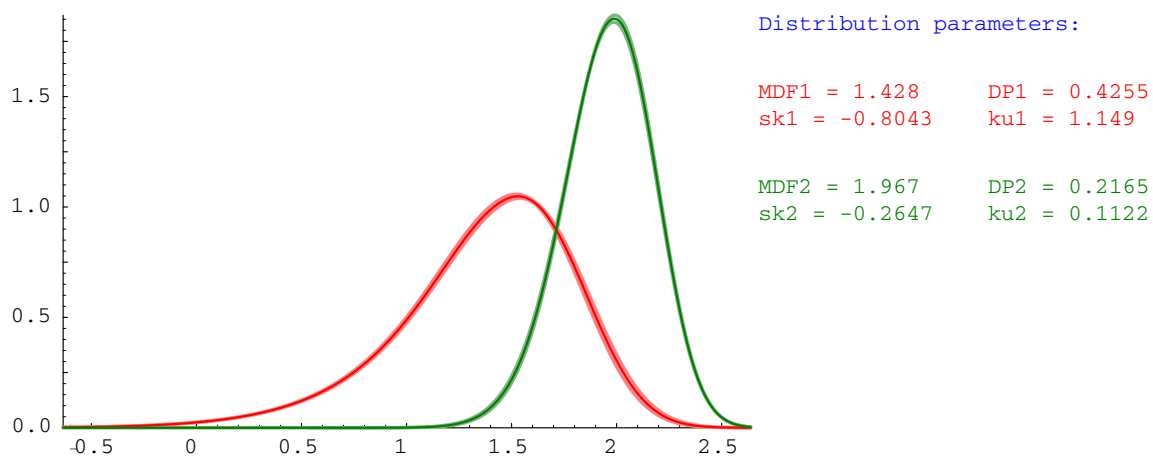
Parameters of component #1:	Parameters of component #2:
a = 0.7989 ± 0.021	a = 0.6556 ± 0.021
μ = 1.428 ± 0.011	μ = 1.967 ± 0.0038
σ = 0.4739 ± 0.0029	σ = 0.4739 ± 0.0029
q = 0.5562 ± 0.0047	q = 0.7265 ± 0.0065
p = 2.108 ± 0.019	p = 2.052 ± 0.0082
MDF = 1.428 ± 0.011	MDF = 1.967 ± 0.0038
mean = 1.374 ± 0.012	mean = 1.957 ± 0.0044
DP = 0.4255 ± 0.0027	DP = 0.2165 ± 0.0024
skewness = -0.8043 ± 0.013	skewness = -0.2647 ± 0.015
kurtosis = 1.149 ± 0.026	kurtosis = 0.1122 ± 0.024

Result of the component analysis:



Calculating the confidence limits of each component. Please wait...

Normalized components with confidence limits:



Preparing data to an export format. Please wait...

[Save results to a log file](#)

Printing results to: components.dat

[Save end-members](#)

Saving the coercivity distributions to WDKarm.comp :
Column #1: magnetic field,
Column #2: component #1 Column #3: error of component #1
Column #4: component #2 Column #5: error of component #2
END

Loading CODICA and running GECA

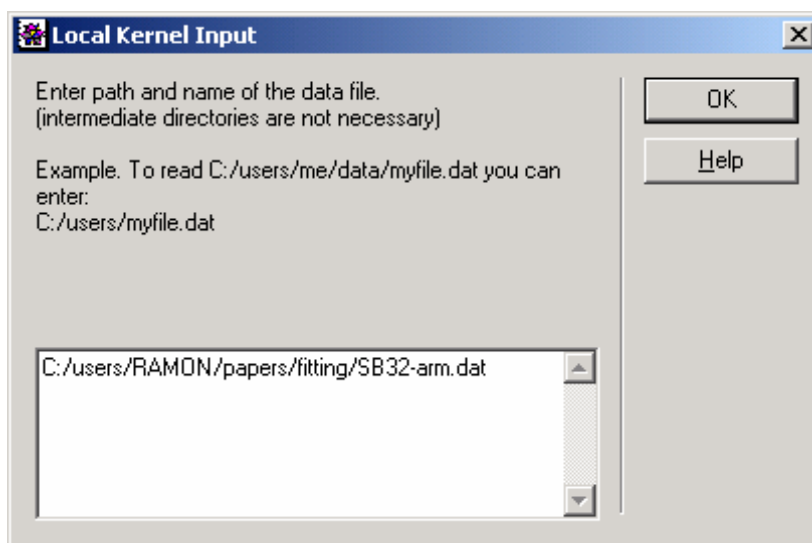
To run GECA, open a new *Mathematica* notebook by clicking on the *Mathematica* program icon. Type `<<Utilities`Codica`` on the input prompt `In[]` and press the keys *Shift + Enter* to load CODICA. On the next input prompt type `Geca` and press the keys *Shift + Enter* to start GECA. From now on, the program asks you to enter specific commands step by step. In the following, all GECA commands are explained in order of appearance.

[Back to the program](#)

Enter the name of the data file

The prompt window on the right asks you to enter the name on the file which contains the coercivity distribution data. Type the path of the data file. You can skip intermediate directories if other files with the same name are not stored. The data file should be an ASCII file with three columns of numbers separated by spaces or tabulators. The file should not contain comment lines or text in general. The first column is the scaled or unscaled field, the second column is the value of the coercivity distribution for the corresponding field. The third column is the relative error of the second column; 0.1 means 10% error. Output files of CODICA with extensions `.slin`, `.slog` and `.spow` are automatically accepted. It is strongly recommended to run GECA only on CODICA output files with extension `.slog`.

[Back to the program example](#)



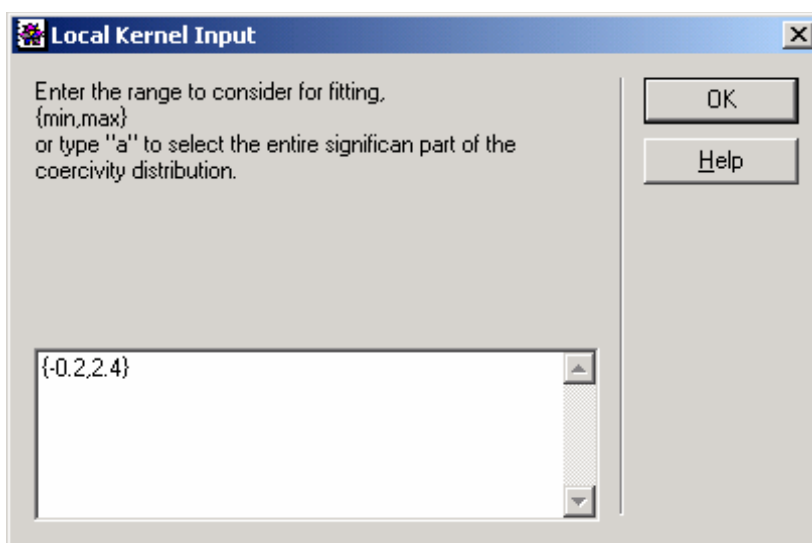
Plot the coercivity distribution

The coercivity distribution is plotted together with the confidence limits given by the error estimation stored in the file. If the maximal measurement error is less than 5% of the peak value of the coercivity distribution, only the confidence limit are plotted as a pair of gray lines. With errors larger than 5%, the coercivity distribution is plotted as a black line, together with the confidence limits. Within the plot, an estimation of the total magnetization is given. This estimation is obtained by integrating the coercivity distribution over the field range given by the data stored in the file. If saturation is not reached within this range, the calculated value is an underestimation of the total magnetization. You can use the estimation of the total magnetization as a reference when you enter the [initial distribution parameters](#).

[Back to the program example](#)

Set the fitting range

GECA estimates a field range where the values of the coercivity distribution are significant. As a significance limit, a maximum relative error of 50% has been chosen for the values of the coercivity distribution. You can enter a different range with the prompt window displayed on the right. If the coercivity distribution was calculated from a demagnetization curve, it is recommended to discard the data near the right end of the field range, because they could be affected by truncation effects. Data outside the range you entered are displayed but are not considered for further calculations.



[Back to the program example](#)

Enter initial distribution parameters

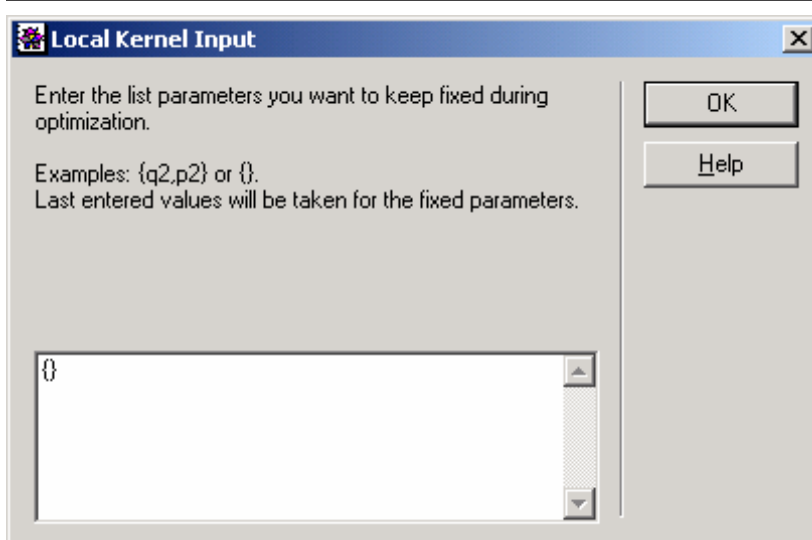
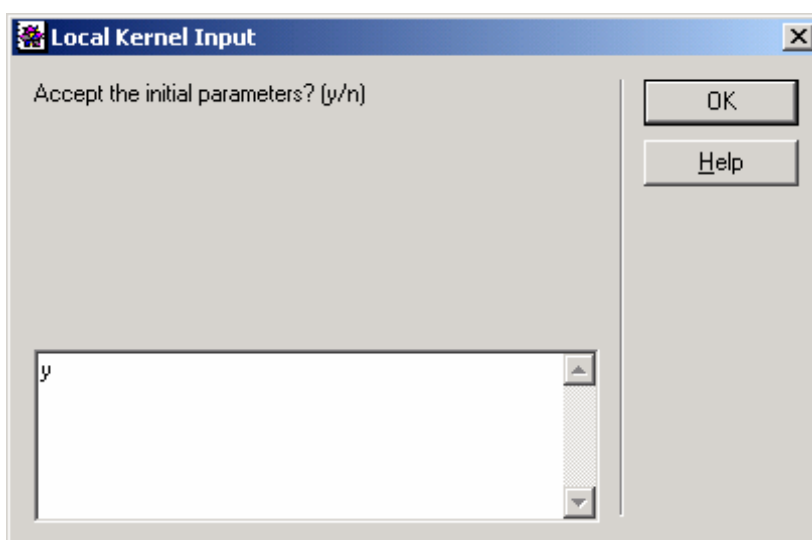
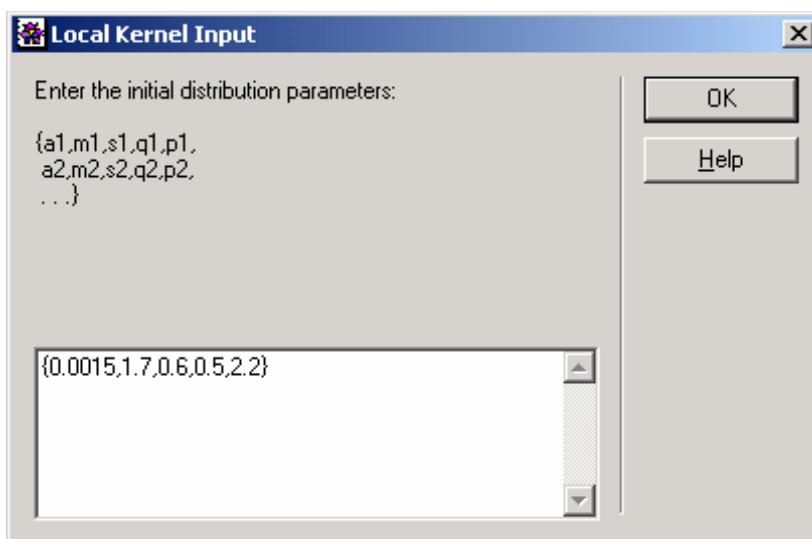
You are asked to enter initial values for the parameters of the [finite mixture model](#) that will be used for the component analysis. GECA uses a set of one to four [SGG functions](#) to fit the measured coercivity distribution. Each SGG function is characterized by following five parameters:

- *amplitude* a : the area under the SGG function, which is equivalent to the magnetization of a component whose coercivity distribution is represented by this function.
- *median* μ : this parameter corresponds to the [median value](#) of the function, also called median destructive field (MDF) or median acquisition field (MAF).
- *parameter for the standard deviation* σ : this is the principal parameter which controls the standard deviation of the SGG function, also called the [dispersion parameter DP](#).
- *parameter for the skewness* q : this is the principal parameter which controls the [skewness](#) of a SGG function, with $-1 \leq q \leq 1$. Positive values of q generate left skewed functions, negative values of q generate right skewed functions. Symmetrical functions are characterized by $q = \pm 1$. Generally, real coercivity distributions are characterized by $0.5 \leq |q| \leq 1$. If you do not have independent informations about the starting parameters, set a vaule of q near 1.
- *parameter for the kurtosis* p : this is the principal parameter which controls the [kurtosis](#) of a SGG function. A logarithmic Gaussian distribution is characterized by $p = 2$. More squared functions are generated with $p > 2$, less squared distributions by $p < 2$. Common values for real coercivity distributions are given by $1.6 < p < 2.4$. If you do not have independent informations about the starting parameters, set $p = 2$.

Enter the parameters as an ordered list: a , μ , σ , q and p of the first component, a , μ , σ , q and p of the second component, and so on, as in the example given in the prompt window shown to the right.

The [end-member distributions](#) defined by the initial parameters you entered are plotted with different colors (red, green, violed and light blue). The modeled coercivity distribution is given by the sum of all end-members and is plotted in blue, together with the measured coercivity distribution (black/gray). The initial parameters should be chosen so, that the modelled coercivity distribution is as close to the measured coercivity distribution as possible. You can enter the initial parameters either with some knowledge about the magnetic components which are contributing to the measured distribution, or by try and error. In this last case you can reenter new initial parameters until you get a satisfying result. After entering the initial parameters, you are asked to keep some parameters fixed during the optimization. If you want to optimize all parameters, type “{}”. Otherwise, enter the symbols for the fixed parameters in the next prompt window. For example, if you want to use a logarithmic Gaussian function for the second end-member, set $q = 1$ and $p = 2$ as initial values for the corresponding SGG function, and keep these parameters fixed by entering “{q2,p2}” in the following prompt window. You can choose every combination of parameters to keep fixed. If you exactly know the parameters of one magnetic component, you can model this component by an end-member function with fixed values of μ , σ , q and p . Then, only the magnetic contribution of this component, given by a , will be optimized.

It is recommended to start with a small number of end-members and a small number of parameters, and to use independent informations about the number of magnetic components and their properties. You can then progressively increase the complexity of your model. Keep in mind that the complexity of a model increases exponentially with the number of parameters to optimize. If you



you exactly know the parameters of one magnetic component, you can model this component by an end-member function with fixed values of μ , σ , q and p . Then, only the magnetic contribution of this component, given by a , will be optimized.

It is recommended to start with a small number of end-members and a small number of parameters, and to use independent informations about the number of magnetic components and their properties. You can then progressively increase the complexity of your model. Keep in mind that the complexity of a model increases exponentially with the number of parameters to optimize. If you

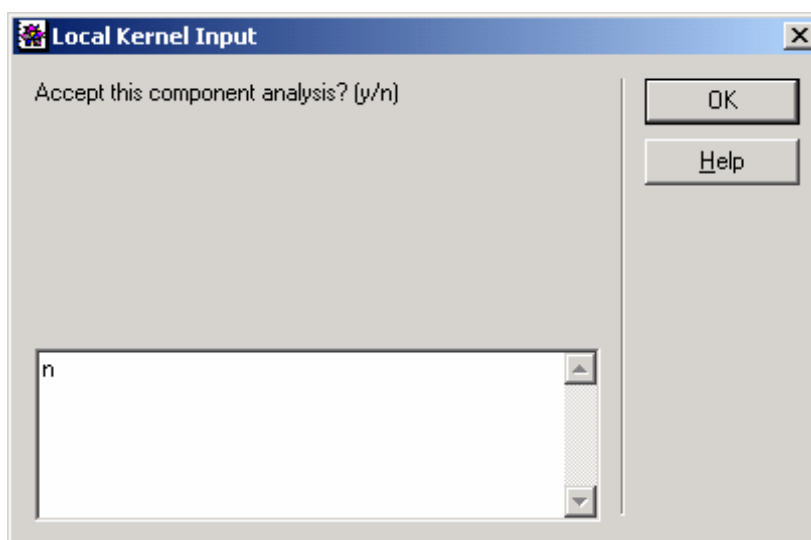
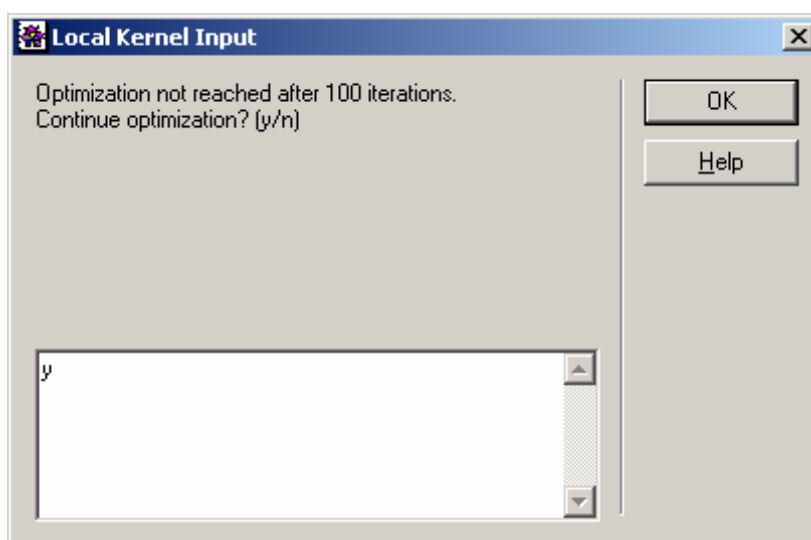
want to perform a component analysis with three SGG functions, you have to deal with a solution space in 15 dimensions. You will not have the possibility to perform a systematic solution search in such a space: if you try 5 initial values for each parameter, you should perform $5^{15} = 3 \times 10^{10}$ optimizations! You would probably find several stable solutions, but only one among them is correct and has a physical meaning. The parameters have a hierarchical structure: a controls the amplitude of an end-member, μ the “position” along the field axis, and σ the “width”, q the symmetry and p the curvature. The amplitude is the most important parameter, the curvature is the less important. You can start with fixed values of p , or fixed values of q and p . Use logarithmic Gaussian functions to model magnetic components which are not saturated in the field range of the measured coercivity distribution.

In the program example, the measured coercivity distribution is similar to an asymmetric unimodal probability density function. There is no direct evidence for more than one magnetic component. Therefore, initial parameters for one SGG function have been entered in the program example. Since only 5 parameters have to be optimized, the component analysis is relatively simple and only one stable solution is expected. Therefore, it is not necessary to start with a modelled coercivity distribution which is very close to the measured data.

[Back to the program example](#)

Perform a component analysis

GECA performs a component analysis by optimizing the initial parameters in a way that minimizes the [squared residuals](#) between model and measurements by using a Levenberg-Marquardt algorithm. The parameters to optimize are displayed together with the corresponding initial values. If the initial values were carefully chosen, the search for a solution is performed in a reasonable time with no more than 100 iterations. Otherwise, the search will take more than 100 iterations or it will converge to an absurd solution. If a global or a local minimum of the squared residuals is not reached within 100 iterations, a warning message appears and you will be asked to continue the search or stop it and plot the solution given by the last iteration. It is recommended to perform at least 200 iterations. The numerical values of the parameters are shown every 100 iterations and you can check how they change and if they converge to a meaningful result.



If a convergence to a stable solution cannot be obtained, interrupt the search for a solution by typing “n” in the prompt window and choose other initial parameters.

The result of the component analysis is displayed exactly like the initial model. The same colors are used to label the end-members. Additionally, the difference between the model and the measurements (blue line) is plotted below the result of the component analysis, together with the measurement errors (pair of gray lines). The difference between model and measurements (called *misfit* in the following) should be of the same order of magnitude as the estimated measurement error. If the misfit is much larger than the estimated measurement error, the model is not able to account for the measurements: other parameters should be added to reduce the misfit. If the misfit is much smaller than the measurement error, the model chosen for the component analysis is able to fit the measurements very well but it is not significant: some of the model parameters do not have any physical meaning. In this case you should decrease the number of model parameters by reducing the number of end-members or by keeping some parameters fixed. If the misfit has the same amplitude as the estimated measurement error, the model may be adequate. Nevertheless, more than one solution which satisfy this condition may exist.

An adequate parameter to test the significance of a component analysis is the χ^2 [statistics](#). GECA gives an estimation of χ^2/l , where l represents the [degrees of freedom](#) of the model. For a correct model, $\chi_{l;\alpha}^2 \leq \chi^2 \leq \chi_{l;1-\alpha}^2$, where $\chi_{l;\alpha}^2$, $\chi_{l;1-\alpha}^2$ are the [confidence limits](#) at a given confidence level (usually 95%). GECA calculates the confidence limits with a 95% confidence level and displays them together with the estimation of χ^2/l . If $\chi^2 > \chi_{l;1-\alpha}^2$, the model is significantly different from the measured data, and GECA suggests you to refine it by adding more parameters. If $\chi^2 < \chi_{l;\alpha}^2$, not all model parameters are significant and a warning message is displayed. In this case you should reduce the number of parameters. If $\chi_{l;\alpha}^2 \leq \chi^2 \leq \chi_{l;1-\alpha}^2$, GECA suggests to accept the model.

In the program example, the component analysis with one SGG function is inadequate to model the measurements within the given error margins.

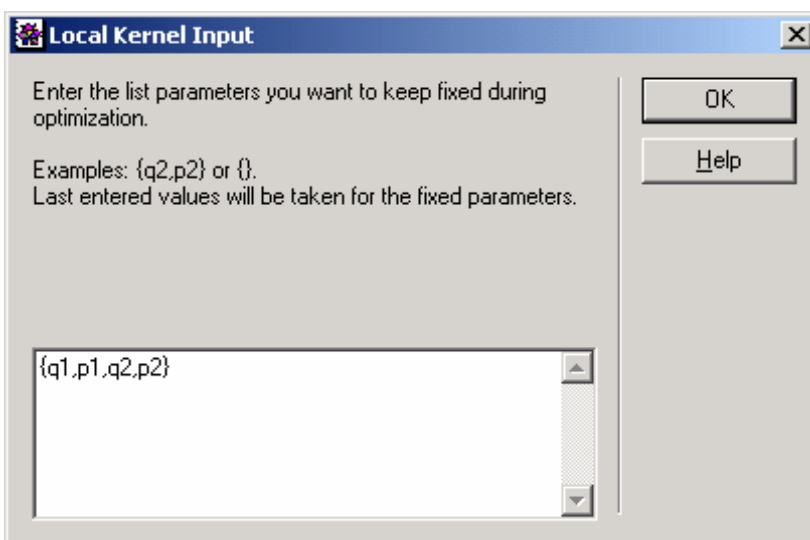
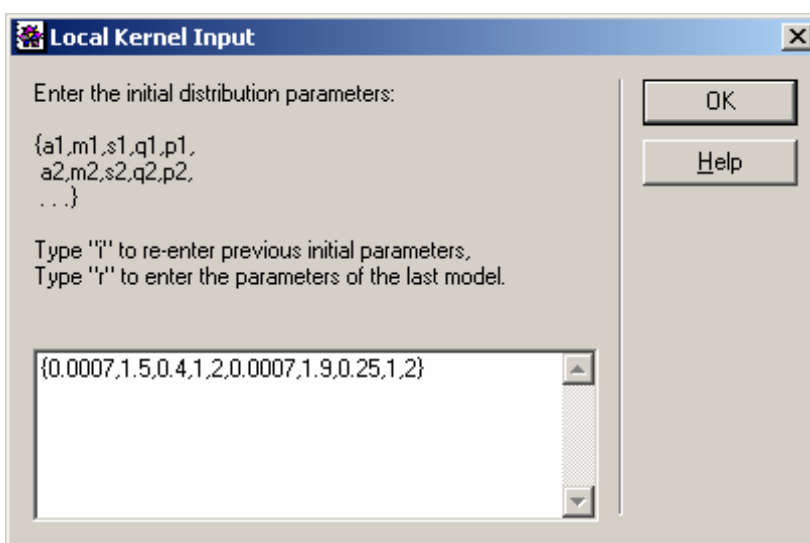
[Back to the program example](#)

Enter initial distribution parameters (2 logarithmic Gaussian functions)

It is possible to run a new component analysis with different initial parameters. After obtaining your first result, the input prompt on the right asks you for one of the following actions: (1) type new initial parameters, (2) re-enter the initial parameters of the previous component analysis, (3) enter the result of previous component analysis as a set of initial parameters.

Since the component analysis with one component was not adequate, a more complex model with two [logarithmic Gaussian functions](#) is used. The initial values for skewness and kurtosis are set to zero by entering $q = 1$ and $p = 2$ for both components. These parameters are kept fixed by entering “{q1,p1,q2,p2}” in the second prompt window. Initial parameters for a , μ and σ are guessed until a relatively good agreement with the measured data is obtained.

[Back to the program example](#)



Perform a component analysis (2 logarithmic Gaussian functions)

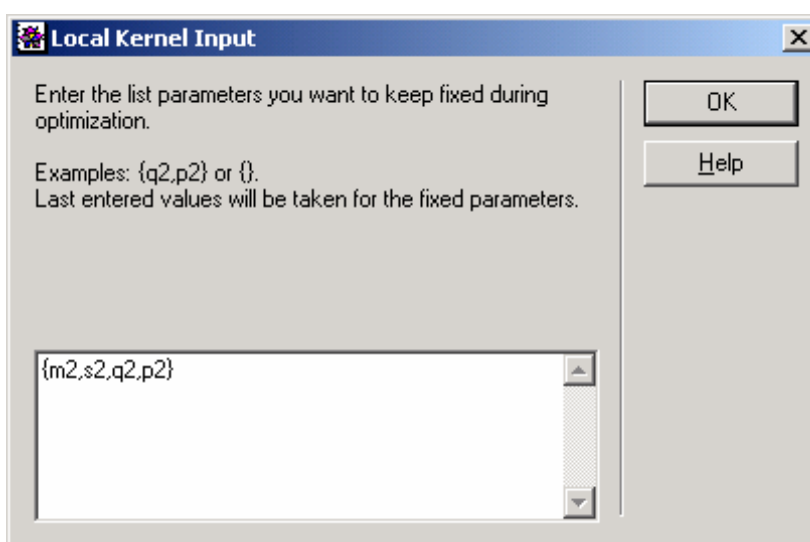
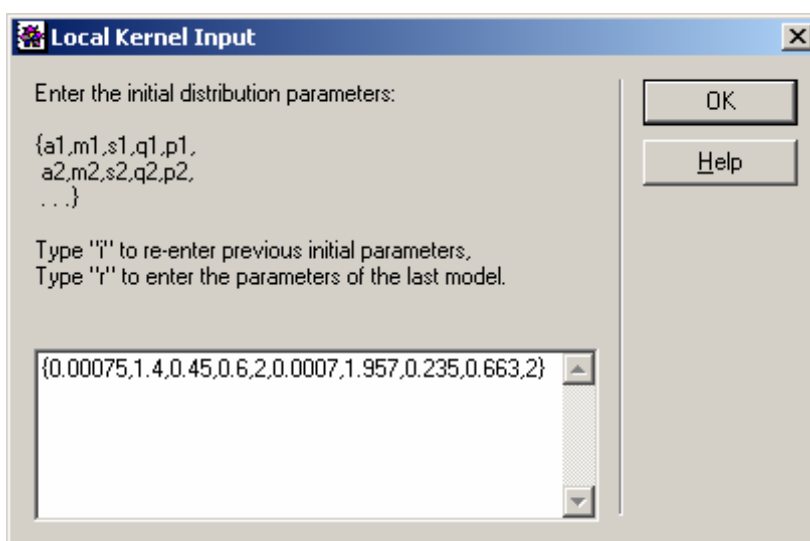
The component analysis with two logarithmic Gaussian functions is inadequate to model the measurements within the given error margins. The misfit between model and measurements is larger than that obtained with one SGG function, even if there is one more parameter to optimize. This example shows that logarithmic Gaussian functions are generally not suitable for modelling the coercivity distribution of single magnetic components. A good agreement between measurements and model is achieved only with 4 logarithmic Gaussian functions ([Figure 3](#)). Unfortunately, these functions cannot be related to the coercivity distributions or real magnetic components.

[Back to the program example](#)

Enter initial distribution parameters (one component is known)

The different sources of magnetic minerals for the sample taken as example are known from independent investigations on urban atmospheric dust samples collected in the same region. The two main sources are given by natural dust and by the products of combustion processes, mainly from motor vehicles and from waste incineration. The coercivity distribution of the combustion products can be modeled by a SGG function with $\mu = 1.96$, $\sigma = 0.235$, $q = 0.66$, $p = 2$. These parameters are kept fixed during the component analysis. Only the magnetization of the combustion products (given by a) is unknown and is optimized, together with the unknown parameters of natural dust.

[Back to the program example](#)



Perform a component analysis (one component is known)

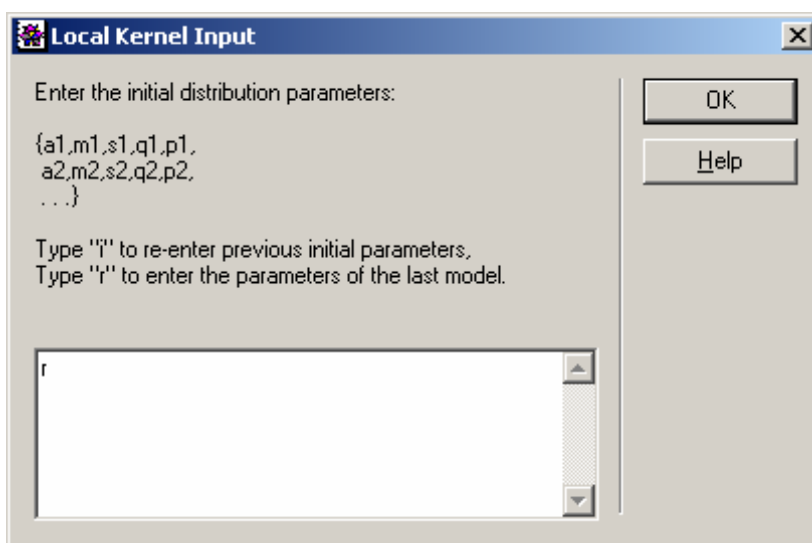
This component analysis is characterized by a much better agreement with the measurements, if compared to the previous results. Six distribution parameters have been optimized. The same number of parameters has been used to perform a component analysis with two logarithmic Gaussian functions: nevertheless, χ^2/l was almost one order of magnitude larger! This model is still significantly different from the measurements. A reason for that could arise from small variations in the properties of the same magnetic component collected from different places.

[Back to the program example](#)

Enter initial distribution parameters (2 SGG functions)

To take into account small variations in the magnetic properties of combustion products, the results of the previous component analysis are taken as initial values for a new component analysis where all 10 distribution parameters are optimized. This is done by typing “r” to recall the result of the last component analysis.

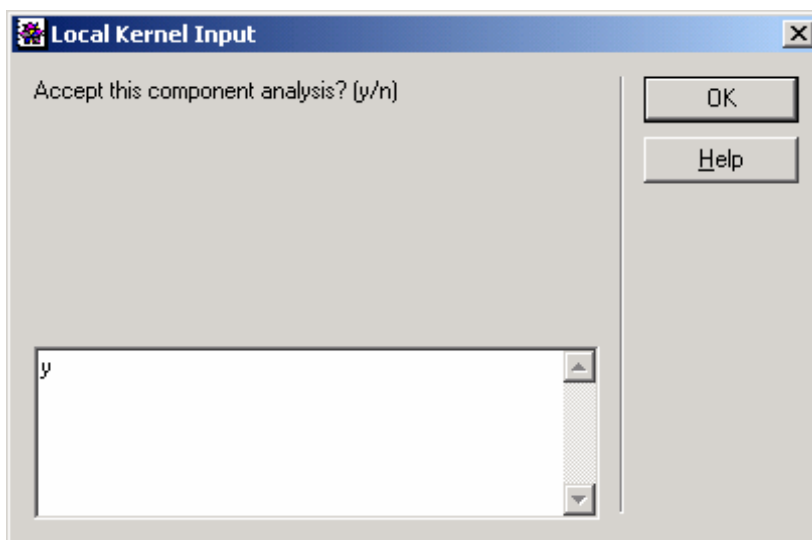
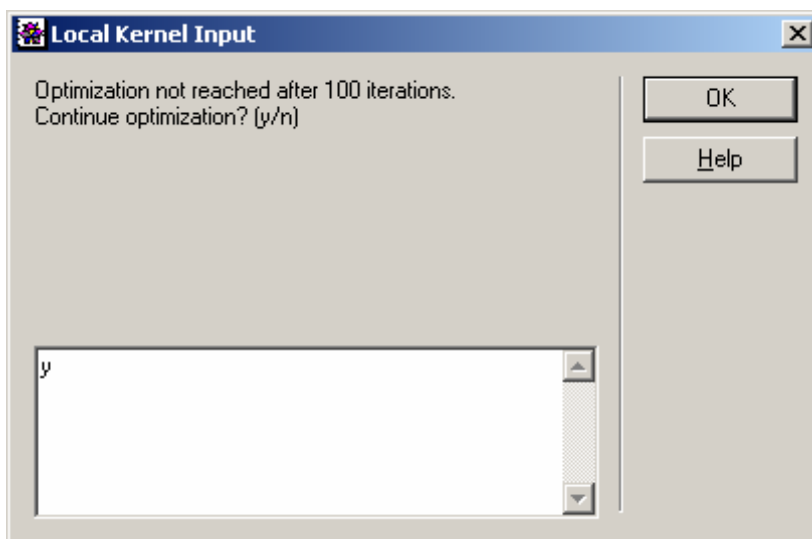
[Back to the program example](#)



Perform a component analysis (2 SGG functions)

The model is now much more complex and the search for a stable solution needs more than 50 iterations. A warning message appears and you are asked to stop or continue for other 50 steps. Finally a stable solution is reached. The distribution parameters of the combustion product did not change more than 10% with respect to the initial values, and the model is now compatible with the measurements within the measurement errors. The solution of this component analysis is accepted, since it is compatible with the magnetic properties assumed initially for the combustion products.

[Back to the program example](#)

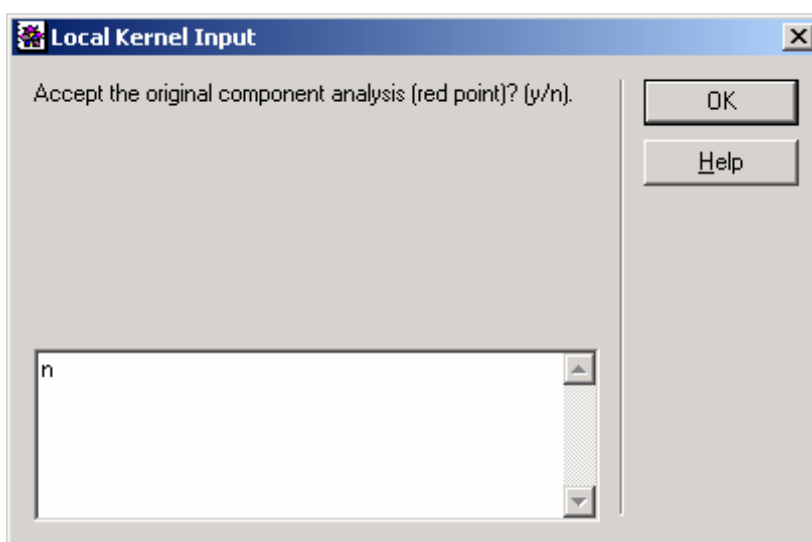
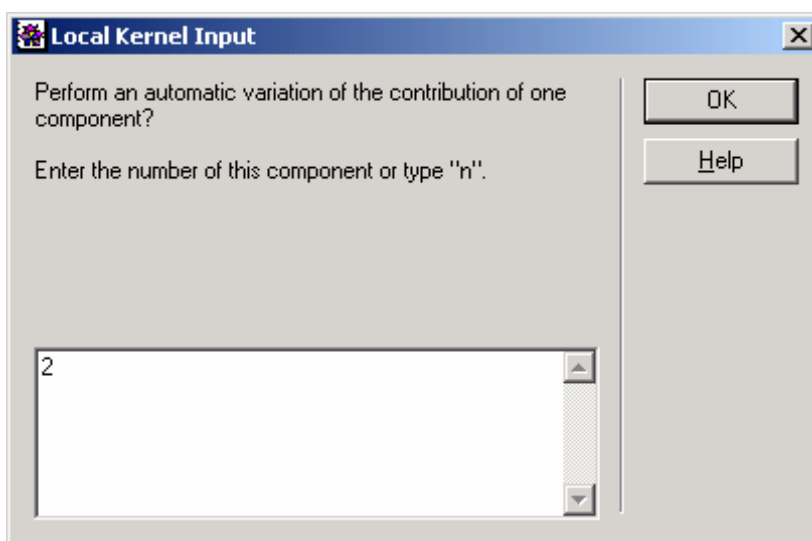


Perform a systematic solution search

Stable solutions of the component analysis correspond to local minima of the [merit function](#). The merit function can have several [local minima](#) for a given finite mixture model. One among them is a [global minimum](#) as well, and is usually considered as the acceptable solution. A solution which corresponds to a global minimum of the merit function is attained if the initial model is chosen to be close enough to the acceptable solution. Since this solution is usually unknown in advance, a sufficient number of initial models has to be tested in order to ensure that at least one will converge to a global minimum of the merit function. If you try 5 initial values for each parameter of a model with two SGG functions, you should perform $5^{10} = 10^7$ optimizations! GECA performs a selected search for a global minimum of the merit function, starting with the result of the last component analysis as initial model. You can select one end-member function, whose amplitude a will be increased and decreased in steps of 1/100 of the

total sample magnetization, starting from the solution of the last component analysis. After each step, the new solution is taken as an initial model for the next component analysis. As a result, the merit function is plotted for all possible amplitudes of the selected SGG function. In the program example, the last solution of the component analysis (red point) is close to the global minimum of the merit function. The sharp steps of the merit function are an effect of the sudden convergence of some distribution parameters to a different local minimum. You are asked to accept the solution of the last component analysis, indicated by a red point, if it corresponds to a global minimum of the merit function.

[Back to the program example](#)

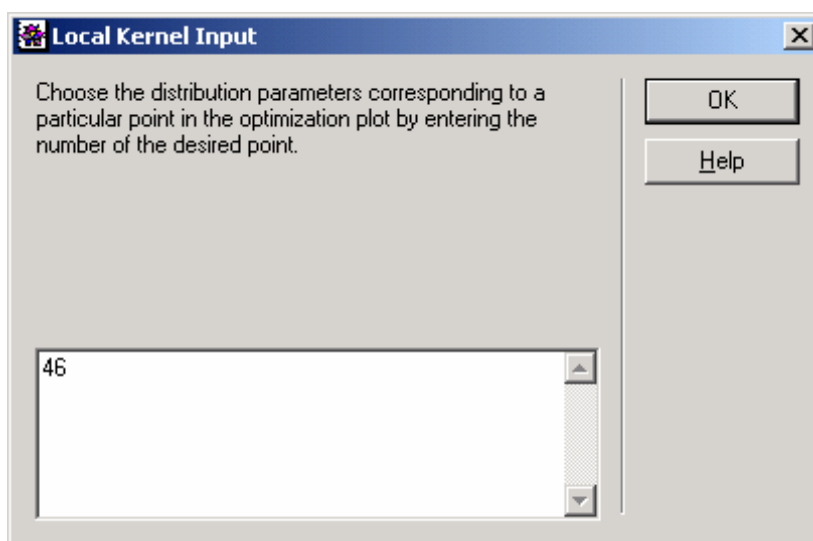


Choose initial distribution parameters from the systematic solution search

You can check the distribution parameters which corresponds to various values of the merit function previously plotted. The merit function was calculated for 100 points (black dots in the last plot, every 10th point is gray). You can enter the number of the point which corresponds to a particular value of the merit function you are interested in. In this way you can explore the solutions which correspond to various local minima and to the global minimum of the merit function. This option is particularly useful in

the case that several local minima exist, which are close to the global minimum. Due to the measurement errors, a meaningful solution could be given by one of these [local minima](#). You may evaluate different solutions with some independent informations about the coercivity distribution expected for the individual magnetic components. In the program example, point number 46 is entered, which corresponds exactly to the global minimum of the merit function.

[Back to the program example](#)

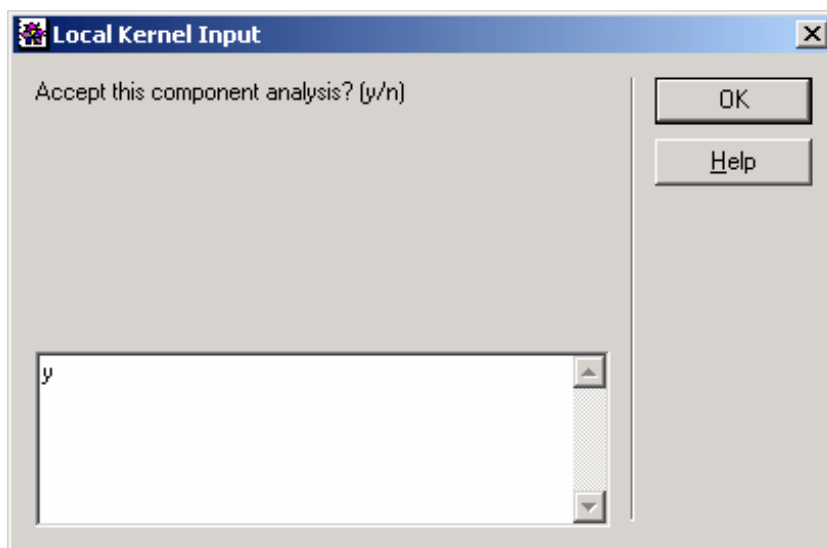


Perform a component analysis (representing a global minimum)

The set of distribution parameters which corresponds to a global minimum of the merit function is taken as initial model for the last component analysis. In the program example, this is the final solution, which represents a finite mixture model which is compatible with the measurements and with independent informations about the properties of the individual magnetic components. In other cases you may not accept this solution and enter a set of initial parameters which

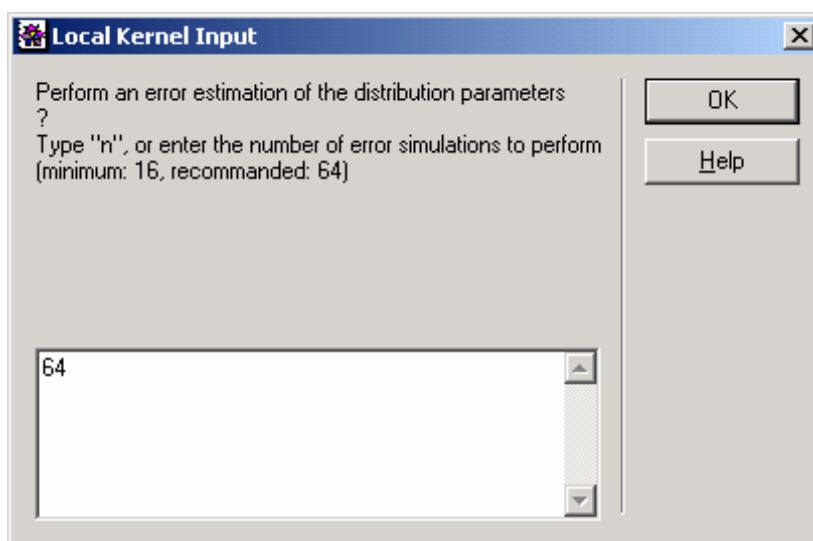
corresponds to other values of the merit function, until a satisfying result is obtained.

[Back to the program example](#)



Perform an error estimation

You can choose to perform an error estimation of the last component analysis. GECA will perform the error estimation by adding a random noise signal to the measured coercivity distribution. The standard deviation of the noise signal is chosen to be identical with the estimated measurement error for each value of the field. The new “noisy” coercivity distribution is fitted with the same set of end-member functions used for the last component analysis, whose solution is taken as the initial model. The result of the

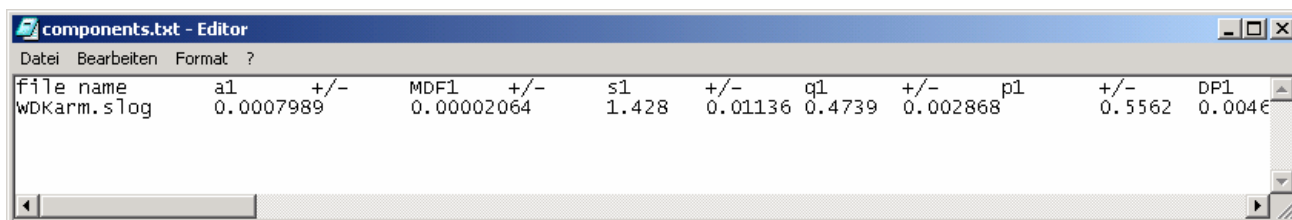
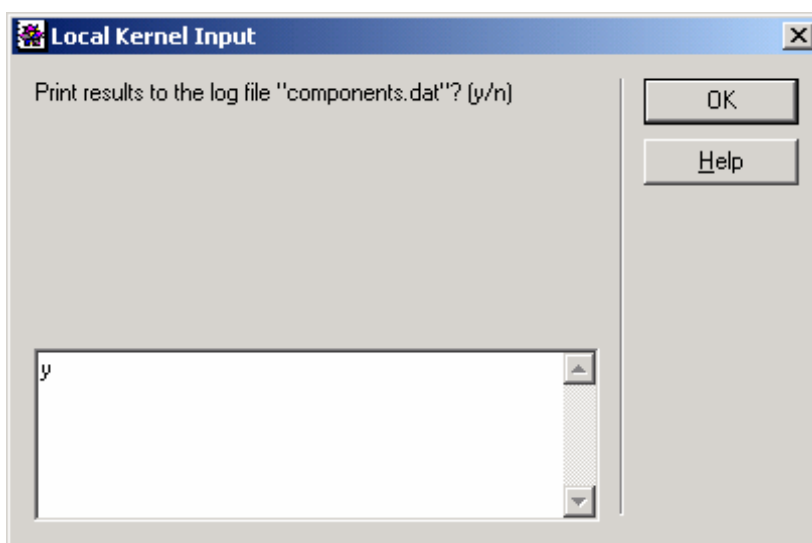


component analysis performed on the “noisy” coercivity distribution differs slightly from the result of the component analysis performed on the original coercivity distribution. The process of adding and adequate noise component to the original data and fitting the resulting coercivity distribution is repeated several times. GECA calculates the standard deviation of the component analysis results for each distribution parameter. These standard deviations are taken as an error estimation. At the same time, some statistical properties of the end-member functions are calculated as well, together with the related errors. The error estimation performed by GECA is quite time consuming, therefore you can choose the number of iterations to perform. With 64 iterations, an accuracy of 12% is expected for the error estimation. The relative accuracy of the error estimation, expressed in %, is given by $100/\sqrt{n}$, where n is the number of iterations used. The error estimation performed by GECA takes into account the effect of the measurement errors on a set of distribution parameters which is related to a particular local minimum of the merit function. The effect of measurement errors on the [convergence of the component analysis](#) to parameters which correspond to other local minima of the merit function is not considered. Therefore, the error estimation performed by GECA has to be considered as a lower limit for the real error of each parameter. Finally, all distribution parameters are displayed together with the estimated errors. Additionally, statistical parameters like the [dispersion parameter DP](#), the [mean](#), the [skewness](#) and the [kurtosis](#) are displayed with the related errors. The result of the component analysis is plotted again, together with the confidence limits of each end-member function. Finally, the normalized end-member distributions are plotted, together with their confidence limits. The area under the curve of each end-member distribution is equal to one in this last plot.

[Back to the program example](#)

Save the component analysis results in a log file

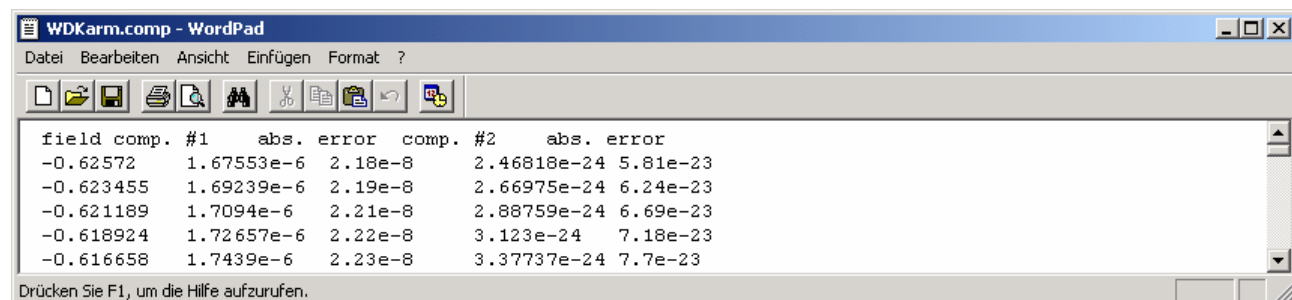
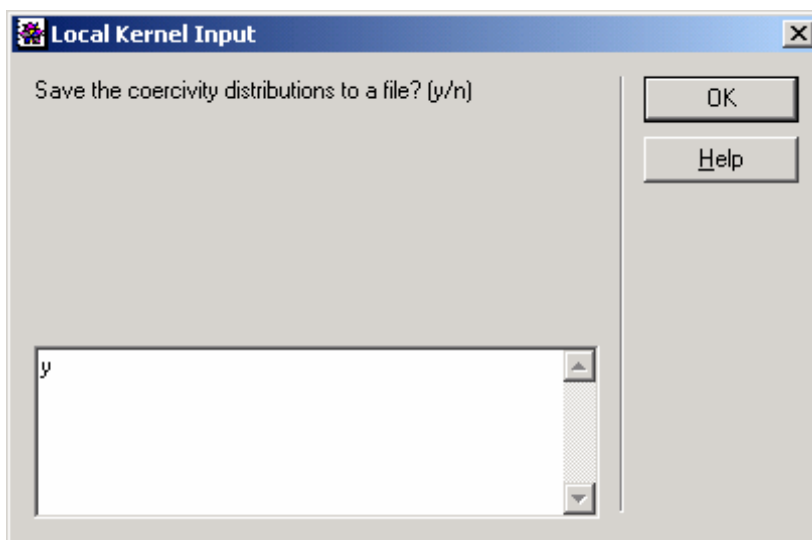
You can save the component analysis results to the log file `components.txt`. You will find this file in the same folder where the program package CODICA is installed. The log file contains the results of all component analysis you decided to save, in form of a list of distribution parameters and statistics for each end-member distribution. The error estimation of each parameter is stored as well, if you decided to run an error estimation with GECA. An example of the content of the log file is displayed below.



[Back to the program example](#)

Save the end-members to a file

The end-member distributions, together with their confidence limits, can be stored in a separated file as a list of columns with the numerical values of each function. The file will have the same name as the file where the original data for the coercivity distribution were stored, with extention `.cum`. This file will be stored in the same directory as the data file. An example is given below.



[Back to the program example](#)

Cautionary note

GECA 2.0 has been tested more than 500 times with coercivity distributions of various artificial and natural samples and the most different combinations of initial parameters. Nevertheless, there is a remote possibility that particular uncommon data or parameter sets will produce evaluation problems. In this case, blue-written warning messages appear on the *Mathematica* front-end. If more than one of these messages is displayed, you may force-quit the Kernel of *Mathematica* as follows: in the top menu bar choose Kernel → Quit Kernel → Local. You can also exit from GECA at any time just by typing “abort” in any input prompt window.

GECA 2.0 does not work on previous versions of *Mathematica 5.0* because of a substantial redefinition of a minimum-search routine embedded in *Mathematica 5.0*. Updates of *Mathematica* will generally not affect the functionality of packages such as GECA, which is expected to run on later versions. Use GECA 1.1 on previous versions of *Mathematica 5.0*.

In case of problems, write to the author (Ramon Egli) at the address given in the `install.txt` file of the installation packet or at the beginning of the source code file `Codica.m`. Please save and send a copy of the *Mathematica* session you were using when the problem arised, together with the data file you analyzed with GECA.

References

Egli, R., Analysis of the field dependence of remanent magnetization curves, *J. Geophys. Res.*, 102, doi 10.1029/2002JB002023, 2003.

Heslop, D., M. J. Dekkers, P. P. Kruiver and I. M. H. Oorschot, Analysis of isothermal remanent magnetization acquisition curves using the expectation-maximization algorithm, *Geophys. J. Int.*, 148, 58-64, 2002.

Kruiver, P. P., M. J. Dekkers, and D. Heslop, Quantification of magnetic coercivity components by the analysis of acquisition curves of isothermal remanent magnetization, *Earth Planet. Sci. Lett.*, 189, 269-276, 2001.