

pyIRM-unmix: An efficient Python tool for IRM unmixing

TingWei Zhang^{1*}, KaiXian Qi¹, Min Zhang², YiFei Hou¹, BiTian Wei¹, ShuHui Cai^{1,3}, and ChengLong Deng^{1,3}

¹State Key Laboratory of Lithospheric and Environmental Coevolution, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China;

²Key Laboratory of Planetary Science and Frontier Technology, Institute of Geology and Geophysics, Chinese Academy of Sciences; Beijing 100029, China;

³College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

Key Points:

- pyIRM-unmix provides unsupervised magnetic coercivity unmixing of isothermal remanent magnetization (IRM) curves with reduced manual tuning.
- A Bayesian modeling framework supports robust parameter initialization and full-parameter fitting.
- The package supports batch processing of IRM acquisition curves, improving efficiency and ensuring consistency.

Citation: Zhang, T. W., Qi, K. X., Zhang, M., Hou, Y. F., Wei, B. T., Cai, S. H., and Deng, C. L. (2026). pyIRM-unmix: An efficient Python tool for IRM unmixing. *Earth Planet. Phys.*, 10(6), 1–11. <http://doi.org/10.26464/epp2026067>

Abstract: Identification and semi-quantification of magnetic mineral assemblages in natural samples are essential for understanding geological processes and paleoenvironmental evolution. Among various rock magnetic approaches, the unmixing of isothermal remanent magnetization (IRM) curves has been widely used to semi-quantify magnetic components and characterize their coercivity distributions. Although various mathematical models and unmixing tools have been developed, existing approaches usually rely on manual parameter setting. This procedure introduces a degree of subjectivity, limits processing efficiency for high-resolution, continuous environmental magnetic datasets, and compromises the consistency of unmixing criteria among samples. To address these limitations, we developed pyIRM-unmix, a Python-based package for IRM unmixing. Compared with conventional workflows, pyIRM-unmix reduces the reliance on manual parameter tuning by enabling robust automatic parameter initialization and full-parameter fitting within a Bayesian modeling framework. Comparative tests demonstrate that the program has high statistical reliability. Furthermore, it substantially reduces processing time while maintaining a fitting accuracy comparable to, or even better than, that of established tools. With its computational efficiency, explicit statistical criteria, and applicability to large datasets, pyIRM-unmix effectively mitigates the subjectivity caused by manual intervention. It provides a practical new tool for high-throughput data analysis in environmental and rock magnetism.

Keywords: magnetic coercivity spectrum; IRM unmixing; Bayesian model; environmental magnetism; rock magnetism

1. Introduction

The magnetic properties of natural materials constitute a fundamental proxy for reconstructing paleoenvironmental records (Evans and Heller, 2003; Liu QS et al., 2012). However, natural samples are typically composed of complex mixtures of magnetic minerals, and their bulk magnetic properties represent the superposition of multiple components. This mixing effect can obscure the primary environmental information carried by individual mineral fractions. For instance, marine and lake sediments commonly contain both detrital and biogenic magnetite (Roberts et al., 2011; Zhang Q et al., 2021; Liu Y et al., 2023; He K et al., 2024; Zhang TW et al., 2025). Although these components share the same mineralogy, they encode fundamentally different environmental signals: detrital magnetite primarily reflects terrigenous

input and regional climatic variability, whereas biogenic magnetite sensitively records bottom-water redox conditions and paleo-productivity. Failure to distinguish between these components may therefore lead to ambiguous or even misleading paleoenvironmental interpretations. Consequently, the accurate identification and separation of magnetic mineral components is a prerequisite for extracting robust environmental information.

Over the past several decades, a variety of magnetic unmixing approaches have been developed and applied to hysteresis loops (Heslop and Roberts, 2012), isothermal remanent magnetization (IRM) curves (Kruiver et al., 2001; Heslop et al., 2002; Egli, 2003; Maxbauer et al., 2016; Zhao XY et al., 2018; Bai F et al., 2022), and first-order reversal curves (FORC) (Lascu et al., 2015; Harrison et al., 2018). Among these approaches, the quantitative decomposition of IRM curves has become widely adopted. This widespread adoption is fundamentally rooted in the robust theoretical premise established by Robertson and France (1994), namely that the bulk IRM of a mixed sample can be regarded as a linear superposition

Correspondence to: T. W. Zhang, zhangtw6@mail.iggcas.ac.cn

Received 01 MAY 2026; Accepted 22 JUL 2026.

First Published online 24 SEP 2026.

©2026 by Earth and Planetary Physics.

of the IRM curves of its independent components when magneto-static interactions are negligible. Furthermore, the acquisition behavior of individual magnetic minerals can be effectively modeled using cumulative log-Gaussian (CLG) functions, which provide a convenient parameterization of their coercivity distributions (Robertson and France, 1994; Eyre, 1996). Building on this linear additivity framework, IRM unmixing methodologies have evolved significantly. Early attempts relied heavily on subjective visual assessments to determine the number and parameters of components. To enhance objectivity, Kruijer et al. (2001) introduced statistical criteria (e.g., *F*-test and *t*-test on the sums of the squared residuals) and developed the IRM-CLG tool in a Microsoft Excel application. However, this approach still required extensive manual parameter tuning and was time-consuming for complex samples. Heslop et al. (2002) implemented an expectation-maximization algorithm that enabled objective and efficient parameter estimation without manual curve matching. Recognizing that natural coercivity distributions are frequently asymmetric and deviate from ideal Gaussian behavior, Egli (2003) further advanced the theoretical framework by introducing the skewed generalized Gaussian (SGG) model. By incorporating additional parameters to describe skewness (*q*) and kurtosis (*p*), the SGG model substantially improved the ability to characterize complex natural coercivity spectra.

Following advances in both theoretical frameworks and computational algorithms, new generation of IRM unmixing tools has been developed to facilitate data analysis. These include MAX UnMix, an R-based platform for implementing the SGG distribution model (Maxbauer et al., 2016); BatchUnMix in the Matlab platform (Bai F et al., 2022); and Python-based packages such as pyIRM (Liu J, 2026). In addition, Zhao XY et al. (2018) proposed a fitting approach based on the Burr Type XII distribution, which provides an alternative framework for decomposing IRM curves. As emphasized by Heslop (2015), the choice of an appropriate unmixing strategy depends strongly on the extent of available *a priori* knowledge about the magnetic mineral assemblage. When sufficient prior information is available, supervised protocols are advantageous because independent constraints can be incorporated into a more meaningful and geologically interpretable unmixing model. Therefore, manual intervention in existing tools should not be regarded simply as a limitation, but rather as a useful way to guide model construction when the magnetic components are reasonably constrained. Nevertheless, practical challenges remain when strong prior constraints are unavailable or large numbers of IRM acquisition curves need to be processed. First, some commonly used tools adopt simplified distribution models or fixed shape parameters (typically skewness of $q = 1$ and kurtosis of $p = 2$), which effectively reduce the SGG distribution to the standard Gaussian case. Although such simplification can improve computational efficiency, it may limit the model's ability to capture complex distribution shapes, particularly those characterized by pronounced kurtosis variations, such as heavy tails or sharp peaks. As noted by Egli (2003), skewness generally exerts a stronger influence than kurtosis in fitting natural samples. MAX UnMix addresses this issue by allowing skewness to be fitted as a free parameter (Maxbauer et al., 2016). However, in practical applications, many users still adopt the default setting (e.g., $q = 1$).

Second, for datasets with weak prior constraints, the manual selection of initial guesses and sample-by-sample parameter adjustment can introduce user-dependent subjectivity and reduce the reproducibility of fitting results (Zhao XY et al., 2018). This issue becomes more pronounced for high-resolution sediment cores or long stratigraphic sequences, where repeated manual tuning is time-consuming and may lead to inconsistencies in component definition across a dataset. Although tools such as BatchUnMix provide batch-processing capabilities, they typically require users to pre-analyze representative samples and define reasonable parameter ranges to guide subsequent fitting (Bai F et al., 2022). Such supervised or semi-supervised workflows are valuable when appropriate prior knowledge exists, but they may be less efficient for exploratory, high-throughput applications. Therefore, an unsupervised and reproducible workflow is needed as a complementary approach, particularly for large IRM datasets or cases where prior constraints are limited.

To address these practical challenges, we propose pyIRM-unmix as an unsupervised tool for IRM unmixing. The tool offers the following advantages. First, it uses a full parameterized SGG model by explicitly incorporating both skewness and kurtosis, thereby allowing a more flexible approximation of complex coercivity distributions. Second, it integrates a Bayesian Gaussian mixture model with a least-squares optimization algorithm to achieve automatic parameter estimation, reducing dependence on manually specified initial parameter settings. Third, it supports both single-sample analysis and batch processing of large datasets, enabling consistent and reproducible unmixing across entire sedimentary sequences.

2. Mathematical Framework

2.1 SGG Model

The SGG model was introduced by Egli (2003) to describe coercivity distributions derived from remanent magnetization acquisition and demagnetization curves. In environmental magnetism studies, this model has particularly been used for analyzing anhysteretic remanent magnetizations (ARM) and IRM acquisition and demagnetization curves, including detailed alternating field demagnetization spectra, to characterize magnetic components in natural samples (Egli, 2003, 2004a; Roberts et al., 2019). Compared with the conventional log-Gaussian model, which assumes a symmetric distribution in log-field space, the SGG model introduces additional flexibility through skewness and kurtosis parameters. The probability density function (PDF) of the SGG model is defined as

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

with $x^* = (x - \mu)/\sigma$, $\sigma > 0$, $0 < |q| \leq 1$, and $p > 0$. Here, x is the logarithmic field of data, whereas μ and σ denote the position and scale parameters of the distribution, corresponding to the mean

coercivity (B_h) and dispersion parameter (DP) of an individual component, respectively. In addition to the scale parameters σ , the SGG distribution includes two additional shape parameters q and p , which control skewness and kurtosis, respectively. The standard Gaussian distribution represents a special case of the SGG distribution where $q = 1$ and $p = 2$ (Figure 1). Specifically, the parameter q controls the asymmetry of the distribution. When $q = 1$, the distribution is symmetric, whereas lower absolute values of q allow the distribution to exhibit pronounced skewness (Egli, 2004b; Figure 1). The parameter p controls the peak and tail behavior of the distribution. For $p < 2$, the distribution becomes more sharply peaked with heavier tails, approaching a Laplace distribution at $p = 1$. Conversely, for $p > 2$, the distribution becomes flatter near the center and has thinner tails (Egli, 2004b; Figure 1). Therefore, the additional parameters q and p make the SGG distribution more flexible than the conventional Gaussian distribution, providing a flexible statistical approximation for asymmetric and non-Gaussian coercivity distributions observed in remanence data from natural samples, rather than being an exact physical representation of individual magnetic mineral components (Egli, 2004b; Zhao XY et al., 2018; Bai F et al., 2022).

Despite this flexibility, the SGG model should be regarded as a statistical approximation of coercivity distributions derived from remanence acquisition or demagnetization curves, rather than being an exact physical model of individual magnetic mineral components (Egli, 2003). Because unmixing results can be nonunique, especially when components overlap or prior constraints are limited, the interpretation of fitted components should be supported by independent rock magnetic, mineralogical, or geological evidence (Heslop, 2015; Roberts et al., 2019).

2.2 Statistical Criteria for Component Selection

To ensure the reliability of the unmixing results, robust parameter initialization is essential. In this study, a Bayesian Gaussian mixture model (BGMM; Blei and Jordan, 2006) is first applied to the IRM gradient curve to generate initial Gaussian components, effectively reducing the need for subjective manual parameter setting. The number of components tested is from 1 to 6, with values greater

than 5 treated as an upper limit because severe component overlap commonly leads to ambiguous physical interpretation. To reduce the risk of convergence to local optima, the BGMM procedure is repeated 10 times by default, yielding 10 candidate parameter sets for each component number.

To identify the optimal parameters and determine the appropriate number of components, we designed a modified Bayesian information criterion (BIC^*) based on the conventional BIC framework (Schwarz, 1978). The BIC^* evaluates the goodness of fit using the residual sum of squares (RSS), while penalizing both model complexity and geologically unreasonable solutions. For a model with j components, BIC^* is defined as

$$BIC^* = n \ln \left(\frac{RSS}{n} \right) + k \ln(n) + \lambda P_\mu + P_\sigma, \quad (2)$$

where n is the number of data points and k is the total number of free parameters. For the standard Gaussian distribution, each component has 3 parameters (μ, σ, w), resulting in $k = 3j$. The term P_μ represents the position penalty, which penalizes solutions in which the mean coercivities of different components are too close to each other. This penalty increases exponentially as the distance between any two component centers (μ_a and μ_b) decreases, and is defined as

$$P_\mu = \sum_{a=1}^{j-1} \sum_{b=a+1}^j p_{\mu_{ab}} \quad (3)$$

$$p_{\mu} = \exp \left(-\frac{(\mu_a - \mu_b)^2}{2c^2} \right), \quad (4)$$

where c is a scaling constant that controls the sensitivity of the overlap penalty. In the present algorithm, c is set to 0.1 as the default value in the log-field, representing an empirical resolution threshold for distinguishing adjacent component centers (Figure 2). The parameter λ in Equation (2) is a weighting factor used to balance the magnitude of the penalty term against the RSS-based likelihood term, and λ is set to $n \ln(10)$ by default. With this setting, a position penalty of P_μ can be offset only when the RSS is reduced by a factor of 10^{-P_μ} . For example, when a BGMM update

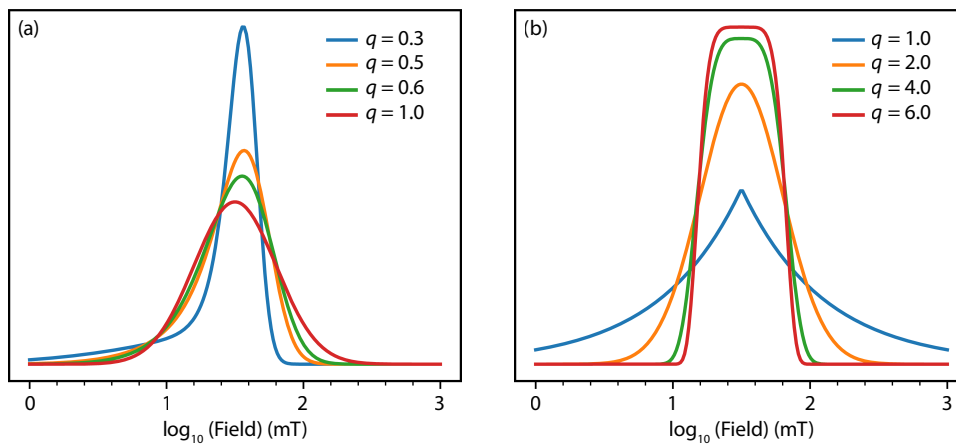


Figure 1. Examples of the SGG model defined by Equation (1). (a) Particular cases of varying skewness ($q = 0.3, 0.5, 0.6$, and 1.0) with kurtosis fixed at $p = 2.0$. (b) Particular cases of varying kurtosis ($p = 1.0, 2.0, 4.0$, and 6.0) with skewness fixed at $q = 1.0$. In all cases, μ and σ are set to 1.35 and 0.25, respectively.

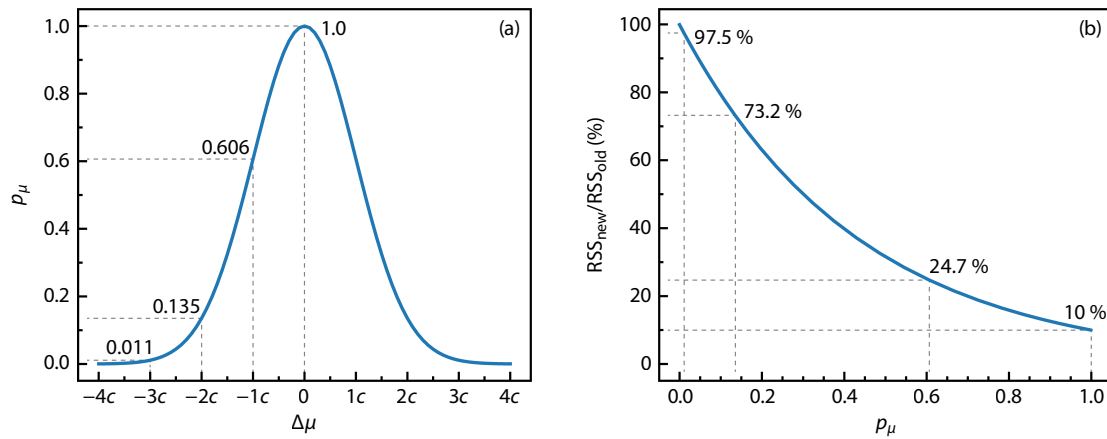


Figure 2. Schematic illustration of the position penalty term and its effect on penalized BIC model selection. (a) Relationship between p_μ and the separation between two component positions $\Delta\mu$ ($\mu_a - \mu_b$). The penalty reaches its maximum value of 1.0 when the two component centers coincide ($\Delta\mu = 0$), and it decreases progressively as their separation increases. (b) Relationship between p_μ and the reduction of RSS required to offset the penalty when $\lambda = n\ln(10)$.

reduces the separation between two component centers to $\Delta\mu = 1c$, with all other component centers remaining well separated, the RSS must decrease to approximately 24.7% of its previous value to offset the position penalty alone (Figure 2). Otherwise, the closely spaced solution will not be favored. In addition, the term P_σ is introduced as a hard dispersion constraint to reject unrealistically narrow or broad components. It is defined as

$$P_\sigma = \begin{cases} 0, & 0.1 < \sigma < 0.6, \\ +\infty, & \sigma \leq 0.1 \text{ or } \sigma \geq 0.6, \end{cases} \quad (5)$$

where $P_\sigma = 0$ if all component dispersion parameters satisfy $0.1 < \sigma < 0.6$, and $P_\sigma = +\infty$ otherwise. In practical implementation, assigning $P_\sigma = +\infty$ means that the corresponding candidate solution is rejected and is not selected as the optimal model. The range of $0.1 < \sigma < 0.6$ follows the commonly used range of dispersion parameters for coercivity distributions in IRM unmixing studies (Egli, 2003, 2004a).

The selection procedure consists of two steps. First, for a fixed number of components, the BIC^* is calculated for all 10 candidate solutions, and the set with the minimum BIC^* is selected as the optimal initialization for that component number. Second, the minimum BIC^* values across models with 1 to 6 components are compared, and the global minimum is taken as the statistically optimal solution. Following Zhao XY et al (2018), the component number suggested by BIC^* should be regarded as an upper limit rather than a definitive estimate of the true number of components. When prior knowledge of magnetic minerals is available, it should be used to further evaluate the selected component number and the interpretation of fitted components. When such constraints are limited or unavailable, the BIC^* -based result can serve as an unsupervised first-order solution, but its geological interpretation should be treated with caution. Researchers may therefore select a component number not exceeding this upper bound for subsequent fitting using the more flexible SGG model.

2.3 Optimization Algorithm

After determining the optimal number of components and their

initial parameters, the final fitting is performed using the full SGG model through nonlinear least-squares (NLS) optimization with the objective of minimizing the RSS between the modeled and observed IRM gradient curves. At the initial stage, the shape parameters of each component are reset to standard Gaussian values ($q = 1$ and $p = 2$). Given that high-dimensional nonlinear optimization is prone to becoming trapped in local minima under multi-parameter coupling conditions, a Gaussian perturbation strategy is introduced to enhance the global search capability (Maxbauer et al., 2016). Specifically, small random perturbations with a standard deviation (std) of 0.02 are applied to the initial parameters of μ, σ, q, p , thereby expanding the search space and improving the likelihood of convergence to the global optimum.

The final optimization is carried out using the Sequential Least Squares Programming (SLSQP) algorithm (Kraft, 1988), which effectively handles bounded and constrained nonlinear problems. To ensure numerical stability and avoid geologically unrealistic solutions, parameter bounds are imposed during the final SGG fitting. These bounds are used as default optimization constraints rather than universal physical limits, and they can be adjusted by users when independent constraints are available. The component position μ is restricted to vary within ± 0.25 log-field units around the initial values estimated by the BGMM, which allows local refinement while preventing large shifts during optimization. The dispersion parameter σ is constrained within the range $0.15 \leq \sigma \leq 0.35$, consistent with commonly observed dispersion ranges of natural samples. The skewness q is constrained to 0.15–0.35 and kurtosis p is constrained to 1.5–4.0, to allow moderate asymmetric and non-Gaussian shapes while avoiding excessively skewed or sharply peaked distributions that are difficult to interpret geologically. These default bounds, together with perturbed initialization, are intended to improve numerical stability and reduce dependence on manual parameter adjustment during SGG fitting and to provide reproducible candidate solutions.

2.4 Error Estimate

A single optimization typically yields only deterministic parameter

estimates. To quantitatively assess the robustness and reliability of the model, a Monte Carlo-based bootstrapping framework is utilized to estimate uncertainties in the component parameters, following a approach similar to Maxbauer et al. (2016). The procedure is driven by a user-defined number of iterations (default is 100). In each iteration, randomness is introduced into both the input data and the initial parameters. Specifically, a random subset of the original IRM data is selected (95% retained by default) to simulate measurement variability, and Gaussian perturbations ($\text{std} = 0.02$) are applied to the initial parameters, serving as new initial seeds for the optimizer.

The SGG model is then re-optimized for each resampled dataset, generating an ensemble of parameter solutions. From this ensemble, the mean and standard deviation of each parameter (w, μ, σ, q, p) are calculated, and the 2.5 and 97.5 percentiles are used to define the 95% confidence intervals. This approach provides a quantitative assessment of uncertainty and significantly enhances the statistical reliability of the interpreted magnetic components.

2.5 pyIRM-unmix Workflow

In this study, we developed pyIRM-unmix, an open-source Python package. The core computational modules are encapsulated within the main script (`pyIRM_unmix_functions.py`). To accommodate different user preferences, the software is provided in both standard Python scripts (.py) and Jupyter Notebook (.ipynb) formats. The workflow is organized into two modes: single-sample analysis and high-throughput batch processing (Figure 3).

The single-sample workflow consists of the following steps. First, the IRM data are imported and the gradient curve is calculated for visualization (Figure 3). Second, preliminary fitting is performed across a range of component numbers, and the optimal number is determined using the modified BIC*. Third, NLS optimization is applied to obtain the final SGG model. Fourth, parameter uncertainties are estimated using an error estimate approach. Finally, the optimized parameters' component contributions are exported as Text files, along with graphical outputs in both PDF and PNG formats. To support large-scale datasets (e.g., long sedimentary profiles), pyIRM-unmix provides an automated batch-processing workflow designed to ensure consistency in component partitioning across samples (Figure 3). First, preliminary fitting is performed for all samples under a predefined number of components (determined from pre-analysis of representative samples) to obtain initial parameters. Second, the median values of these parameters are used as unified initial parameters for all samples, followed by NLS optimization. Finally, the complete set of fitting results is exported for subsequent paleoenvironmental analysis.

3. Synthetic Data Tests

3.1 Synthetic Data Generation

To evaluate the stability of component number selection and parameter recovery in pyIRM-unmix, a simplified test was performed using a four-component synthetic IRM acquisition curve. The predefined parameters for the four components are listed in Table 1. These values were selected to represent a synthetic IRM curve with moderately separated coercivity compo-

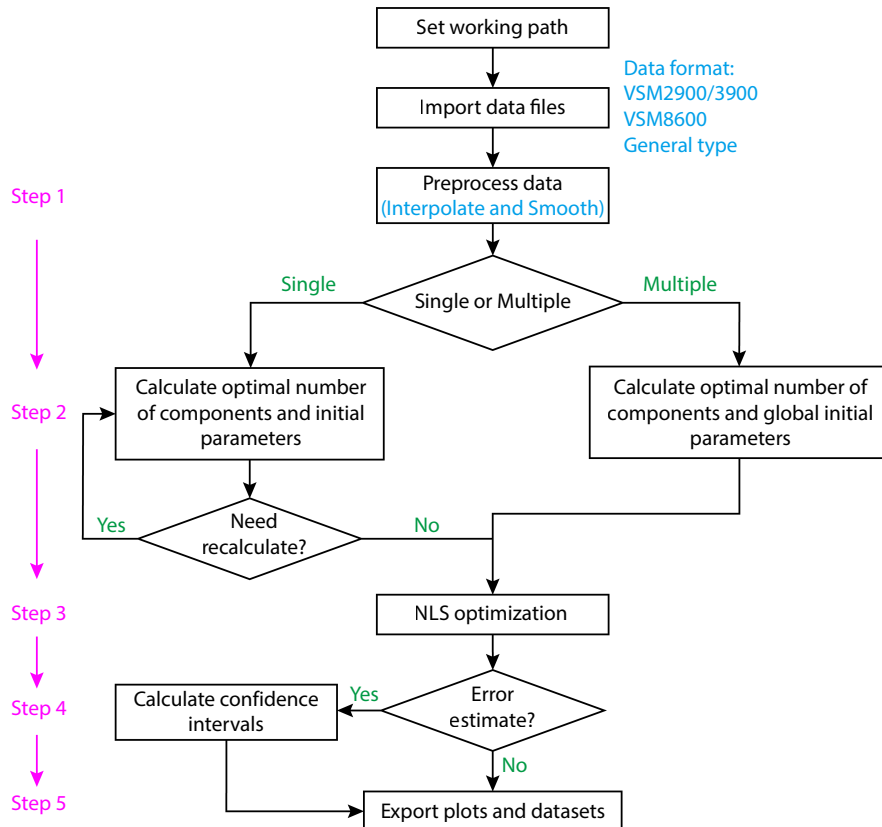


Figure 3. Workflow of pyIRM-unmix, including both single-sample analysis and batch processing procedures.

Table 1. Predefined and recovered parameters for the synthetic IRM datasets.^a

Parameters	Predefined values	Synthetic	Synthetic (+ 0.1% noise)	Synthetic (+ 0.2% noise)	Synthetic (+ 0.5% noise)	Synthetic (+ 1.0% noise)
μ_{Comp1} (mT)	19.953	20.815 $\pm$ 1.380	21.088 $\pm$ 1.475	20.321 $\pm$ 2.081	20.039 $\pm$ 2.394	17.119 $\pm$ 2.475
μ_{Comp2} (mT)	39.811	39.254 $\pm$ 0.250	39.095 $\pm$ 0.535	38.995 $\pm$ 0.395	39.081 $\pm$ 0.899	40.448 $\pm$ 1.315
μ_{Comp3} (mT)	125.893	140.033 $\pm$ 7.178	130.600 $\pm$ 6.772	131.639 $\pm$ 9.050	131.893 $\pm$ 9.396	157.297 $\pm$ 8.720
μ_{Comp4} (mT)	316.228	378.108 $\pm$ 26.579	355.234 $\pm$ 19.641	363.882 $\pm$ 25.956	372.934 $\pm$ 26.086	464.173 $\pm$ 26.094
σ_{Comp1}	0.3	0.307 $\pm$ 0.009	0.313 $\pm$ 0.010	0.313 $\pm$ 0.011	0.314 $\pm$ 0.018	0.257 $\pm$ 0.038
σ_{Comp2}	0.2	0.199 $\pm$ 0.004	0.206 $\pm$ 0.003	0.213 $\pm$ 0.005	0.217 $\pm$ 0.012	0.225 $\pm$ 0.026
σ_{Comp3}	0.3	0.295 $\pm$ 0.016	0.292 $\pm$ 0.013	0.289 $\pm$ 0.011	0.278 $\pm$ 0.018	0.307 $\pm$ 0.029
σ_{Comp4}	0.3	0.261 $\pm$ 0.008	0.266 $\pm$ 0.008	0.253 $\pm$ 0.009	0.224 $\pm$ 0.010	0.159 $\pm$ 0.010
w_{Comp1}	0.2	0.217 $\pm$ 0.022	0.225 $\pm$ 0.024	0.214 $\pm$ 0.035	0.217 $\pm$ 0.044	0.192 $\pm$ 0.050
w_{Comp2}	0.3	0.298 $\pm$ 0.027	0.285 $\pm$ 0.020	0.301 $\pm$ 0.032	0.307 $\pm$ 0.049	0.365 $\pm$ 0.073
w_{Comp3}	0.3	0.348 $\pm$ 0.032	0.326 $\pm$ 0.024	0.329 $\pm$ 0.028	0.329 $\pm$ 0.035	0.363 $\pm$ 0.037
w_{Comp4}	0.2	0.137 $\pm$ 0.030	0.164 $\pm$ 0.023	0.157 $\pm$ 0.030	0.146 $\pm$ 0.026	0.080 $\pm$ 0.008
q_{Comp1}	1	0.993 $\pm$ 0.016	0.974 $\pm$ 0.029	0.999 $\pm$ 0.002	0.982 $\pm$ 0.025	0.943 $\pm$ 0.054
q_{Comp2}	1	0.989 $\pm$ 0.015	0.995 $\pm$ 0.009	0.990 $\pm$ 0.017	0.999 $\pm$ 0.003	0.964 $\pm$ 0.066
q_{Comp3}	1	0.984 $\pm$ 0.021	0.982 $\pm$ 0.025	0.985 $\pm$ 0.013	0.978 $\pm$ 0.022	0.958 $\pm$ 0.067
q_{Comp4}	1	0.984 $\pm$ 0.019	0.987 $\pm$ 0.016	0.997 $\pm$ 0.006	0.970 $\pm$ 0.046	0.959 $\pm$ 0.117
p_{Comp1}	2	2.015 $\pm$ 0.008	2.000 $\pm$ 0.023	2.014 $\pm$ 0.021	1.990 $\pm$ 0.032	1.731 $\pm$ 0.188
p_{Comp2}	2	1.977 $\pm$ 0.011	1.992 $\pm$ 0.010	2.002 $\pm$ 0.016	2.051 $\pm$ 0.076	2.278 $\pm$ 0.190
p_{Comp3}	2	2.031 $\pm$ 0.010	2.034 $\pm$ 0.014	2.039 $\pm$ 0.016	2.087 $\pm$ 0.050	3.081 $\pm$ 0.580
p_{Comp4}	2	2.001 $\pm$ 0.003	2.002 $\pm$ 0.006	2.016 $\pm$ 0.009	2.113 $\pm$ 0.106	2.935 $\pm$ 0.502

^a μ , σ , w , q , and p are the center coercivity (field in mT), dispersion parameter, relative contribution, skewness, and kurtosis, respectively.

nents. The μ values were set to 1.30, 1.60, 2.10, and 2.50, corresponding to mean coercivity (B_h) values of approximately 20, 40, 126, and 316 mT, respectively.

These four components were designed to mimic typical coercivity populations that may occur in natural magnetic assemblages. The Comp1 ($B_h \approx 20$ mT) may represent a mixture of pedogenic, extra-cellular, and detrital magnetite (Egli, 2004a). Comp2 ($B_h \approx 40$ mT) is broadly consistent with a biogenic magnetite (Egli, 2004a; Yamazaki et al., 2019). Comp3 ($B_h \approx 126$ mT) may reflect pedogenic hematite or partially oxidized magnetite (Hu et al., 2013; Zhang Q et al., 2018). The highest coercivity component ($B_h \approx 316$ mT) represents a high-coercivity antiferromagnetic mineral such as hematite or goethite (Heslop et al., 2002; Jiang et al., 2022). These interpretations are used only as mineralogical analogues for the synthetic test, rather than as unique assignments of individual components to specific mineral phases.

3.2 Component Number Selection and Noise Robustness Test

A noise-free cumulative IRM acquisition curve was first generated based on the predefined parameters. To test the effect of noise, four additional synthetic curves were produced by adding Gaussian noise to the cumulative IRM curve, with noise levels of 0.1%, 0.2%, 0.5%, and 1%, respectively. Therefore, five datasets were

used in total: one noise-free synthetic curve and four noisy synthetic curves. The corresponding gradient curves are shown in Figure 4a–4e. The curves with 0.1% and 0.2% noise are close to the typical quality of measured IRM data in daily measurement, whereas the curves with 0.5% and 1% noise show stronger fluctuations and may represent lower signal-to-noise conditions.

For each synthetic curve, the BGMM-based BIC* component selection procedure was repeated 10 times. This design was used to examine whether the automatic initialization and optimization procedure would produce similar results when repeatedly applied to the same dataset. After automatic component selection, a four-component NLS optimization was also performed because the true number of components in the synthetic data was known. Table 1 summarizes the mean parameters recovered from the 10 repeated runs for each synthetic curve.

The BIC* results show that the preferred component number is mostly four for both the noise-free and noisy synthetic curves, although the five-component solution is selected in a few runs (Figure 4f–4j). This result is reasonable because adding an extra component can further reduce the fitting residual, especially when small perturbations are present in the data. It reflects a common trade-off in IRM unmixing that achieves improving model agreement while avoiding overparameterization. There-

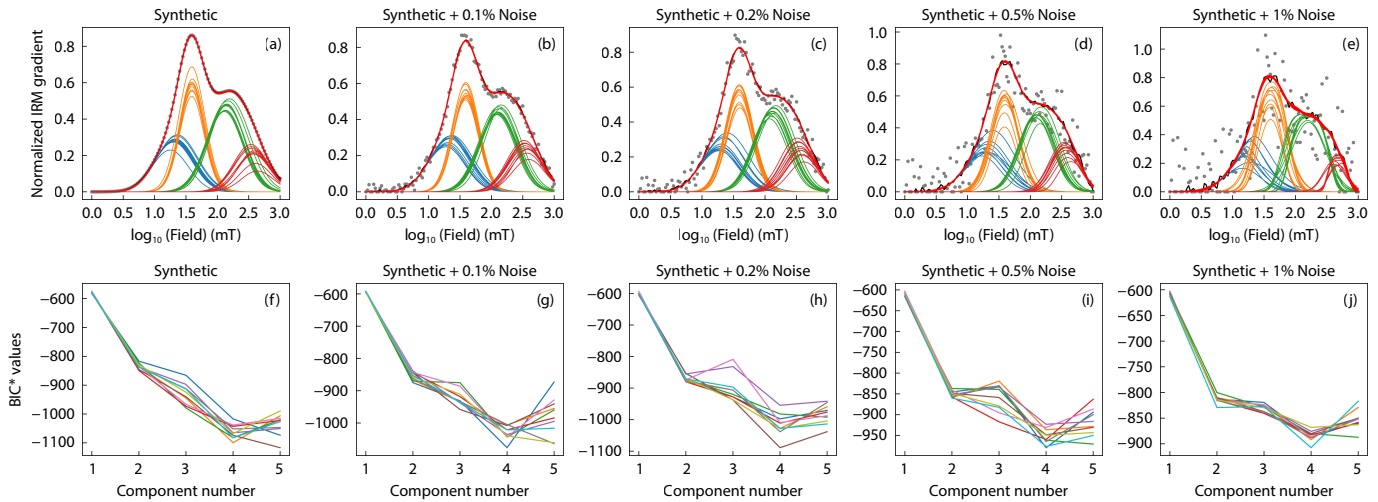


Figure 4. Synthetic curve test results. (a–e) IRM gradient curves generated from the predefined four-component model and noise levels. The four-components NLS optimization results are also shown. (f–j) Component selection results of five datasets.

fore, the statistically preferred component number should be regarded as a guide for model selection rather than as direct proof of the true number of magnetic components.

The NLS results provide a direct test of parameter recovery. For both the noise-free and noisy curves, the four components are generally well separated, and the recovered parameters are close to the predefined values (Figure 4 and Table 1). The Comp1 and Comp2 show particularly stable results. Their centers, dispersion parameters, and relative contributions remain close to the predefined values, even when relatively strong noise is added. In contrast, the highest-coercivity component shows larger deviations, and the difference becomes more pronounced as the noise level increases. This may be related to several factors. First, the high-coercivity component has a broader distribution, which makes its position and contribution less tightly constrained. Second, the high-field part of the gradient curve may remain partly open, so the tail of the high-coercivity component is more difficult to define. These effects are expected in IRM unmixing and indicate that high coercivity components should be interpreted with more caution, especially when the data are noisy or when the applied field range does not fully capture the high-coercivity tail.

Overall, the synthetic test shows that pyIRM-unmix can recover the main components and their relative contributions with good stability when the coercivity components are moderately separated. The repeated fitting results also suggest that the automatic initialization and optimization procedure is not strongly controlled by a single set of starting parameters. However, the occurrence of both four- and five-component preferred solutions confirms that statistical model selection should be used as a guide rather than as a definitive determination of the true number of magnetic components. In practical applications, the statistically preferred solution should therefore be evaluated together with rock magnetic evidence.

4. Application and Discussion

4.1 Single-Sample Analysis

To evaluate the practical applicability of pyIRM-unmix, three

representative data from Figure 2i–2k of Zhang TW et al (2025) were reanalyzed using pyIRM-unmix. These three samples represent different types of magnetic mineral mixtures in the studied lake sediments, as discussed by Yang XQ et al. (2022) and Zhang TW et al. (2025).

The component number was first evaluated using the BIC* values (Figure 5), which indicate that a five-components model provides a reasonable balance between fitting performance and model complexity for the three samples. This selection is further supported by independent mineralogical and rock magnetic evidence. In the studied lake sediments, detrital magnetite and biogenic magnetite were identified by SEM and TEM observations, whereas goethite and hematite were confirmed by diffuse reflectance spectroscopy (DRS). In addition, principal components analysis of FORC (FORC-PCA) results suggested the presence of at least two detrital magnetite populations with different grain-size characteristics, which may correspond to coarse-grained and fine-grained detrital magnetite with different coercivity ranges. Taken together, these observations indicate that five coercivity components are geologically reasonable for these samples.

The five-component solution obtained by pyIRM-unmix shows well-resolved coercivity distributions that are broadly consistent with the mineralogical and rock magnetic evidence (Figure 5). The B_h values of the five components are mainly distributed within 7–14, 22–30, 40–60, 100–200, and 400–550 mT, respectively. These coercivity ranges can be reasonably linked to coarse-grained detrital magnetite, fine-grained detrital magnetite, biogenic magnetite, and high-coercivity antiferromagnetic minerals. The component parameters obtained by pyIRM-unmix are also broadly consistent with those derived from MAX UnMix (Table 2), although some differences occur for individual components mainly in the high-coercivity components. For example, the most notable difference in DP is observed for Comp4, for which MAX UnMix gives a relatively broad distribution, whereas pyIRM-unmix constrains the parameter within the predefined bounds. These differences likely reflect the sensitivity of high-coercivity components to parameter bounds and broad high-field tails,

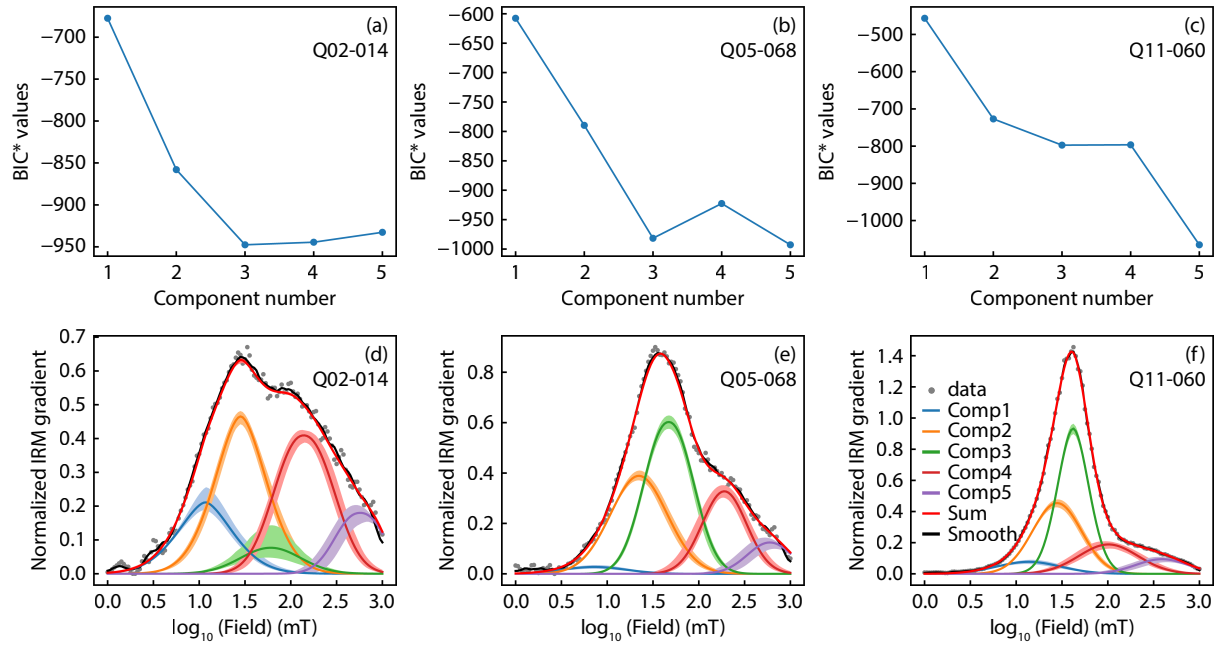


Figure 5. Component number selection and unmixing results. (a–c) Component selection results of three samples. (d–f) Unmixing results obtained using pyIRM-unmix. Data are from Zhang TW et al. (2025). Q02-014, Q05-068, and Q11-060 are sample names.

Table 2. Comparison of fitting parameters for single samples obtained using two unmixing tools.

Parameters	MAX UnMix			pyIRM-unmix		
	Q02-014	Q05-068	Q11-060	Q02-014	Q05-068	Q11-060
B_h _Comp1 (mT)	8.249	8.575	10.601	13.320	7.194	13.699
B_h _Comp2 (mT)	24.274	29.007	25.297	30.149	22.431	27.645
B_h _Comp3 (mT)	38.468	42.687	40.401	58.597	47.727	41.877
B_h _Comp4 (mT)	137.034	185.832	138.006	140.661	194.486	101.056
B_h _Comp5 (mT)	828.673	985.746	752.206	490.269	547.604	387.126
DP_Comp1	0.248	0.230	0.374	0.341	0.350	0.350
DP_Comp2	0.279	0.297	0.287	0.314	0.310	0.266
DP_Comp3	0.236	0.224	0.186	0.350	0.272	0.169
DP_Comp4	0.365	0.382	0.440	0.289	0.232	0.316
DP_Comp5	0.363	0.351	0.432	0.241	0.196	0.281
weight_Comp1 (%)	11.559	6.793	4.402	18.730	2.836	6.767
weight_Comp2 (%)	26.970	25.683	19.620	36.281	30.913	30.332
weight_Comp3 (%)	12.374	33.141	51.817	4.306	40.349	40.854
weight_Comp4 (%)	42.794	31.219	22.709	28.755	19.544	15.209
weight_Comp5 (%)	6.304	3.165	1.454	11.929	6.358	6.839

rather than a fundamental inconsistency between the two methods. In addition, pyIRM-unmix provides a more reproducible workflow by reducing the need for repeated manual parameter adjustment.

4.2 Large-Scale Dataset Batch Processing

The batch-processing workflow of pyIRM-unmix was applied to the complete IRM dataset published by Zhang TW et al. (2025).

Before batch unmixing, all IRM acquisition curves were normalized, smoothed, and converted into gradient curves. The standard deviation of the gradient values at each field step was calculated across all samples, producing a standard deviation curve as a function of the applied field (Figure 6a). This curve is conceptually similar to the standard deviation approach used in grain-size analysis (Boulay et al., 2003), and it highlights the coercivity intervals that show the strongest variation among samples.

The standard deviation curve was further normalized and treated as a representative gradient curve for component-number evaluation. The BIC* values indicate that at least five components are required to describe the main coercivity variability in the dataset (Figure 6b). A five-component NLS optimization was then performed on this normalized standard deviation curve (Figure 6c). The optimized parameters were subsequently used as initial parameters for batch NLS optimization of all individual samples.

Among the five components, Comp3 is interpreted to represent the biogenic magnetite contribution in the lake sediments (Zhang TW et al., 2025). The down-core variation of Comp3 obtained by pyIRM-unmix shows a strong similarity to the biogenic magnetite record obtained from the FORC-PCA method (Figure 6d). It is also broadly consistent with the results obtained using MAX UnMix. All records show increased biogenic magnetite contributions during the M1 and M2 intervals, which correspond to periods of higher effective moisture and lake levels. This pattern suggests that the activity of magnetotactic bacteria was closely linked to lake level

changes and water column conditions in the studied lake system (Zhang TW et al., 2025).

These results demonstrate that pyIRM-unmix can be used for efficient and reproducible batch processing of large IRM datasets. The use of a standard deviation curve provides a practical way to identify the main coercivity intervals responsible for down-core variability and to guide component number selection. However, this preprocessing approach has not yet been fully validated and should be used with caution. It is most appropriate for datasets in which the samples share broadly similar sources and magnetic mineral assemblages. If the sediment source or magnetic mineralogy changes substantially within a sequence, the standard deviation curve may combine signals from different processes and may not provide a reliable basis for defining common components across all samples.

5. Conclusions

pyIRM-unmix provides a practical and reproducible framework for

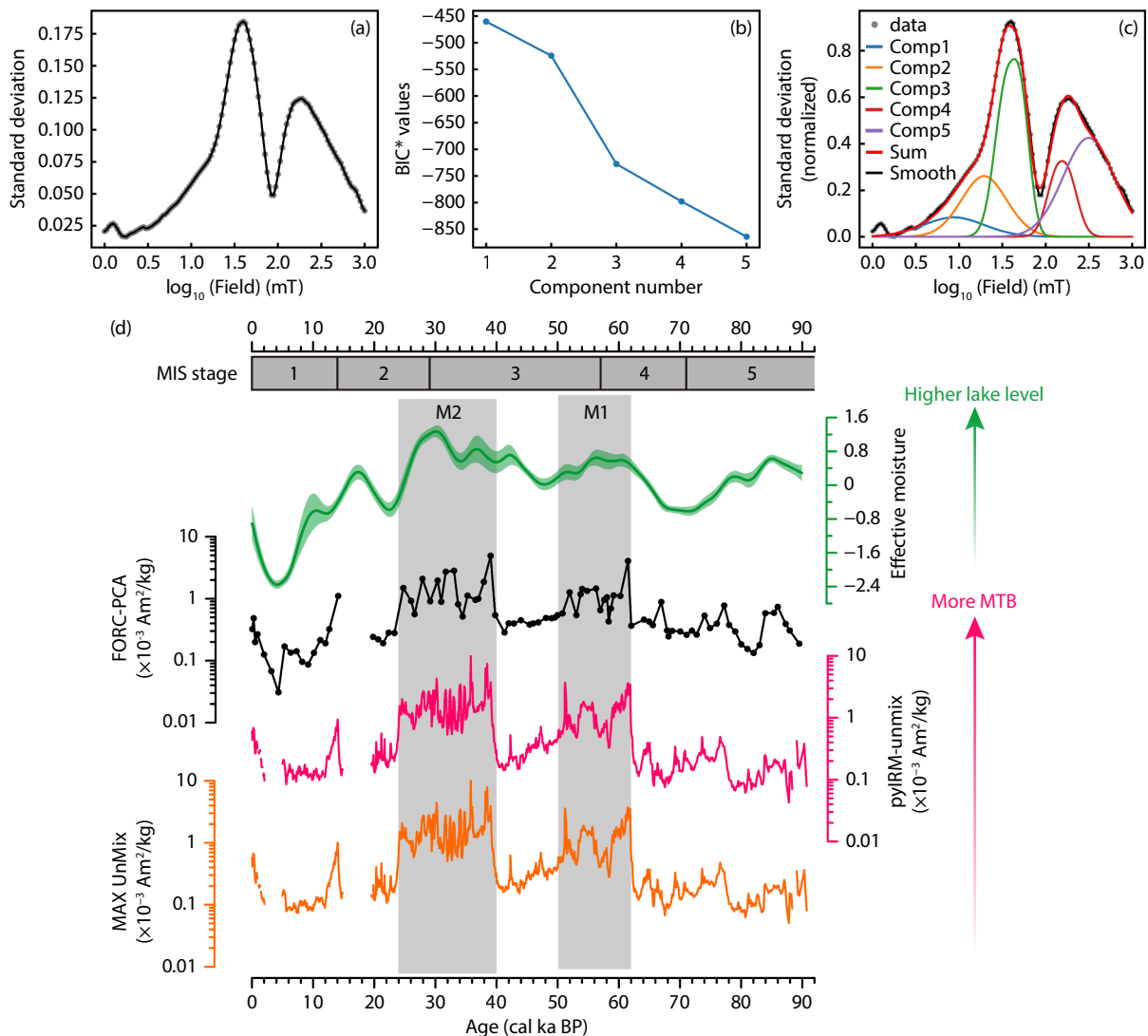


Figure 6. Results of batch processing of multiple samples. (a) Standard deviation values versus a field diagram of all samples. (b) Component selection results of the normalized standard deviation curve in (a). (c) Five-component NLS optimization results of the normalized standard deviation curve in (a). (d) Comparison of the biogenic magnetite component with related records. Data are from Zhang TW et al. (2025).

the statistical unmixing of magnetic coercivity distributions from IRM acquisition curves. Developed within the open-source Python environment, the package offers flexible and cross-platform compatibility through both standard scripts and interactive Jupyter Notebooks. The application results show that pyIRM-unmix can efficiently process both individual samples and large sets of IRM acquisition curves, providing reproducible unmixing outputs for geological interpretation.

In contrast to traditional approaches that rely on extensive manual parameter tuning, pyIRM-unmix enables semi-automated parameter initialization and multi-threaded batch processing, thereby reducing user-dependent subjectivity and improving analytical efficiency. The workflow can be applied in an unsupervised or semi-supervised mode. Because environmental magnetic studies increasingly involve large, high-resolution datasets, such semi-automated and flexible workflows can facilitate consistent analysis across entire sedimentary or stratigraphic sequences. Owing to its computational efficiency, statistical rigor, and scalability, pyIRM-unmix represents a practical and accessible tool for future applications in environmental and rock magnetism.

Future developments could further improve the workflow by incorporating user-defined geological priors, adaptive parameter constraints, and machine-learning algorithms trained on large IRM datasets. Such extensions would allow automated unmixing to better integrate prior knowledge with data-driven optimization, thereby improving the geological interpretability of IRM coercivity spectra.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant Nos. 42525802, 42488201, 42404086, 42241101), the China Postdoctoral Science Foundation (Grant No. 2023M743470), the Bureau of Frontier Sciences and Basic Research, Chinese Academy of Sciences (CAS) (Grant NO. QYJ-2025-01), the Key Research Program of the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS-202401).

Conflicts of Interest

The authors declare no conflicts of interest relevant to this study.

Open Research

The associated Python codes can be found in Zhang TW (2026) or <https://github.com/zhangtw6/pyIRM-unmix>.

References

- Bai, F., Chang, L., Xue, P. F., and Wang, S. S. (2022). Decomposing isothermal remanent curves and development of a batch processing tool BatchUnMix. *Chin. J. Geophys. (in Chinese)*, 65(12), 4789–4801. <https://doi.org/10.6038/cjg2022Q0232>
- Blei, D. M., and Jordan, M. I. (2006). Variational inference for dirichlet process mixtures. *Bayesian Anal.*, 1(1), 121–143. <https://doi.org/10.1214/06-BA104>
- Boulay, S., Colin, C., Trentesaux, A., Pluquet, F., Bertaux, J., Blamart, D., Buehring, C., and Wang, P. (2003). Mineralogy and sedimentology of Pleistocene sediment in the South China Sea (ODP Site 1144). In Prell, W. L., Wang, P., Blum, P., Rea, D. K., and Clemens, S. C. (Eds.), *Proc. ODP, Sci. Results*, 184: College Station, TX (Ocean Drilling Program), 1–21. <http://doi.org/10.2973/odp.proc.sr.184.211.2003>
- Egli, R. (2003). Analysis of the field dependence of remanent magnetization curves. *J. Geophys. Res.: Solid Earth*, 108(B2), 2002JB002023. <https://doi.org/10.1029/2002JB002023>
- Egli, R. (2004a). Characterization of individual rock magnetic components by analysis of remanence curves. 1. Unmixing natural sediments. *Stud. Geophys. Geod.*, 48(2), 391–446. <https://doi.org/10.1023/B:SGEG.0000020839.45304.6d>
- Egli, R. (2004b). Characterization of individual rock magnetic components by analysis of remanence curves. 3. Bacterial magnetite and natural processes in lakes. *Phys. Chem. Earth Parts A/B/C*, 29(13–14), 869–884. <https://doi.org/10.1016/j.pce.2004.03.010>
- Evans, M. E., and Heller, F. (2003). *Environmental Magnetism: Principles and Applications of Enviromagnetics*. Amsterdam: Academic Press.
- Eyre, J. K. (1996). The application of high resolution IRM acquisition to the discrimination of remanence carriers in Chinese loess. *Stud. Geophys. Geod.*, 40(3), 234–242. <https://doi.org/10.1007/BF02300740>
- Harrison, R. J., Muraszko, J., Heslop, D., Lascu, I., Muxworthy, A. R., and Roberts, A. P. (2018). An improved algorithm for unmixing first-order reversal curve diagrams using principal component analysis. *Geochem. Geophys. Geosyst.*, 19(5), 1595–1610. <https://doi.org/10.1029/2018GC007511>
- He, K., Zhao, X. Y., Jiang, Z. X., and Li, S. Z. (2024). Paleoenvironmental controls on the abundances of magnetofossils in the southwestern Iberian margin. *J. Geophys. Res.: Solid Earth*, 129(4), e2023JB027983. <https://doi.org/10.1029/2023JB027983>
- Heslop, D., Dekkers, M. J., Kruiver, P. P., and Van Oorschot, I. H. M. (2002). Analysis of isothermal remanent magnetization acquisition curves using the expectation-maximization algorithm. *Geophys. J. Int.*, 148(1), 58–64. <https://doi.org/10.1046/j.0956-540x.2001.01558.x>
- Heslop, D., and Roberts, A. P. (2012). A method for unmixing magnetic hysteresis loops. *J. Geophys. Res.: Solid Earth*, 117(B3), B03103. <https://doi.org/10.1029/2011JB008859>
- Heslop, D. (2015). Numerical strategies for magnetic mineral unmixing. *Earth Sci. Rev.*, 150, 256–284. <https://doi.org/10.1016/j.earscirev.2015.07.007>
- Hu, P., Liu, Q., Torrent, J., Barrón, V., and Jin, C. (2013). Characterizing and quantifying iron oxides in chinese loess/paleosols: Implications for pedogenesis. *Earth and Planetary Science Letters*, 369–370, 271–283. <https://doi.org/10.1016/j.epsl.2013.03.033>
- Jiang, Z., Liu, Q., Roberts, A. P., Dekkers, M. J., Barrón, V., Torrent, J., and Li, S. (2022). The magnetic and color reflectance properties of hematite: From Earth to Mars. *Reviews of Geophysics*, 60(1), e2020RG000698. <https://doi.org/10.1029/2020RG000698>
- Kraft, D. (1988). A software package for sequential quadratic programming. *Tech. Rep. DFVLR-FB 88-28*, DLR German Aerospace Center-Institute for Flight Mechanics, Koln, Germany.
- Kruiver, P. P., Dekkers, M. J., and Heslop, D. (2001). Quantification of magnetic coercivity components by the analysis of acquisition curves of isothermal remanent magnetisation. *Earth Planet. Sci. Lett.*, 189(3–4), 269–276. [https://doi.org/10.1016/S0012-821X\(01\)00367-3](https://doi.org/10.1016/S0012-821X(01)00367-3)
- Lascu, I., Harrison, R. J., Li, Y. T., Muraszko, J. R., Channell, J. E. T., Piotrowski, A. M., and Hodell, D. A. (2015). Magnetic unmixing of first - order reversal curve diagrams using principal component analysis. *Geochem. Geophys. Geosyst.*, 16(9), 2900–2915. <https://doi.org/10.1002/2015GC005909>
- Liu, J. (2026). pyIRM: A Python package for isothermal remanent magnetization acquisition curve decomposition (Version v1.0.0) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.20265129>
- Liu, Q. S., Roberts, A. P., Larrasoana, J. C., Banerjee, S. K., Guyodo, Y., Tauxe, L., and Oldfield, F. (2012). Environmental magnetism: Principles and applications. *Rev. Geophys.*, 50(4), RG4002. <https://doi.org/10.1029/2012RG000393>
- Liu, Y., Kissel, C., Mazaud, A., Pan, Y. X., and Li, J. H. (2023). Glacial - interglacial circulation and climatic changes in the south Indian Ocean (Kerguelen Plateau region) recorded by detrital and biogenic magnetic minerals. *J. Geophys. Res.: Solid Earth*, 128(11), e2023JB027741. <https://doi.org/10.1029/2023JB027741>
- Maxbauer, D. P., Feinberg, J. M., and Fox, D. L. (2016). MAX UnMix: a web

- application for unmixing magnetic coercivity distributions. *Comput. Geosci.*, 95, 140–145. <https://doi.org/10.1016/j.cageo.2016.07.009>
- Roberts, A. P., Florindo, F., Villa, G., Chang, L., Jovane, L., Bohaty, S. M., Larrasoana, J. C., Heslop, D., and Fitz Gerald, J. D. (2011). Magnetotactic bacterial abundance in pelagic marine environments is limited by organic carbon flux and availability of dissolved iron. *Earth Planet. Sci. Lett.*, 310(3–4), 441–452. <https://doi.org/10.1016/j.epsl.2011.08.011>
- Roberts, A. P., Hu, P. X., Harrison, R. J., Heslop, D., Muxworthy, A. R., Oda, H., Sato, T., Tauxe, L., and Zhao, X. (2019). Domain state diagnosis in rock magnetism: evaluation of potential alternatives to the day diagram. *J. Geophys. Res.: Solid Earth*, 124(6), 5286–5314. <https://doi.org/10.1029/2018JB017049>
- Robertson, D. J., and France, D. E. (1994). Discrimination of remanence-carrying minerals in mixtures, using isothermal remanent magnetisation acquisition curves. *Phys. Earth Planet. Inter.*, 82(3–4), 223–234. [https://doi.org/10.1016/0031-9201\(94\)90074-4](https://doi.org/10.1016/0031-9201(94)90074-4)
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Yamazaki, T., Suzuki, Y., Kouduka, M., and Kawamura, N. (2019). Dependence of bacterial magnetosome morphology on chemical conditions in deep-sea sediments. *Earth Planet. Sci. Lett.*, 513, 135–143. <https://doi.org/10.1016/j.epsl.2019.02.015>
- Yang, X. Q., Zhang, T. W., Zhang, E. L., Toney, J., Zhou, Q. X., and Xie, Y. X. (2022). Paleosecular variations during the last glacial period from Tengchong Qinghai lake, Yunnan province, China. *J. Geophys. Res.: Solid Earth*, 127(3), e2021JB023459. <https://doi.org/10.1029/2021JB023459>
- Zhang, Q., Liu, Q. S., Li, J. H., and Sun, Y. B. (2018). An integrated study of the Eolian dust in pelagic sediments from the North Pacific Ocean based on environmental magnetism, transmission electron microscopy, and diffuse reflectance spectroscopy. *J. Geophys. Res.: Solid Earth*, 123(5), 3358–3376. <https://doi.org/10.1002/2017JB014951>
- Zhang, Q., Liu, Q. S., Roberts, A. P., Yu, J. M., Liu, Y., and Li, J. H. (2021). Magnetotactic bacterial activity in the North Pacific Ocean and its relationship to Asian dust inputs and primary productivity since 8.0 ma. *Geophys. Res. Lett.*, 48(15), e2021GL094687. <https://doi.org/10.1029/2021GL094687>
- Zhang, T. W., Yang, X. Q., and Deng, C. L. (2025). Magnetotactic bacterial activity indicates redox conditions in an upland lake in southwest China since the last glaciation. *Quat. Sci. Rev.*, 356, 109298. <https://doi.org/10.1016/j.quascirev.2025.109298>
- Zhang, T. W. (2026). zhangtw6/pyIRM-unmix: V0.1.1 (Version v0.1.1) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.20749213>
- Zhao, X. Y., Fujii, M., Suganuma, Y., Zhao, X., and Jiang, Z. X. (2018). Applying the burr type XII distribution to decompose remanent magnetization curves. *J. Geophys. Res.: Solid Earth*, 123(10), 8298–8311. <https://doi.org/10.1029/2018JB016082>