RESEARCH ARTICLE

Key Points:
- We develop a probabilistic framework for estimating FORC distributions
- Using Bayesian model selection, we identify underfitting and overfitting of FORC distributions
- FORCsensei software can be used to determine smoothing parameters to estimate a FORC distribution

Correspondence to:
D. Heslop,
david.heslop@anu.edu.au

An Automatic Model Selection-Based Machine Learning Framework to Estimate FORC Distributions

D. Heslop[1,2] [iD], A. P. Roberts[1,2] [iD], H. Oda[2] [iD], X. Zhao[1,2] [iD], R. J. Harrison[3] [iD], A. R. Muxworthy[4], P.-X. Hu[1,2] [iD], and T. Sato[2,5] [iD]

[1]Research School of Earth Sciences, Australian National University, Canberra, Australian Capital Territory, Australia, [2]Research Institute of Geology and Geoinformation, Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan, [3]Department of Earth Sciences, University of Cambridge, Cambridge, UK, [4]Department of Earth Science and Engineering, Imperial College London, South Kensington Campus, London, UK, [5]Earthquake Research Institute (ERI), University of Tokyo, Tokyo, Japan

Abstract First-order reversal curve (FORC) distributions are a powerful diagnostic tool for characterizing and quantifying magnetization processes in fine magnetic particle systems. Estimation of FORC distributions requires the computation of the second-order mixed derivative of noisy magnetic hysteresis data. This operation amplifies measurement noise, and for weakly magnetic systems, it can compromise estimation of a FORC distribution. Previous processing schemes, which are based typically on local polynomial regression, have been developed to smooth FORC data to suppress detrimental noise. Importantly, the smoothed FORC distribution needs to be consistent with the measurement data from which it was estimated. This can be a challenging task even for expert users, who must adjust subjectively parameters that define the form and extent of smoothing until a "satisfactory" FORC distribution is obtained. For nonexpert users, estimation of FORC distributions using inappropriate smoothing parameters can produce distorted results corrupted by processing artifacts, which can lead to spurious inferences concerning the magnetic system under investigation. We have developed a statistical machine learning framework based on a probabilistic model comparison to guide the estimation of FORC distributions. An intuitive approach is presented that reveals regions of a FORC distribution that may have been smoothed inappropriately. An associated metric can also be used to compare data preparation and local regression schemes to assess their suitability for processing a given FORC data set. Ultimately, our approach selects FORC smoothing parameters in a probabilistic fashion, which automates the derivative estimation process regardless of user expertise.

1. Introduction

First-order reversal curves (FORCs) are a form of magnetic hysteresis measurement that provide diagnostic information for characterizing fine magnetic particle systems. FORCs can, for example, reveal the domain state of magnetic particles (Carvallo et al., 2003; Newell, 2005; Pike et al., 2001; Roberts et al., 2000), distinguish between different forms of magnetic anisotropy (Harrison & Lascu, 2014; Harrison et al., 2019; Newell, 2005; Valdez-Grijalva & Muxworthy, 2019), discriminate mineral subpopulations in mixtures (Harrison et al., 2018; Heslop et al., 2014; Lascu et al., 2015; Muxworthy et al., 2005; Roberts et al., 2000, 2014), quantify magnetic interactions (Carvallo et al., 2006; Muxworthy et al., 2004), and reveal thermal relaxation in single-domain (SD) particles (Lanci & Kent, 2018; Pike et al., 2001). Thus, FORCs can play a crucial role in paleomagnetic, rock magnetic, and environmental magnetic studies where natural materials contain a variety of magnetic minerals with distributions of sizes, interparticle spacings, and so forth. Beyond Earth science, FORCs have been used extensively in solid-state physics, materials science, and industry, where a broad range of fine magnetic particle systems, such as magnetic recording media (Miyazawa et al., 2019; Papusoi et al., 2011; Valcu et al., 2011), requires quantitative characterization.

Since their introduction to the geophysics community by Pike et al. (1999) and Roberts et al. (2000), effort has been made to improve experimental measurement protocols, to refine data processing, and to optimize graphical representation of FORC distributions (see Roberts et al., 2014 for a recent review). Parallel to these practical developments, experimental, theoretical, and modeling studies have provided insights into the

expression of different magnetic particle systems in FORCs. Thus, over 20 years of research has provided a framework for the measurement, processing, representation, and quantitative interpretation of FORC data.

FORCs are partial hysteresis curves constructed by taking a sample from a positive saturating field (B_{sat}) to a predefined reversal field, B_r, and returning to B_{sat}. An individual FORC is the magnetization (M) measured as a function of B_r and B as the applied field returns to B_{sat}. FORC diagrams are constructed from a collection of FORCs and are transformed into a distribution, which is defined by the mixed second derivative of the magnetization with respect to B_r and B (Mayergoyz, 1986; Pike et al., 1999):

$$\rho(B, B_r) = -\frac{1}{2}\frac{\partial^2 M}{\partial B_r \partial B}. \tag{1}$$

Traditionally, FORC distributions are displayed in a rotated coordinate system of coercivity, $B_c = (B - B_r)/2$, versus interaction field, $B_u = (B + B_r)/2$ (Pike et al., 1999). While Equation 1 is readily stated, estimation of ρ is not trivial. Standard finite difference approximations to Equation 1 amplify measurement noise and obscure the underlying FORC distribution. The challenge of estimating ρ is the focus of this paper.

Pike et al. (1999) estimated ρ using a local regression framework (Cleveland & Devlin, 1988; Loader, 1999). Specifically, the magnetization around a point of interest is approximated by a second-order polynomial surface fitted in a least-squares sense to the magnetization data deemed "local" to that point. The second-order polynomial surface takes the form

$$M(B, B_r) = a_1 + a_2 B + a_3 B_r + a_4 B^2 + a_5 B_r^2 + a_6 B_r B, \tag{2}$$

which in turn provides an estimate of ρ, denoted as $\hat{\rho}$:

$$\hat{\rho}(B, B_r) = -\frac{a_6}{2}. \tag{3}$$

The choice of a second-order polynomial can be justified from statistical and physical standpoints. First, a second-order surface corresponds to the lowest order polynomial where $\hat{\rho}$ will not be zero everywhere. Therefore, a second-order surface can be viewed as an appropriately parsimonious polynomial for local regression of FORC data because it employs a minimum number of parameters. Second, a more complex polynomial, such as a third-order surface, would have continuous second derivatives, which is incompatible with some magnetic systems (e.g., Stoner & Wohlfarth (1948) particles). Once the estimation process is repeated for each point of interest, the resulting FORC distribution is a combination of all local regression models. Smoothing can be increased by enlarging the local region to include more data points in each regression model. However, selecting the appropriate smoothing level is a subjective decision. If local regions are too small, noise cancelation will be ineffective, and the underlying structure of the FORC distribution may remain hidden. Alternatively, if local regions are too large, the magnetization will be overly smoothed, and the estimated FORC distribution will be distorted (Roberts et al., 2000, 2014). The effects of variable smoothing are illustrated by Roberts et al. (2000, 2014). Oversmoothing will not only provide a poor representation of the true FORC distribution, but it may also obscure the presence of subtle diagnostic features (Harrison & Feinberg, 2008).

Pike et al. (1999) employed local regression with a square array of data centered on each point of interest. The local squares have a constant size across the FORC diagram, which is defined by a smoothing factor (SF), whereby each side of the square has a length of $2SF + 1$ points with respect to the $\{B_r, B\}$ grid of measurements. Heslop and Muxworthy (2005) developed a statistical framework to accompany the Pike et al. (1999) regression scheme, whereby a sample-specific SF is selected based on estimating the maximum level of noise cancelation that can be achieved before the underlying FORC distribution becomes distorted. Heslop and Muxworthy (2005) further showed how the local regression approach of Pike et al. (1999) could be accelerated using two-dimensional convolution (Savitzky & Golay, 1964). While this approach accelerated estimation of a FORC distribution by a factor of ~500, it required the measured magnetizations to be interpolated onto a regular field grid, which may produce correlated errors that unduly affect the least-squares estimate of ρ.

An alternative approach was proposed by Acton et al. (2007), which suppressed measurement noise by first filtering the FORC magnetization data before estimating ρ via a central finite difference scheme. The

initial filtering step involved a Gaussian function, the width of which could be enlarged to increase the level of data smoothing. Harrison and Feinberg (2008) adopted an alternative locally weighted regression (LOESS—locally estimated scatterplot smoothing) scheme, where regression is based on data that lie within a circle around a point of interest. The relative importance of each data point included in a local regression estimate is then represented using a tricube function (Cleveland, 1979), whereby data closer to the point of interest have a greater influence on the estimation of the second-order polynomial and, therefore, $\hat{\rho}$. The level of smoothing is controlled by changing the size of the circular local regression region to include more (greater smoothing) or fewer (less smoothing) points. Through analysis of fitting residuals, Harrison and Feinberg (2008) provided a scheme whereby different smoothing levels could be compared to inform the selection of the size of the circular regression region to be applied across the FORC diagram. Recently, Berndt and Chang (2019) showed that LOESS-based FORC processing could be accelerated using a fast Fourier transform algorithm.

The estimation techniques of Pike et al. (1999) and Harrison and Feinberg (2008) both employ local regression regions of fixed size throughout a given FORC diagram. Experimental and theoretical analysis, however, has demonstrated that characteristic features in FORC distributions can be anisotropic. For example, noninteracting stable SD particles produce a narrow horizontal "central ridge" feature (Egli et al., 2010; Newell, 2005; Pike et al., 1999), while multidomain particles produce vertical contours close to $B_c = 0$ (Pike et al., 2001). To account for these issues, Egli (2013) developed a flexible locally weighted scheme, named VARIFORC, where both the size and shape of rectangular regression regions aligned with the $\{B_c, B_u\}$ coordinate system are adjusted as a function of their location in a FORC diagram. Therefore, smoothing can be extended horizontally or vertically, depending on the features in a given FORC distribution. The size and shape of regression rectangles in the VARIFORC scheme are controlled by user-defined parameters, which are typically adjusted interactively based on the form of the FORC distribution under consideration.

A statistical framework to estimate significance levels and confidence intervals for locations in FORC distributions was developed by Heslop and Roberts (2012). These statistics are calculated readily for weighted local regression problems; however, estimated significance levels and confidence intervals depend on the selected smoothing parameters. As pointed out by Heslop and Roberts (2012), it is feasible to select a smoothing scheme that will ensure that any given location in a FORC distribution is statistically significant. Clearly, such an approach would be inappropriate, and Heslop and Roberts (2012) recommended that smoothing parameters are selected independently to ensure objective estimation of significance levels and confidence intervals.

Recently, Cimpoesu et al. (2019) developed a more advanced statistical framework for comparing estimated FORC distributions via goodness of fit metrics and model comparison techniques. Building on the local regression framework described by Loader (1999), Cimpoesu et al. (2019) considered different methods for estimating ρ and illustrated how information criteria, which measure the level of information loss during regression (Burnham & Anderson, 2002), can be used to compare different smoothing parameters to guide calculation of a FORC distribution. The doFORC software of Cimpoesu et al. (2019) provides a flexible local regression tool where estimated FORC distributions can be compared statistically. The doFORC analysis framework does not, however, accommodate schemes such as VARIFORC that allow variable smoothing as a function of position in a FORC distribution. For natural materials, which often contain a number of magnetic particle subpopulations with drastically different hysteresis properties, such variable smoothing is crucial to estimate different regions of a FORC distribution in a manner appropriate to data in that region (Egli, 2013).

As FORC processing techniques have become more complex, there is an increasing challenge for end users to make decisions concerning the method by which ρ is estimated and the choice of technique-specific settings, such as smoothing parameters. Poorly chosen processing settings may result in spurious inferences being drawn from inappropriately processed FORC distributions. Ultimately, FORC processing should reduce the influence of measurement noise without overly distorting the underlying FORC distribution. To address the problem of estimating FORC distributions that provide a balance between noise reduction and signal distortion, we adopt a Bayesian machine learning framework that considers learning as a probabilistic inference problem. Specifically, the structure and parameters of a local regression scheme to estimate ρ for a given FORC diagram can be learned directly from the data using probabilistic inference. Thus, our aim is to design a framework that guides the selection of VARIFORC smoothing parameters based on ensuring consistency

with the measured magnetization data. Such an objective approach is particularly beneficial for nonexpert users who may not be familiar with FORC processing or who simply want to automate the processing of large numbers of FORC data sets.

2. Probabilistic Comparison of FORC Regression Models

When estimating ρ via local regression, it is important to consider the balance between noise removal and signal distortion (Egli, 2013; Heslop & Muxworthy, 2005; Roberts et al., 2000). In the case of "overfitting", the polynomial regression model applied to a local set of points is too complex (i.e., it contains too many terms) and will fit the data too closely. Thus, the noise will have an inappropriately strong influence on $\hat{\rho}$. Overfitting is the reason why FORC distributions are noisy if they are estimated using SFs that are too low. Alternately, "underfitting" occurs when the local regression model is too simple (i.e., it contains too few terms) to fit the data appropriately. For example, this is the case when employing SFs that are too high and FORC distributions become distorted because the local model cannot fit the data adequately. Thus, when selecting appropriate smoothing parameters for a given FORC data set, it is crucial to develop a strategy that considers potentially detrimental effects of both underfitting and overfitting.

The simplest implementation of VARIFORC is based on five smoothing parameters; $s_{c,0}$ (minimum horizontal smoothing), $s_{u,0}$ (minimum vertical smoothing), $s_{c,1}$ (minimum horizontal smoothing away from $H_c = 0$), $s_{u,1}$ (minimum vertical smoothing away from a central ridge if one is present), and λ (linear increase in smoothing with increasing field). If we consider the estimation of a collection of FORC distributions using an ensemble of candidate VARIFORC smoothing parameters, how can we rank the performance of the estimated distributions in terms of their consistency with the measured data?

2.1. Polynomial Model Selection to Estimate ρ

Consider a single location in a FORC distribution for which we wish to estimate ρ. For a given set of VARIFORC smoothing parameters, a local subset of measurement points is chosen and assigned weights as a function of their location relative to the point of interest (Egli, 2013). A second-order polynomial (Equation 2) is then fitted to the selected points to determine $\hat{\rho}$ (Equation 3) via weighted least-squares regression. Thus, it is assumed implicitly that a second-order polynomial is consistent with the selected data. If, however, the selected VARIFORC smoothing parameters are inappropriate, estimation of $\hat{\rho}$ with a second-order polynomial could result in overfitting or underfitting. Therefore, an ability to detect either overfitting or underfitting in the local regression problem provides a means to assess the suitability of selected VARIFORC smoothing parameters. To identify overfitting or underfitting, we compare regression models of different complexity probabilistically to determine which is most consistent with the data. Specifically, the theorem of Bayes (1763) provides a naturally parsimonious framework for model selection, whereby the performance of alternative models can be compared probabilistically and more complex models are penalized automatically (MacKay, 1992).

We consider three candidate polynomial surfaces to approximate the n magnetizations, M, in a local regression region within the $\{B, B_r\}$ coordinate system. The first surface is a second-order polynomial (H_2^0) that is constrained to have zero mixed second derivative defined by

$$M(B, B_r) = a_1 + a_2 B + a_3 B_r + a_4 B^2 + a_5 B_r^2 \quad \text{and} \tag{4}$$

$$\hat{\rho}(B, B_r) = 0. \tag{5}$$

H_2^0 will always yield $\hat{\rho} = 0$ so its behavior is inconsistent with the main body of a FORC distribution, where $\rho = 0$ is only expected when passing between positive and negative regions (e.g., in SD systems Newell, 2005). Thus, within the main body of a FORC distribution, an H_2^0 polynomial will tend to underfit (i.e., oversmooth). A full second-order polynomial surface (H_2) is defined by

$$M(B, B_r) = a_1 + a_2 B + a_3 B_r + a_4 B^2 + a_5 B_r^2 + a_6 B_r B \quad \text{and} \tag{6}$$

$$\hat{\rho}(B, B_r) = -\frac{a_6}{2}. \tag{7}$$

This polynomial surface is standard in FORC processing (Egli, 2013; Harrison & Feinberg, 2008; Pike et al., 1999), and its use was justified in section 1. Finally, a third-order polynomial surface (H_3) is defined by

$$M(B, B_r) = a_1 + a_2 B + a_3 B_r + a_4 B^2 + a_5 B_r^2 + a_6 B_r B$$
$$+ a_7 B^3 + a_8 B_r^3 + a_9 B_r B^2 + a_{10} B_r^2 B \quad \text{and} \tag{8}$$

$$\hat{\rho}(B, B_r) = -\frac{a_6}{2} - a_9 B - a_{10} B_r. \tag{9}$$

Importantly, the second derivatives of H_3 are continuous, which is inconsistent with the properties of Stoner and Wohlfarth (1948) particles. Therefore, an H_3 polynomial surface may lead to overfitting (i.e., undersmoothing).

When performing local regression at a single location of interest, we compare the candidate polynomial surfaces probabilistically to infer which is most consistent with the data (Appendix A1). If this comparison indicates that an H_2^0 surface is most consistent with the data, then we infer that a FORC estimate with a second-order surface may result in overfitting (i.e., insufficient smoothing because H_2 is more complex than H_2^0) and noise will unduly influence $\hat{\rho}$. Alternatively, if the model comparison reveals that an H_3 surface is most compatible with the data, then the application of a second-order surface will result in underfitting (i.e., excessive smoothing because H_2 is less complex than H_3) that potentially introduces bias into $\hat{\rho}$. Thus, there are two levels of inference necessary for each point of interest in a FORC distribution. The first involves fitting a candidate model to the data, which requires estimation of the coefficients that describe a given polynomial surface. The second involves assigning a probability-based ranking to the alternative models given their compatibility with the data. Model development and comparison are readily undertaken in a Bayesian setting (Bishop, 2006; Gull, 1989; MacKay, 1992), where the model that is most consistent with the data can be selected in a probabilistic fashion through estimation of Bayes factors (BFs) (Hoeting et al., 1999; Sambridge et al., 2006). Therefore, Bayesian model selection (Appendix A1) provides a means to determine whether a local regression-based ρ estimate will be influenced by overfitting or underfitting when $\hat{\rho}$ is evaluated based on a second-order polynomial surface.

2.2. Selection of an Appropriate VARIFORC Scheme

As discussed, model comparison at a single location in a FORC distribution provides a means to assess if a ρ estimate based on a second-order polynomial will be affected by either overfitting (ineffective noise removal) or underfitting (distorting the underlying signal). Thus, the tendency of a given local regression scheme to overfit or underfit the data can be assessed by examining the frequency with which the collection of regression models used to estimate a FORC distribution is most consistent with a given order polynomial. If $\hat{\rho}$ is to be obtained via Equations 2 and 3, then a regression scheme must be selected that maximizes the number of cases most consistent with a second-order polynomial. If there are regions of the main body of the FORC distribution where H_2^0 models are favored, then $\hat{\rho}$ based on H_2 polynomials can be expected to be noisy in these areas (i.e., insufficient smoothing). Alternatively, $\hat{\rho}$ based on H_2 polynomials can be expected to be biased in areas where H_3 polynomials are preferred (i.e., excessive smoothing). For a FORC distribution estimated for a given set of VARIFORC smoothing parameters, we, therefore, determine the proportion of cases in the local regression process where an H_2 model was most consistent with the data. We refer to this proportion as ψ. If a set of VARIFORC smoothing parameters yields local data consistent with second-order polynomials at every position under consideration, then $\psi = 1$.

While ψ could be estimated across the entire measurement space of a FORC diagram, this would be inappropriate because outside the main body of the distribution, ρ is expected to be zero and, therefore, more compatible with H_2^0 than H_2. We restrict evaluation of ψ to the triangular region of a FORC distribution that corresponds to where the major hysteresis loop is open. This involves estimating the upper ($M^+(B)$) and lower ($M^-(B)$) branches of the major hysteresis loop from the FORC data. A field, B_{open}, is then determined, which corresponds to

$$M^+(B_{\text{open}}) - M^-(B_{\text{open}}) = \omega \left[M^+(0) - M^-(0) \right], \tag{10}$$

where ω is set to a small value, such as 0.05. Therefore, B_{open} corresponds to the field at which the difference between the upper and low major hysteresis branches is ω times the separation between the branches at

zero field. The triangular region of a FORC distribution in which the major hysteresis loop is considered to be open is then defined by $\{B_c, B_u\}$ vertices of $\{0, +B_{\text{open}}\}$, $\{0, -B_{\text{open}}\}$, and $\{+B_{\text{open}}, 0\}$.

For a given specimen, it may not be feasible to select VARIFORC smoothing parameters that yield local data consistent with second-order polynomials at every position under consideration in a FORC distribution (i.e., $\psi = 1$). For example, the FORC response due to a noninteracting SD particle assemblage is represented by an infinitely sharp ridge (Egli et al., 2010; Newell, 2005; Pike et al., 1999) that by definition cannot be fitted with a second-order polynomial. Structured deviations from locally second-order data may indicate that the selected regression scheme is not completely appropriate for a given sample, but it is impractical to reject the whole VARIFORC scheme simply because $\psi = 1$ cannot be achieved. We offer a pragmatic solution to this problem by considering a large ensemble (typically thousands) of combinations of VARIFORC smoothing parameters. We then select a suitably smoothed FORC distribution based on the combination of VARIFORC smoothing parameters that yields the highest proportion of H_2 models (i.e., it maximizes ψ).

For a given combination of VARIFORC smoothing parameters, H_2^0, H_2, and H_3 polynomials are estimated for a given point within the triangle of FORC space defined by B_{open}. These polynomial models are then compared probabilistically using BFs. Kass and Raftery (1995) recommended that when comparing two hypotheses, H_A and H_B, if $\text{BF}(H_A, H_B) > 10$, there is "strong" support for H_A, and when $\text{BF}(H_A, H_B) < 1/10$, there is "strong" support for H_B. Using this approach, we estimate ψ for a given combination of VARIFORC parameters by assuming that each local regression problem should be represented by an H_2 model unless there is "strong" support for H_2^0 or H_3. For positions in the FORC distribution where H_2^0 is favored over H_2, $\hat{\rho}$ is set to zero. At all other positions, $\hat{\rho}$ is estimated using H_2. Heslop and Roberts (2012) showed that an ensemble of null hypothesis tests could be used in a local regression setting to define statistical significance throughout a FORC distribution. Regions of a FORC distribution where $\rho \neq 0$ at a given significance level can then be demarcated using contours. Our probabilistic model selection scheme removes the need for the null hypothesis scheme of Heslop and Roberts (2012). Instead, positions in the FORC distribution where H_2^0 is favored are set as $\hat{\rho} = 0$.

We have developed an open-source Python package, FORCsensei (https://forcaist.github.io), which includes an interactive Jupyter Notebook to automate FORC processing and selection of VARIFORC smoothing parameters. The FORCsensei package performs standard FORC preprocessing tasks and then undertakes a grid-search through an ensemble of combinations of the VARIFORC smoothing parameters defined by Egli (2013). This ensemble is constructed based on user-defined limits on $s_{c,0}$, $s_{u,0}$, $s_{c,1}$, $s_{u,1}$, and λ. The smoothing parameters and λ are restricted to lie in the intervals [2, 10] and [0, 0.2], respectively. Smoothing values to be considered are spaced logarithmically between the selected minimum and maximum, rounded to the nearest integer, and duplicate values are removed. λ values are spaced linearly with a separation of 0.04. If the full ranges of the parameter intervals are selected, FORCsensei will compare ψ values of 2,646 FORC distributions generated with different combinations of VARIFORC parameters effectively lying on a five-dimensional grid. Such a large comparison is computationally expensive, so FORCsensei provides functionality to employ downsampling in the estimation of ψ and compare FORC distributions in parallel using Dask (Dask Development Team, 2016), which is an open-source package that can deploy Python code across a multiprocessor system. The number of VARIFORC distributions to be compared can also be reduced by assuming $s_{c,0} = s_{u,0}$, or $s_{c,0} = s_{u,0} = s_{c,1} = s_{u,1}$, and $\lambda = 0$, which corresponds to the approach of Pike et al. (1999).

There are two important caveats to our proposed approach. First, we employ a simple five-parameter VARIFORC scheme; however, Egli (2013) demonstrated the challenges associated with smoothing artifacts in estimated FORC distributions and devised more detailed schemes to remove them. For example, the five-parameter VARIFORC scheme may not be appropriate in regions of FORC space where $B_u \approx -B_c$, when FORC magnetizations are changing rapidly, or when there are large differences between neighboring FORCs. Egli (2013) devised schemes for minimizing the effects of such smoothing artifacts, and users are encouraged to employ sample-specific VARIFORC solutions as necessary. Second, users should place importance on the acquisition of high-quality FORC measurements. No FORC processing algorithm can be expected to produce high-quality FORC distributions from poor-quality measurements. For example, it is crucial that a measurement field step is chosen that is appropriate to resolve features of interest (e.g., see Egli et al., 2010 for a discussion of appropriate FORC parameters for measurement of central ridge features).

Where possible, instrument drift should be minimized, and samples may need to be remeasured if sudden impulse drift events occur (Roberts et al., 2014).

3. Case Studies

We present two case studies to demonstrate our proposed ψ metric and its use in selecting VARIFORC smoothing parameters. These examples were chosen because they provide an intuitive demonstration of our approach. We do not question the analyses or conclusions of the original studies in which the following examples were published.

3.1. Biogenic Magnetite

We present results for a case study from marine sediment core MD00-2361, which was recovered from a water depth of 1,805 m off the coast of northwestern Western Australia (22°04.92′S, 113°28.63′E). Magnetostratigraphic analysis of core MD00-2361 located the Matuyama-Brunhes paleomagnetic reversal at a depth of ~16 m (Heslop et al., 2013). Environmental magnetic analysis and FORC measurements (Heslop et al., 2013) reveal that the MD00-2361 sediments contain both magnetofossils and coarse-grained detrital magnetite. The MD00-2361 FORC distributions have well-defined central ridges (Heslop et al., 2013, 2014), which are indicative of noninteracting stable SD particles (Egli et al., 2010; Newell, 2005). Transmission electron microscope observations revealed abundant magnetofossils (Heslop et al., 2013), and analysis of FORC central ridge signals indicated that the morphology of the magnetofossils changed with the environmental alternations between glacial and interglacial periods (Heslop et al., 2014). Here we reanalyze high-resolution FORC measurements (field spacing of 0.5 mT) (Figure 1a) performed on a sediment sample from a depth of 1.25 m below the seafloor. For an ensemble of VARIFORC models, based on 2,646 smoothing parameter configurations, a maximum of $\psi = 0.38$ (Figure 1b) is obtained when $s_{c,0} = 4$, $s_{u,0} = 2$, $s_{c,1} = 4$, $s_{u,1} = 4$, and $\lambda = 0.04$. Inspection of the highest-ranked VARIFORC model through the space of the FORC distribution (Figure 1c) reveals areas where the data are most consistent with H_2; however, there is a region in the lower half-plane in which H_3 is preferred, where underfitting may occur. Furthermore, there are regions along the lower coercivity $B_u = 0$ region where underfitting occurs because even with high-resolution FORC measurements, a second-order polynomial cannot resolve fully the infinitely sharp central ridge (Egli et al., 2010; Newell, 2005). The resulting FORC distribution (Figure 1d) contains a main central ridge and vertical spreading at lower coercivities produced by the coarser, potentially vortex state (Lascu et al., 2018; Roberts et al., 2017), detrital particles. The FORC distribution also contains anomalous streaking along the so-called remanence diagonal ($B_u = -B_c$).

In a detailed analysis of FORC estimation procedures, Egli (2013) identified anomalous remanence diagonal behavior. Artifacts appear in this region of a FORC distribution when reversible magnetization processes, for example, due to superparamagnetic particles, produce a sigmoidal response as the applied field crosses zero, which cannot be well approximated by a second-order polynomial surface. Egli (2013) recommended an elegant solution to this problem, where the subtraction of the last FORC from all preceding FORCs removes the sigmoidal signal from the magnetization data. The derivative of the last FORC with respect to B_r is zero, so this subtraction does not change the underlying ρ but transforms the magnetization data into a form that is expected to be more consistent with a series of local polynomial surfaces. In practice, this preprocessing step, which has become known as "lower branch subtraction", increases noise in the data, but the ρ estimate may be improved because the form of the model on which $\hat{\rho}$ is based is more consistent with the data. In many cases, the effect of increased noise is negligible, and the Egli (2013) lower branch subtraction technique removes artifacts produced by reversible magnetizations.

When using lower branch subtracted magnetizations for the MD00-2361 sample (Figure 2a), ψ is maximized at 0.70 when $s_{c,0} = 3$, $s_{u,0} = 2$, $s_{c,1} = 3$, $s_{u,1} = 5$, and $\lambda = 0.08$ (Figures 2b and 2c). Much of the tendency to underfit is removed by lower branch subtraction; however, the overfitting region around the central ridge remains because it is related to measurement resolution. Diagonal overfitting features appear in the lower branch subtracted model, which are probably due to uncompensated instrument drift (Figure 2d). The full FORC diagram and lower branch subtracted FORC diagram have similar distributions (Figures 1d and 2d, respectively); however, lower branch subtraction has successfully removed the fitting artifact along the remanence diagonal.

To demonstrate the effects of overfitting and underfitting, we also provide examples where extreme VARIFORC parameters were selected for the lower branch subtracted FORCs. A case of overfitting
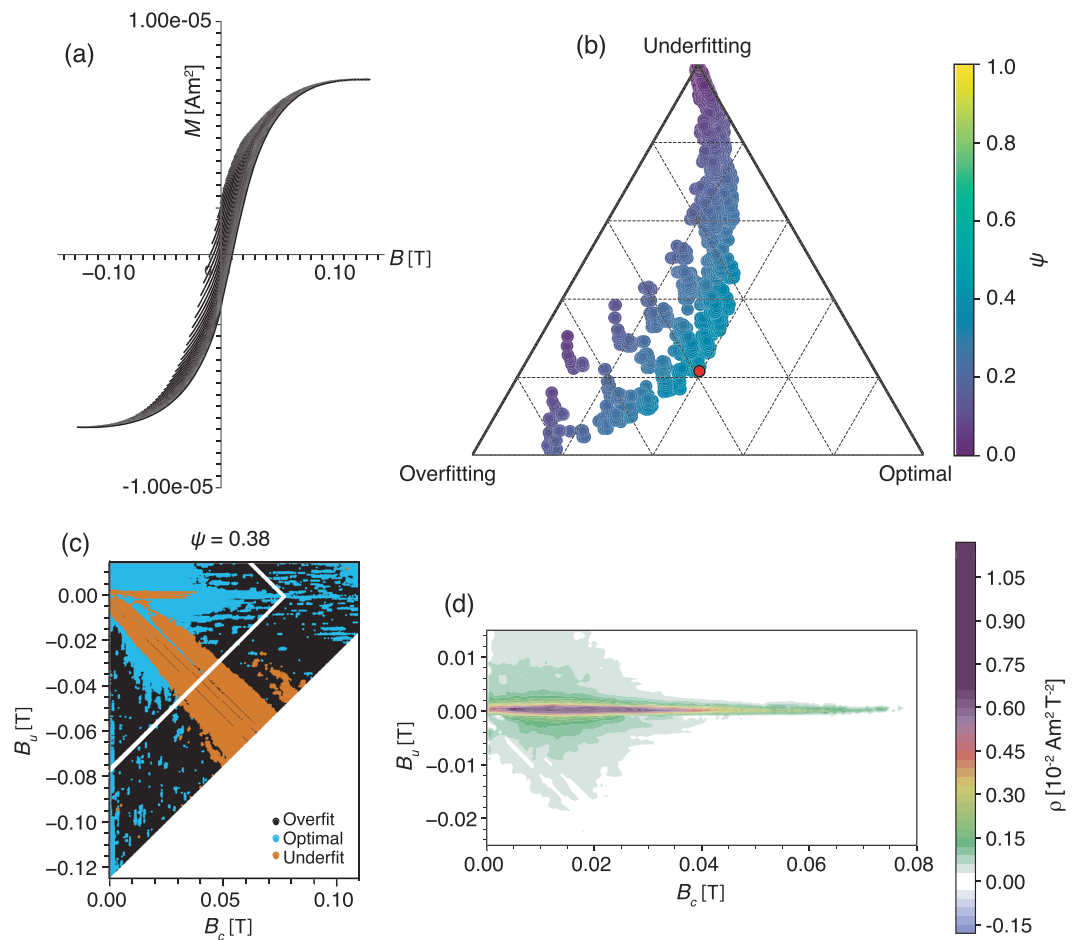
Figure 1. (a) High-field slope-corrected FORCs for a marine sediment rich in biogenic magnetite from core MD00-2361 (sample depth of 1.25 m below seafloor). Every fifth FORC is plotted for clarity. (b) Distribution of model-selection statistics for 2,646 combinations of VARIFORC smoothing parameters. Each point represents a FORC distribution for a given combination of VARIFORC smoothing parameters, where the "overfitting", "optimal", and "underfitting" proportions represent the estimated relative number of H_2^0, H_2, and H_3 cases in the local regression scheme, respectively. The selected VARIFORC configuration, in which ψ is maximized, is shown by a red point; the remaining points are color-coded according to their ψ values. Model comparison was based on 2,000 random locations selected within the triangle of the FORC space defined by B_{open}. (c) Color-coded distribution of selected model order across the space of the selected MD00-2361 FORC distribution. The local regression scheme contains cases of overfitting (black points), where the data are most consistent with H_2^0, optimal cases (blue points), where the data are consistent with H_2, and underfitting (orange points), where the data are most consistent with H_3. The white line corresponds to the B_{open} triangle. (d) Final FORC distribution estimated with selected VARIFORC parameters: $s_{c,0} = 4$, $s_{u,0} = 2$, $s_{c,1} = 4$, $s_{u,1} = 4$, and $\lambda = 0.04$.

(i.e., insufficient smoothing) is shown in Figure 3, where $s_{c,0} = 2$, $s_{u,0} = 2$, $s_{c,1} = 2$, $s_{u,1} = 2$, and $\lambda = 0.0$ yield $\psi = 0.49$. The VARIFORC solution contains a high proportion of local regression models where the data are consistent with H_2^0 rather than the desired H_2 (Figure 3a). In these areas, the data support a hypothesis of $\rho = 0$, and the regions of the FORC distribution that flank the central ridge are lost (Figure 3b). This demonstrates that when considering the performance of given smoothing parameters, it is important to compare the estimated FORC distribution with the distribution of the selected model order. VARIFORC parameters $s_{c,0} = 8$, $s_{u,0} = 8$, $s_{c,1} = 8$, $s_{u,1} = 8$, and $\lambda = 0.12$ are selected to demonstrate underfitting ($\psi = 0.10$), where most of the local regression models involve data more consistent with H_3 than H_2 (Figure 3c). This results in oversmoothing of the underlying signal where the central ridge becomes too broad and the low coercivity flanks widen (compare Figure 3d to Figure 2d).
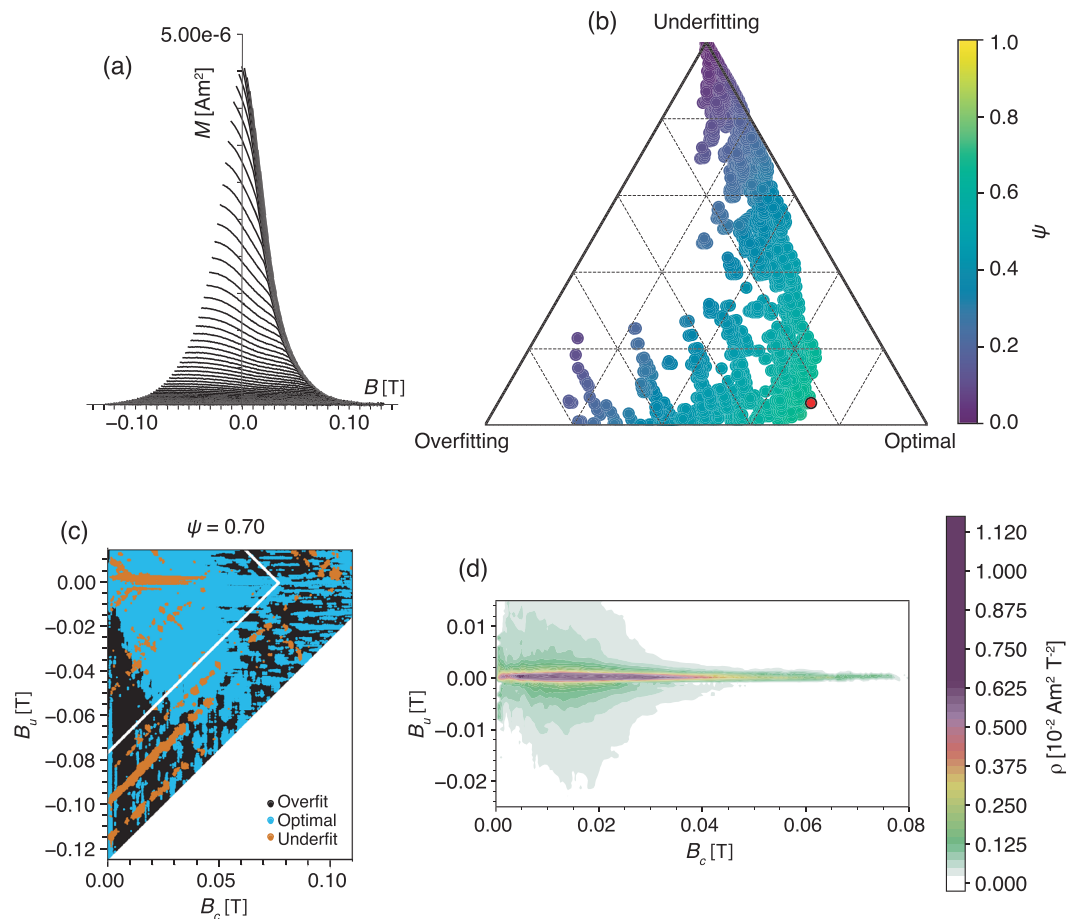
Figure 2. (a) Lower branch subtracted FORCs for the MD00-2361 sample illustrated in Figure 1a. (b) As in Figure 1b for the lower branch subtracted FORCs. (c) As in Figure 1c for the lower branch subtracted FORCs. (d) As in Figure 1d for the lower branch subtracted FORCs with VARIFORC parameters: $s_{c,0} = 3$, $s_{u,0} = 2$, $s_{c,1} = 3$, $s_{u,1} = 5$, and $\lambda = 0.08$.

3.2. Sedimentary Greigite

In their discussion of noise suppression, Roberts et al. (2014) demonstrated changes in the FORC distribution of a sedimentary greigite sample from South Island, New Zealand, as a function of SF (sample UB199B from Roberts & Turner, 1993). The sediment contains strongly interacting stable SD greigite particles, which produce a FORC distribution that is characterized by broad coercivity and interaction field distributions. With increasing SF, the main body of the FORC distribution becomes smoother and broader as the size of the local regression regions grows. To develop rigorous interpretations, appropriate smoothing is required to ensure that the FORC distribution is not overfitted or underfitted. The UB199B magnetization data are smooth (Figure 4a), and the lower branch subtracted curves contain little visible noise (Figure 4b). Comparison of an ensemble of 2,646 candidate VARIFORC models leads to selection of a model that corresponds to the lowest smoothing level ($s_{c,0} = 2$, $s_{u,0} = 2$, $s_{c,1} = 2$, $s_{u,1} = 2$, and $\lambda = 0.0$), yielding $\psi = 0.70$ (Figures 4c and 4d). The resulting FORC distribution (Figure 4e) has a somewhat noisy appearance, which reflects a balance between noise removal and signal distortion. While an increased level of underfitting would produce a smoother FORC distribution, the underlying signal would be distorted to a point where it would be less consistent with the data. There are diagonal underfitting regions in Figure 4d that originate at $B_u \approx \pm 0.05$ T and converge toward $B_c \approx 0.05$ T. Such fitting artifacts can appear in lower branch subtracted data sets if there are rapid magnetization changes close to the coercive field of the hysteresis loop and can be corrected by limiting the size of local regression regions in these areas. Such functionality is available in the VARIFORC package of Egli (2013).
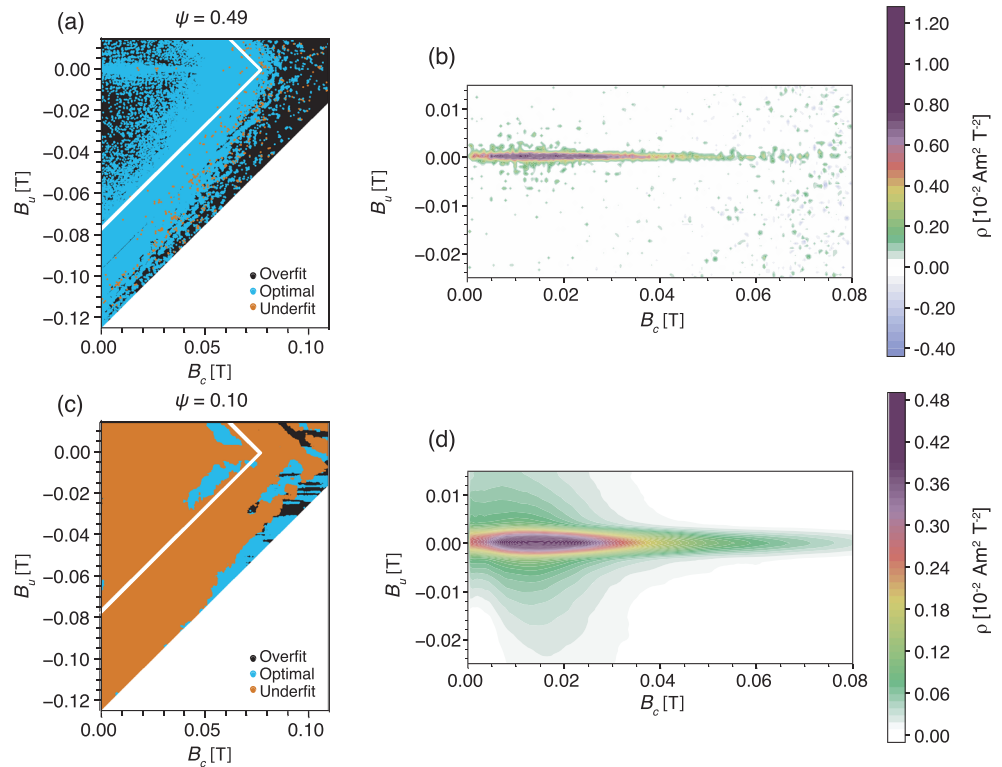
Figure 3. (a, b) Example of a predominantly overfitted model produced by reducing the level of VARIFORC smoothing to $s_{c,0} = 2$, $s_{u,0} = 2$, $s_{c,1} = 2$, $s_{u,1} = 2$, and $\lambda = 0.0$. (a) Color-coded distribution of selected model order across the space of the MD00-2361 sample FORC distribution. (b) The corresponding FORC distribution is overfitted and is unduly influenced by noise. (c, d) Example of a predominantly underfitted model produced by increasing the level of VARIFORC smoothing to $s_{c,0} = 8$, $s_{u,0} = 8$, $s_{c,1} = 8$, $s_{u,1} = 8$, and $\lambda = 0.12$. (c) Color-coded distribution of selected model order across the space of the MD00-2361 sample FORC distribution. (d) The corresponding FORC distribution is underfitted and is unduly distorted by smoothing.

The Bayesian regression framework provides uncertainties for the estimated polynomial coefficients, which can be converted readily to confidence intervals on $\hat{\rho}$ via standard uncertainty propagation (Heslop & Roberts, 2012). Therefore, once an independent estimation of the appropriate VARIFORC smoothing parameters is made, additional quantification of the FORC distribution can be undertaken. For example, profiles of $\hat{\rho}$ with associated confidence intervals can be constructed (Figure 4f).

To further demonstrate our proposed approach, we provide examples involving the addition of noise to the UB199B lower branch subtracted FORCs. As expected, higher smoothing parameters are selected as noise is added to the data (Figure 5). However, with the addition of noise and the need for greater smoothing, artifacts start to appear in the FORC distribution, for example, streaking (compare Figures 4e and 5f). It is important to note that FORC processing algorithms may not be able to resolve appropriate smoothing parameters when FORC data are corrupted by high noise levels. Our approach is designed to find a balance between noise removal and minimizing signal distortion. When noise is high, it may be infeasible to find smoothing parameters that remove the noise contribution without overly distorting the signal. Therefore, users should strive to reduce experimental noise by making the best possible measurements (Roberts et al., 2014).

4. Discussion

In our case studies, we focus on selecting smoothing parameters for the VARIFORC framework of Egli (2013). More broadly, however, our proposed approach provides an intuitive performance metric that can be applied readily to compare different local regression schemes. In our approach, smoothing parameters that yield the highest ψ value (i.e., the greatest proportion of second-order cases) can be determined automatically. Use of ψ requires further justification because it is reasonable to ask whether with the
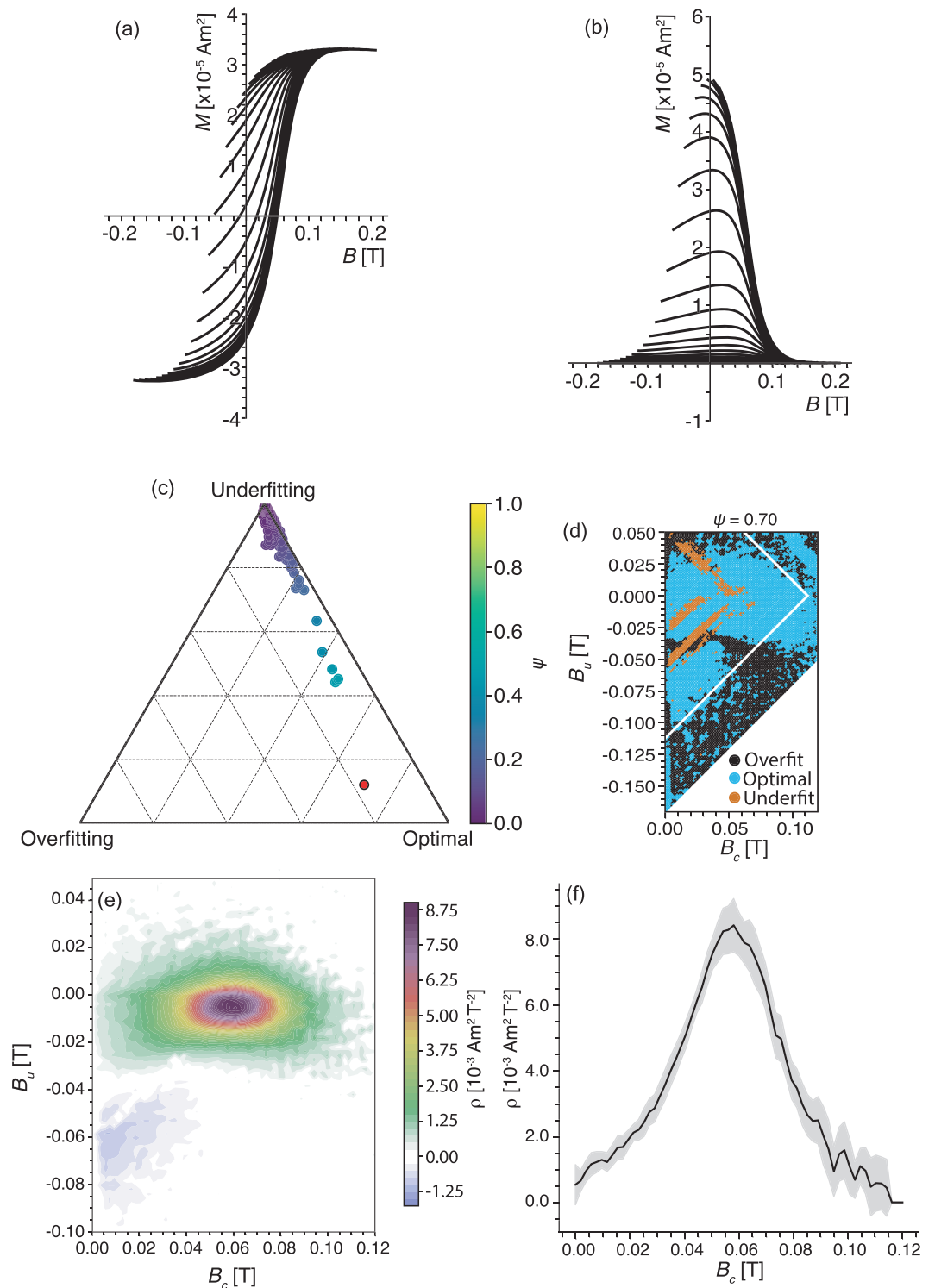
Figure 4. (a) High-field slope-corrected FORCs for sedimentary greigite sample UB199B (Roberts & Turner, 1993; Roberts et al., 2014). Every fifth FORC is plotted for clarity. (b) Lower branch subtracted curves estimated from the FORCs shown in panel (a). (c) As in Figure 1b for the UB199B lower branch subtracted FORCs, model comparison was based on 2,000 random locations within the triangular region (white line) defined by B_{open} in panel (d). (d) As in Figure 1c for the UB199B lower branch subtracted FORC curves. (e) As in Figure 1d for the UB199B lower branch subtracted FORCs with VARIFORC parameters: $s_{c,0} = 2$, $s_{u,0} = 2$, $s_{c,1} = 2$, $s_{u,1} = 2$, and $\lambda = 0.0$. (f) Coercivity profile estimated through the FORC distribution for sample UB199B at $B_u = 0$ T. Shading corresponds to the 95% confidence interval around ρ (Heslop & Roberts, 2012).
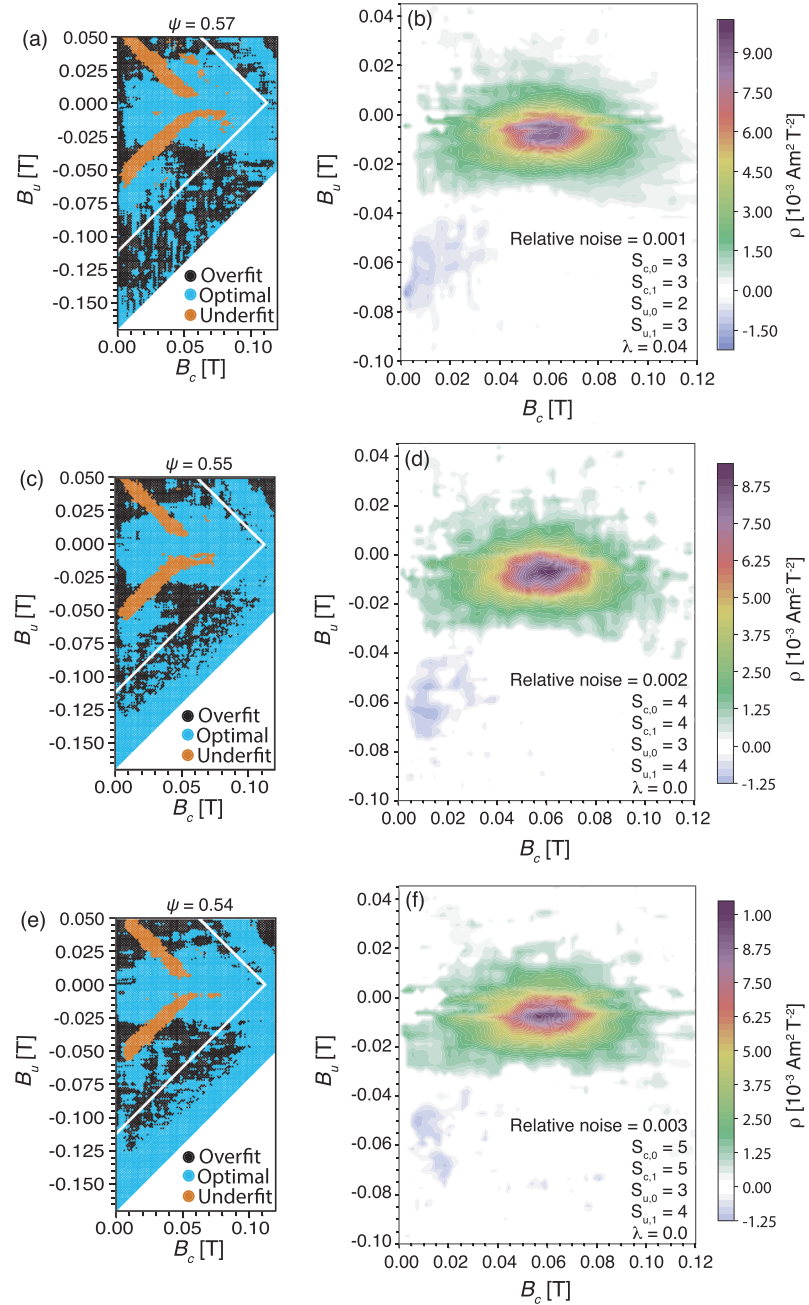
Figure 5. Comparison of the effects of adding Gaussian noise to the UB199B lower branch subtracted data with a standard deviation (σ_n) defined relative to the maximum data value. VARIFORC parameters were then estimated for cases with $\sigma_n = 0.001$, 0.002, and 0.003. (a) Color-coded distribution of selected model order across the space of the UB199B lower branch subtracted sample FORC distribution with $\sigma_n = 0.001$. (b) The corresponding FORC distribution and VARIFORC parameters for $\sigma_n = 0.001$. (c) Distribution of selected model order for $\sigma_n = 0.002$ and (d) the corresponding FORC distribution. (e) Distribution of selected model order for $\sigma_n = 0.003$ and (f) the corresponding FORC distribution. As expected, as the magnitude of the added noise increases, the selected VARIFORC smoothing parameters also increase to compensate for the reduced signal to noise ratio.

polynomial model selection framework outlined in section 2.1 it would be more appropriate to simply estimate ρ for each point in the FORC space based on the optimal local polynomial. There is a key hurdle to such an approach. Model selection techniques should be parsimonious and only favor more complex (i.e., higher-order) models if sufficient information (i.e., data) is available to support them (MacKay, 1992). Therefore, the size of a local regression region and selected polynomial order are not independent of each other. If we were to simply employ the optimal polynomial model in each local problem, how do we initially select the size of the local regression region on which the model selection will be based? To provide a concrete example of this issue, consider an analysis that employs small boxes for the local regression problem. Little information is available in each box, so higher-order polynomials cannot be supported, and H_2^0 would be selected for each box, leading to the inference that $\rho = 0$ throughout the FORC space. Clearly, such a result would be meaningless and simply an outcome of selecting an arbitrary region size a priori. We have circumvented this problem by assuming the order of the local polynomials with which ρ will be estimated (i.e., H_2, which is common to other FORC regression schemes) and searching for the VARIFORC scheme that produces regions most consistent with this assumption. Our ψ metric measures directly this level of consistency. An ability to estimate a FORC distribution where the effects of underfitting and overfitting are minimized also provides an independent technique to ensure that confidence intervals on ρ can be estimated rigorously.

It is important to note that the FORC distributions produced by our model selection-based approach may appear noisier than those typically published in the literature. We do not consider this to be a shortcoming of our algorithm; rather, it is an indication that user-defined smoothing will tend to underfit FORC data. This is not surprising. Noise in FORC distributions is easy to identify visually, while the distortion of the underlying signal is more subtle. Therefore, it is natural for users to oversmooth their data to reduce the visible effect of noise without having an obvious means to judge the corresponding signal distortion. This emphasizes the importance of making high-quality FORC measurements, where possible. Removal of noise comes at the cost of potentially distorting the underlying signal, and advanced FORC processing schemes can only be expected to reduce, but not solve entirely, this trade-off.

We have entered an era of high-resolution FORC analysis, with larger data sets containing in excess of 10^5 measurement points. This makes estimation of FORC smoothing using our proposed ψ metric computationally intensive. However, an ability to quantify overfitting and underfitting as a function of position in a FORC distribution can guide users toward an appropriate solution. Our proposed ψ metric is used to perform a grid search of candidate smoothing parameters to automatically select a suitable VARIFORC scheme for a given FORC data set. Such an approach is computationally intensive even for a simplified VARIFORC scheme defined by five parameters. We, therefore, recommend estimating a downsampled version of ψ, where all FORC measurement points are included in the local regression analysis, but the model comparison is only performed at a limited number of locations through the triangle of FORC space defined by B_{open}. Although downsampling reduces computation time, ψ estimation becomes more uncertain as fewer data points are included in its evaluation. Tests on a variety of samples suggest that a downsampled ψ estimate should involve approximately 2,000 points or more. FORCsensei employs Dask so that searching through the ensemble of smoothing parameters is performed automatically in a parallel fashion. To give a concrete example, the results in Figure 4 involved a data set with 15,364 measurement points, and FORCsensei considered 2,646 VARIFORC schemes and performed model comparison at 2,000 locations. On a 2015 MacBook Pro with a 2.8 GHz processor and 16 GB of memory, this analysis took 30 min. While this might seem like a long time compared to some accelerated algorithms (Berndt & Chang, 2019; Heslop & Muxworthy, 2005), FORCsensei is automatic and does not rely on subjective user-defined smoothing parameters.

5. Conclusions

FORC distributions have traditionally been estimated via a collection of local second-order polynomial models. Thus, if a given local regression region is most consistent with a lower-order polynomial, it indicates that overfitting may occur when the FORC distribution is estimated based on a second-order polynomial. Similarly, if local regression data are most consistent with a higher-order polynomial, then underfitting may occur when the FORC distribution is estimated. We have developed an intuitive approach to aid FORC distribution estimation, which is based on measuring the tendency of a given local regression scheme to overfit

or underfit measured magnetization data. Bayesian model selection is employed to rank the consistency of a collection of candidate polynomial regression models with the data in a local regression scheme. This information is combined into a metric, ψ, that provides a measure of the appropriateness of the local regression scheme. Thus, ψ provides a means to compare automatically the suitability of different local regression schemes and smoothing parameters to estimate the FORC distribution of a given specimen.

The FORCsensei software automates the process of VARIFORC smoothing parameter selection. Regardless of user expertise, an appropriate FORC distribution can be estimated that has a balance between noise cancelation and preservation of the underlying signal. Such an independent technique also enables rigorous estimation of confidence intervals on ρ. FORCsensei is not, however, a panacea to all FORC processing problems. FORC measurements must be of high quality with appropriate resolution to resolve features of interest. Furthermore, investigators must be ready to employ more refined processing, for example, using the full VARIFORC approach of Egli (2013), to resolve sample-specific artifacts.

Appendix A: Bayesian Model Comparison

In the theorem of Bayes (1763), the posterior probability of a hypothesis, H, given a collection of observations, X, is

$$p(H|X) = \frac{p(H)p(X|H)}{p(X)}, \tag{A1}$$

where $p(X)$ is the probability of observing X independently of any specific hypothesis and $p(H)$ is the probability of hypothesis H being true prior to any observations being made. Thus, the relative probabilities of two competing hypotheses, H_A and H_B, can be estimated by (Gallagher et al., 2009; Sambridge et al., 2006):

$$\frac{p(H_A|X)}{p(H_B|X)} = \frac{p(H_A)}{p(H_B)} \cdot \frac{p(X|H_A)}{p(X|H_B)}. \tag{A2}$$

The BF is the final term in Equation A2 and can be expressed directly as (Burnham & Anderson, 2002)

$$\mathrm{BF}(H_A, H_B) = \frac{p(X|H_A)}{p(X|H_B)} = \frac{p(H_A|X)}{p(H_B|X)} \Big/ \frac{p(H_A)}{p(H_B)}. \tag{A3}$$

Given the observed data, the BF measures the relative posterior probabilities of H_A and H_B being true. Recent work has developed approaches for estimating BF for a given linear regression scheme, such as those defined in Equations 4, 6, and 8, versus a null model where the observations are independent of the covariates. Specifically, the BF when comparing a linear regression model, H_i, to the null model, H_0, can be given in a closed-form (Maruyama & George, 2011; Wang & Maruyama, 2018):

$$\mathrm{BF}_{i0} = \frac{p(X|H_i)}{p(X|H_0)} = \frac{\Gamma(j/2 + a + 1)\Gamma((n - j - 1)/2)}{\Gamma(a + 1)\Gamma((n - 1)/2)} (1 - R^2)^{-(n-j-1)/2+a+1}, \tag{A4}$$

where Γ denotes the gamma function, n is the number of observations, R^2 is the unadjusted coefficient of determination of the linear regression model defined by H_i, j is the number of terms included in H_i, and a is an adjustable parameter (set to $a = -3/4$, as recommended by Maruyama & George, 2011). When estimating a FORC distribution via local regression, the null model, H_0, is a zero-order polynomial surface:

$$M(B, B_r) = a_1. \tag{A5}$$

Once BFs are calculated comparing each model to H_0, the values can be combined to provide pairwise comparisons for specific models. For example, $\mathrm{BF}(H_2, H_0)$ compares H_2 to H_0, and $\mathrm{BF}(H_3, H_0)$ compares H_3 to H_0, which can be combined to compare H_2 to H_3:

$$\mathrm{BF}(H_2, H_3) = \frac{\mathrm{BF}(H_2, H_0)}{\mathrm{BF}(H_3, H_0)}. \tag{A6}$$

If we assume a priori that each of the proposed models is equally probable, the BF can be used directly for probabilistic model selection.

Data Availability Statement

The data presented in this paper were processed using the open-source FORCsensei package, which is available freely from https://forcaist.github.io. The FORCsensei package code, including the FORC measurement data files used in this study, are archived on Zenodo (https://doi.org/10.5281/zenodo.4049014).

References

Acton, G., Roth, A., & Verosub, K. L. (2007). Analyzing micromagnetic properties with FORCIT software. *Eos*, *88*, 230. https://doi.org/10.1029/2007EO210004

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418. https://doi.org/10.1098/rstl.1763.0053

Berndt, T. A., & Chang, L. (2019). Waiting for Forcot: Accelerating FORC processing 100X using a fast-Fourier-transform algorithm. *Geochemistry, Geophysics, Geosystems*, *20*, 6223–6233. https://doi.org/10.1029/2019GC008380

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Berlin: Springer-Verlag Inc.

Carvallo, C., Muxworthy, A. R., Dunlop, D. J., & Williams, W. (2003). Micromagnetic modeling of first-order reversal curve (FORC) diagrams for single-domain and pseudo-single-domain magnetite. *Earth and Planetary Science Letters*, *213*, 375–390. https://doi.org/10.1016/S0012-821X(03)00320-0

Carvallo, C., Roberts, A. P., Leonhardt, R., Laj, C., Kissel, C., Perrin, M., & Camps, P. (2006). Increasing the efficiency of paleointensity analyses by selection of samples using first-order reversal curve diagrams. *Journal of Geophysical Research*, *111*, B12103. https://doi.org/10.1029/2005JB004126

Cimpoesu, D., Dumitru, I., & Stancu, A. (2019). doFORC tool for calculating first-order reversal curve diagrams of noisy scattered data. *Journal of Applied Physics*, *125*, 23906. https://doi.org/10.1063/1.5066445

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*, 829–836. https://doi.org/10.2307/2286407

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*, 596–610. https://doi.org/10.2307/2289282

Dask Development Team (2016). Dask: Library for dynamic task scheduling. https://dask.org

Egli, R. (2013). VARIFORC: An optimized protocol for the calculation of non-regular first-order reversal curve (FORC) diagrams. *Global and Planetary Change*, *110*, 302–320. https://doi.org/10.1016/j.gloplacha.2013.08.003

Egli, R., Chen, A. P., Winklhofer, M., Kodama, K. P., & Horng, C.-S. (2010). Detection of noninteracting single domain particles using first-order reversal curve diagrams. *Geochemistry, Geophysics, Geosystems*, *11*, Q01Z11. https://doi.org/10.1029/2009GC002916

Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., & Stephenson, J. (2009). Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for earth science problems. *Marine and Petroleum Geology*, *26*, 525–535. https://doi.org/10.1016/j.marpetgeo.2009.01.003

Gull, S. F. (1989). *Maximum entropy and Bayesian methods, Cambridge, 1988 Edited by* Skilling, J., Developments in maximum entropy data analysis, pp. 53–71: Kluwer, Dordrecht.

Harrison, R. J., & Feinberg, J. M. (2008). FORCinel: An improved algorithm for calculating first-order reversal curve distributions using locally weighted regression smoothing. *Geochemistry, Geophysics, Geosystems*, *9*, Q05016. https://doi.org/10.1029/2008GC001987

Harrison, R. J., & Lascu, I. (2014). FORCulator: A micromagnetic tool for simulating first-order reversal curve diagrams. *Geochemistry, Geophysics, Geosystems*, *15*, 4671–4691. https://doi.org/10.1002/2014GC005582

Harrison, R. J., Muraszko, J., Heslop, D., Lascu, I., Muxworthy, A. R., & Roberts, A. P. (2018). An improved algorithm for unmixing first-order reversal curve diagrams using principal component analysis. *Geochemistry, Geophysics, Geosystems*, *19*, 1595–1610. https://doi.org/10.1029/2018GC007511

Harrison, R. J., Zhao, X., Hu, P., Sato, T., Heslop, D., Muxworthy, A. R., et al. (2019). Simulation of remanent, transient, and induced FORC diagrams for interacting particles with uniaxial, cubic, and hexagonal anisotropy. *Journal of Geophysical Research: Solid Earth*, *124*, 12,404–12,429. https://doi.org/10.1029/2019JB018050

Heslop, D., & Muxworthy, A. R. (2005). Aspects of calculating first-order reversal curve distributions. *Journal of Magnetism and Magnetic Materials*, *288*, 155–167. https://doi.org/10.1016/j.jmmm.2004.09.002

Heslop, D., & Roberts, A. P. (2012). Estimation of significance levels and confidence intervals for first-order reversal curve diagrams. *Geochemistry, Geophysics, Geosystems*, *13*, Q12Z40. https://doi.org/10.1029/2012GC004115

Heslop, D., Roberts, A. P., & Chang, L. (2014). Characterizing magnetofossils from first-order reversal curve (FORC) central ridge signatures. *Geochemistry, Geophysics, Geosystems*, *15*, 2170–2179. https://doi.org/10.1002/2014GC005291

Heslop, D., Roberts, A. P., Chang, L., Davies, M., Abrajevitch, A., & De Deckker, P. (2013). Quantifying magnetite magnetofossil contributions to sedimentary magnetizations. *Earth and Planetary Science Letters*, *382*, 58–65. https://doi.org/10.1016/j.epsl.2013.09.011

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, *14*, 382–401. https://doi.org/10.1214/ss/1009212519

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. https://doi.org/10.2307/2291091

Lanci, L., & Kent, D. V. (2018). Forward modelling of thermally activated single-domain particles applied to first-order reversal curves. *Journal of Geophysical Research: Solid Earth*, *123*, 3287–3300. https://doi.org/10.1002/2018JB015463

Lascu, I., Einsle, J. F., Ball, M. R., & Harrison, R. J. (2018). The vortex state in geologic materials: A micromagnetic perspective. *Journal of Geophysical Research: Solid Earth*, *123*, 7285–7304. https://doi.org/10.1029/2018JB015909

Lascu, I., Harrison, R. J., Li, Y., Muraszko, J. R., Channell, J. E. T., Piotrowski, A. M., & Hodell, D. A. (2015). Magnetic unmixing of first-order reversal curve diagrams using principal component analysis. *Geochemistry, Geophysics, Geosystems*, *16*, 2900–2915. https://doi.org/10.1002/2015GC005909

Loader, C. (1999). *Local regression and likelihood*. New York: Springer.

MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation*, *4*, 415–447. https://doi.org/10.1162/neco.1992.4.3.415

Maruyama, Y., & George, E. I. (2011). Fully Bayes factors with a generalized g-prior. *The Annals of Statistics*, *39*, 2740–2765. https://doi.org/10.1214/11-AOS917

Mayergoyz, I. D. (1986). Mathematical models of hysteresis. *Physical Review Letters*, *56*, 1518–1561. https://doi.org/10.1103/PhysRevLett.56.1518

Miyazawa, K., Okamoto, S., Yomogita, T., Kikuchi, N., Kitakami, O., Toyoki, K., et al. (2019). First-order reversal curve analysis of a Nd-Fe-B sintered magnet with soft X-ray magnetic circular dichroism microscopy. *Acta Materialia*, *162*, 1–9. https://doi.org/10.1016/j.actamat.2018.09.053

Muxworthy, A. R., Heslop, D., & Williams, W. (2004). Influence of magnetostatic interactions on first-order reversal curve (FORC) diagrams: A micromagnetic approach. *Geophysical Journal International*, *158*, 888–897. https://doi.org/10.1111/j.1365-246X.2004.02358.x

Muxworthy, A. R., King, J. G., & Heslop, D. (2005). Assessing the ability of first-order reversal curve (FORC) diagrams to unravel complex magnetic signals. *Journal of Geophysical Research*, *110*, B01105. https://doi.org/10.1029/2004JB003195

Newell, A. J. (2005). A high-precision model of first-order reversal curve (FORC) functions for single-domain ferromagnets with uniaxial anisotropy. *Geochemistry, Geophysics, Geosystems*, *6*, Q05010. https://doi.org/10.1029/2004GC000877

Papusoi, C., Srinivasan, K., & Acharya, R. (2011). Study of grain interactions in perpendicular magnetic recording media using first order reversal curve diagrams. *Journal of Applied Physics*, *110*, 83908. https://doi.org/10.1063/1.3652846

Pike, C. R., Roberts, A. P., Dekkers, M. J., & Verosub, K. L. (2001). An investigation of multi-domain hysteresis mechanisms using FORC diagrams. *Physics of the Earth and Planetary Interiors*, *126*, 11–25. https://doi.org/10.1016/S0031-9201(01)00241-2

Pike, C. R., Roberts, A. P., & Verosub, K. L. (1999). Characterizing interactions in fine magnetic particle systems using first order reversal curves. *Journal of Applied Physics*, *85*, 6660–6667. https://doi.org/10.1063/1.370176

Pike, C. R., Roberts, A. P., & Verosub, K. L. (2001). First-order reversal curve diagrams and thermal relaxation effects in magnetic particles. *Geophysical Journal International*, *145*, 721–730. https://doi.org/10.1046/j.0956-540x.2001.01419.x

Roberts, A. P., Almeida, T. P., Church, N. S., Harrison, R. J., Heslop, D., Li, Y., et al. (2017). Resolving the origin of pseudo-single domain magnetic behavior. *Journal of Geophysical Research: Solid Earth*, *122*, 9534–9558. https://doi.org/10.1002/2017JB014860

Roberts, A. P., Heslop, D., Zhao, X., & Pike, C. R. (2014). Understanding fine magnetic particle systems through use of first-order reversal curve (FORC) diagrams. *Reviews of Geophysics*, *52*, 557–602. https://doi.org/10.1002/2014RG000462

Roberts, A. P., Pike, C. R., & Verosub, K. L. (2000). First-order reversal curve diagrams: A new tool for characterizing the magnetic properties of natural samples. *Journal of Geophysical Research*, *105*, 28,461–28,475. https://doi.org/10.1029/2000JB900326

Roberts, A. P., & Turner, G. M. (1993). Diagenetic formation of ferrimagnetic iron sulphide minerals in rapidly deposited marine sediments, South Island, New Zealand. *Earth and Planetary Science Letters*, *115*, 257–273. https://doi.org/10.1016/0012-821X(93)90226-Y

Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P. (2006). Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, *167*, 528–542. https://doi.org/10.1111/j.1365-246X.2006.03155.x

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *8*, 1627–1639. https://doi.org/10.1021/ac60214a047

Stoner, E. C., & Wohlfarth, E. P. (1948). A mechanism of magnetic hysteresis in heterogeneous alloys. *Philosophical Transactions of the Royal Society*, *240*, 599–642. https://doi.org/10.1098/rsta.1948.0007

Valcu, B. F., Gilbert, D. A., & Liu, K. (2011). Fingerprinting inhomogeneities in recording media using the first-order reversal curve method. *IEEE Transactions on Magnetics*, *47*, 2988–2991. https://doi.org/10.1109/TMAG.2011.2146241

Valdez-Grijalva, M. A., & Muxworthy, A. R. (2019). First-order reversal curve (FORC) diagrams of nanomagnets with cubic magnetocrystalline anisotropy: A numerical approach. *Journal of Magnetism and Magnetic Materials*, *471*, 359–364. https://doi.org/10.1016/j.jmmm.2018.09.086

Wang, M., & Maruyama, Y. (2018). Posterior consistency of *g*-prior for variable selection with a growing number of parameters. *Journal of Statistical Planning and Inference*, *196*, 19–29. https://doi.org/10.1016/j.jspi.2017.10.007