Computers & Geosciences 95 (2016) 140–145

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

MAX UnMix: A web application for unmixing magnetic coercivity distributions

Daniel P. Maxbauer [a,b,*], Joshua M. Feinberg [a,b], David L. Fox [b]

[a] Institute for Rock Magnetism, University of Minnesota, Minneapolis, MN, United States
[b] Department of Earth Sciences, University of Minnesota, Minneapolis, MN, United States

ARTICLE INFO

Article history:
Received 24 February 2016
Received in revised form
19 July 2016
Accepted 20 July 2016
Available online 21 July 2016

Keywords:
Unmixing
Rock magnetism
R studio
shiny

ABSTRACT

It is common in the fields of rock and environmental magnetism to unmix magnetic mineral components using statistical methods that decompose various types of magnetization curves (e.g., acquisition, demagnetization, or backfield). A number of programs have been developed over the past decade that are frequently used by the rock magnetic community, however many of these programs are either outdated or have obstacles inhibiting their usability. MAX UnMix is a web application (available online at http://www.irm.umn.edu/maxunmix), built using the shiny package for R studio, that can be used for unmixing coercivity distributions derived from magnetization curves. Here, we describe in detail the statistical model underpinning the MAX UnMix web application and discuss the programs functionality. MAX UnMix is an improvement over previous unmixing programs in that it is designed to be user friendly, runs as an independent website, and is platform independent.

1. Introduction

Magnetic minerals are ubiquitous in a variety of natural systems. Progress in the fields of environmental and rock magnetism has increasingly led to an ability to quantify the abundance, grain size, and chemical composition of various magnetic minerals, which has been critical in enhancing our understanding of an array of natural and anthropogenic processes (see recent reviews by Maher, 2011; Liu et al., 2012; Hatfield, 2014; Maxbauer et al., 2016). In particular, there are a variety of methods available that allow for the statistical unmixing of measured magnetization curves (Robertson and France, 1994; Stockhausen, 1998; Kruiver et al., 2001; Heslop et al., 2002; Egli, 2003; Heslop and Dillon, 2007; Heslop, 2015 provide an excellent review). These methods are widely applied in the literature and have helped to advance our understanding of the processes which govern magnetic mineral formation, transformation, and deposition.

Robertson and France (1994) made the seminal observation that the shape of isothermal remanent magnetization (*IRM*) acquisition curves for an assemblage of grains of a single magnetic mineral could be approximated by a cumulative log-Gaussian function given three parameters: the mean coercivity of an individual grain population (B_h), the component saturation magnetic remanence (M_r), and the dispersion parameter (*DP*; given by one standard deviation in log space). For a given field value of B, the *IRM* of an individual component is given by Robertson and France (1994):

$$IRM(B) = \frac{M_r}{DP(2\pi)^{(1/2)}} \int_{-\infty}^{\infty} \exp\left[\frac{(\log(B) - \log(B_h))^2}{2DP^2}\right] d\log(B) \quad (1)$$

In the case that a specimen is composed of multiple magnetic mineral components, the individual *IRM* acquisition functions (given by Eq. (1)) for each component can be added linearly to approximate the measured data (Robertson and France, 1994; Kruiver et al., 2001). Kruiver et al. (2001) popularized the use of a gradient acquisition plot (GAP) to assist in curve fitting. Subsequent studies refer to the GAP as the coercivity distribution (or spectra; e.g., Heslop et al., 2002, 2004; Egli, 2003), which is the absolute value of the first derivative of the magnetic acquisition dataset (Egli, 2003). Coercivity distributions can be modeled in a similar way to *IRM* acquisition curves by approximation of a probability density function using the same three parameters (Kruiver et al., 2001; Heslop et al., 2002):

$$f(B) = \sum_{i=1}^{n} M_{ri} k(B; B_{hi}; DP_i) \quad (2)$$

where n is the number of magnetic mineral components within a specimen and k corresponds to a log-normal probability density function. From Eq. (2), it is possible to calculate a function that represents the continuous realization of the discrete measured data. Various statistical procedures are used to determine the

* Corresponding author at: 310 Pillsbury Dr. SE, Department of Earth Sciences, University of Minnesota, Minneapolis, MN, 55408, United States.
E-mail address: maxba001@umn.edu (D.P. Maxbauer).

http://dx.doi.org/10.1016/j.cageo.2016.07.009
0098-3004/

goodness of fit for a particular model compared to the measured data using either statistical tests (*F*-test and *t*-test; Kruiver et al., 2001) or automated iterative approaches (Expectation Algorithm; Heslop et al., 2002). These models are accessible for readers to use through downloads of an excel workbook (IRM-CLG; Kruiver et al., 2001) and a Fortran90 executable program (IRM UnMix, available for PCs; Heslop et al., 2002). Fitting is achieved through either manual entry (Kruiver et al., 2001) or through automated optimization (Heslop et al., 2002).

The functions described by Eqs. (1) and (2) operate under the assumption that coercivities of a given magnetic mineral grain population can be closely approximated by a log-normal distribution (Robertson and France, 1994; Kruiver et al., 2001; Heslop et al., 2002; Egli, 2003). However, it is well known that many natural samples contain magnetic mineral components whose coercivities are not log-normal (Egli, 2003, 2004b; Heslop et al., 2004). To account for non-normality, Egli (2003) introduced the skew generalized Gaussian (SGG) function:

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

where x is equivalent to B in Eqs. (1) and (2), μ is the equivalent of B_h, σ is equivalent to DP, q is related to skewness, and p is related to kurtosis (Egli, 2003). The variable x_* arises from a substitution of x with x_*, where $x_* = g(x, q)$ (see Egli, 2003, for details). A Gaussian distribution is equivalent to the SGG when $q=1$ and $p=2$ (decreasing q from 1 to 0 creates left skewed distributions, changing the sign creates right skewed distributions; decreasing p increases peakedness and increasing p enhances squaredness, Egli, 2003). The SGG function has major advantages over simple Gaussian distributions because it can better account for non-normal behavior that is common in natural samples. Deviations from normality can necessitate the need for additional normal or log-normal components within a model to achieve a satisfactory fit, whereas a single skew-component may prove sufficient (see Egli, 2003; Heslop, 2015). The MAG-MIX method of Egli (2003) is available as a set of Mathematica notebooks (CODICA, for deriving coercivity distributions and GECA, for analyzing coercivity distributions) that include graphical user interfaces to assist in data processing. MAG-MIX has been used to analyze the coercivity spectra from a wide range of natural samples and details of those results can be found in Egli (2004a,b,c).

The methods provided by Kruiver et al. (2001), Heslop et al. (2002), and Egli (2003) have proven to be an excellent basis for more detailed interpretation of the magnetic mineralogy of sediments and other geologic samples. However, despite the certain advances presented by Egli (2003), which continues to be utilized by researchers (e.g., Lascu and Plank, 2013; Li et al., 2013; Ludwig et al., 2013; Liu et al., 2014), many studies continue to utilize older methods from Kruiver et al. (2001) (recent examples include Font et al., 2012; Yamazaki and Ikehara, 2012; Ao et al., 2013; Hu et al., 2013; Abrajevitch et al., 2015) and Heslop et al. (2002) (e.g., Roberts et al., 2012; Channell and Hodell, 2013; Weil et al., 2014; Dorfman et al., 2015). This may be in response to difficulties in applying the SGG method, or in response to the software being available only for Mathematica users (which requires expensive licensure). Here, we present a new program, MAX UnMix, that was designed in the statistical computing language `R` (which is open source and available for MAC, PC, and Linux; R-Core-Team, 2015) and built using `shiny` for `R` studio (Chang et al., 2015). The application functions as a web application (available online at http://www.irm.umn.edu/maxunmix) where users interact with the model via a graphical user interface. Supporting information, including instructional videos and a user manual, are available on the MAX UnMix webpage. Below, we describe the statistical model underpinning MAX UnMix and provide a number of examples to highlight aspects of the model's performance.

2. Model description

The observed coercivity distribution, C, of a measured set of magnetization data (M; may be acquisition, demagnetization, or backfield curves) is defined as the absolute value of the first derivative of the raw data:

$$C = \left|\frac{dM}{d\log(B)}\right| \quad (4)$$

where M and B are the respective magnetization and field values for a given dataset. Note we define C in Eq. (4) using the $\log(B)$ scaling, however various field scalings can be used by simple substitution (e.g., Egli, 2003). MAX UnMix utilizes the `predict()` function to calculate C on either a $\log_{10}$ or linear scale, depending on user selection. In line with previous methods, we recommend fitting magnetization curves with a minimum of 25 data points, although generally it is advantageous to have more if possible (Kruiver et al., 2001).

It is often necessary to remove measurement noise within datasets by either application of a spline function (Heslop et al., 2002) or more sophisticated filtering (e.g., the CODICA program described by Egli, 2003). In MAX UnMix, a simple spline function, `smooth.spline()`, allows the user to determine the appropriate level of smoothing. The smoothing factor, *sf*, can be varied between 0 and 1, where $sf=0$ is equivalent to no smoothing and $sf=1$ is the maximum degree of smoothing for a given dataset. Spline fitting prevents large influences of measurement noise, however over smoothing of data can result in spurious features (typically at low and high-fields; see Heslop et al., 2002; Heslop, 2015) and careful observation of this balance should be monitored by users. To avoid complications resulting from smoothing, users have the option to perform smoothing on either raw magnetization data ("Magnetization smoother", C derived from smoothed magnetization data) or raw coercivity data ("Coercivity smoother", C is smoothed directly from raw coercivity data). These choices work variously well at low and high fields and users can determine which method is optimal for a given dataset. As a general rule, the effects of measurement noise are best reduced by maximizing the degree of smoothing imposed on a data set, while taking special care to avoid 'over-smoothing', which can create artifacts.

When a suitable C has been determined from the measured data, the aim is to determine a model function that approximates C for a given set of field values, B. Within the MAX UnMix framework this is achieved using a skew-normal distribution from the `fGarch` package in `R` (Wuertz and Chalabi, 2015). The `dsnorm()` function within the package creates skew-normal probability density functions that we use within our model in the following form:

$$C_m(B) = \sum_{i=1}^{n} p_i w(B;\ B_{hi};\ DP_i;\ S_i) \quad (5)$$

where p_i is a proportion factor that describes the height of the distribution for each component (p_i can range from 0 to 1, normalized such that a value of 1 is equivalent to the maximum of C), $w()$ is the skew-normal probability density function, S_i is a parameter describing skewness (for S_i less than 1 distributions skew left, and vice versa), and C_m represents the modeled approximation of C. In the special case that $S=1$, $w()$ is equivalent to the