



Estimation of significance levels and confidence intervals for first-order reversal curve distributions

David Heslop and Andrew P. Roberts

Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia
(david.heslop@anu.edu.au)

[1] First-order reversal curve (FORC) distributions provide a means with which to describe a magnetic mineral assemblage in terms of coercivities and interaction fields. In recent years the use of experimentally derived FORC distributions has increased dramatically and they are being placed in an increasingly quantitative interpretational framework. An outstanding issue for calculation and interpretation of FORC data sets is the statistical significance that can be assigned to structures in experimentally determined distributions. Without this knowledge, the selection and characterization of structures that can be deemed interpretable within a FORC distribution is a subjective process. We demonstrate how FORC processing algorithms can be adapted to provide a measure of statistical significance and a confidence interval for each point in a FORC distribution. This information can guide measurement protocols and provides a more quantitative framework for interpretation of FORC distributions.

Components: 7100 words, 8 figures.

Keywords: confidence interval; first-order reversal curves; significance level.

Index Terms: 1512 Geomagnetism and Paleomagnetism: Environmental magnetism; 1540 Geomagnetism and Paleomagnetism: Rock and mineral magnetism.

Received 20 February 2012; **Revised** 19 April 2012; **Accepted** 19 April 2012; **Published** 18 May 2012.

Heslop, D., and A. P. Roberts (2012), Estimation of significance levels and confidence intervals for first-order reversal curve distributions, *Geochem. Geophys. Geosyst.*, 13, Q12Z40, doi:10.1029/2012GC004115.

Theme: Magnetism From Atomic to Planetary Scales: Physical Principles and Interdisciplinary Applications in Geosciences and Planetary Sciences

1. Introduction

[2] Since development of an efficient measurement protocol [Pike *et al.*, 1999], first-order reversal curve (FORC) diagrams have become a popular hysteresis-based method with which to characterize magnetic mineral assemblages. FORC distributions provide a means with which to estimate the distributions of coercivities and interaction fields within a magnetic particle system and, therefore, provide information

on both mineralogical composition and domain state [Roberts *et al.*, 2000; Winklhofer and Zimanyi, 2006; Muxworthy and Roberts, 2007].

[3] Recent studies have addressed several issues concerned with both calculation and interpretation of FORC distributions. Statistical analysis of measured magnetization data now allows estimation of a FORC distribution with an optimized signal-to-noise ratio [Heslop and Muxworthy, 2005; Harrison and Feinberg, 2008]. An extended FORC

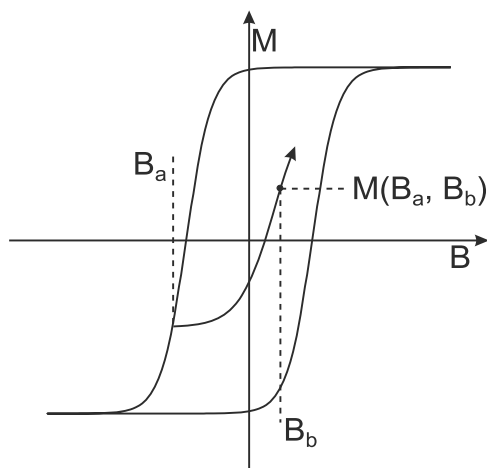


Figure 1. An example of a FORC originating from the reversal field, B_a , and measured at various values of B_b as the field returns to positive saturation.

formalism has been developed so that undefined regions of FORC distributions can be estimated via imputation [Pike, 2003; Winklhofer *et al.*, 2008]. Theoretical and experimental studies have produced type FORC distributions for systems such as superparamagnetic, single domain and multidomain ferromagnets as well as for interacting single domain systems [e.g., Pike *et al.*, 1999; Roberts *et al.*, 2000; Pike *et al.*, 2001a, 2001b; Muxworthy *et al.*, 2004; Newell, 2005; Egli, 2006; Egli *et al.*, 2010]. Artificial samples have been examined to understand the expression of mixed magnetic mineral assemblages and magnetostatic interactions in FORC distributions [Muxworthy *et al.*, 2005; Carvallo *et al.*, 2006a; Krása *et al.*, 2009, 2011]. Fitting parametric functions to profiles taken through FORC distributions has helped to elucidate the behavior of complex systems and quantify the relative abundance of different magnetic components within natural mixing systems [Egli, 2006; Chen *et al.*, 2007; Yamazaki, 2009; Egli *et al.*, 2010]. Finally, the magnetic mineral assemblages contained within a wide variety of geological, biological and extra-terrestrial materials have been characterized using the FORC method [Roberts *et al.*, 2000; Pan *et al.*, 2005; Carvallo *et al.*, 2006b; Roberts *et al.*, 2006; Acton *et al.*, 2007a; Chen *et al.*, 2007; Carvallo *et al.*, 2009; Yamazaki, 2009].

[4] In recent years, there has been an effort to place FORC data into a more quantitative framework to enable interpretation of subtle features of calculated distributions [Newell, 2005; Egli, 2006; Winklhofer and Zimanyi, 2006; Egli *et al.*, 2010]. It is therefore necessary to provide an objective assessment of the

fidelity of structures observed in experimental FORC distributions. Without such information, the decision as to which features of a FORC distribution are a real expression of the magnetic particle system becomes a subjective decision.

[5] We demonstrate how FORC algorithms can be adapted to provide measures of statistical significance and confidence intervals for each point in a calculated distribution. Significance levels can be mapped across a FORC distribution to quantify the structure of large-scale characteristic features, for example, the field to which a central ridge produced by non-interacting single domain grains propagates [Egli *et al.*, 2010]. When fitting parametric functions to FORC data it is essential to consider the uncertainty associated with the points in the selected profile and thus the level of correspondence that is required between the model and data [Egli, 2006; Chen *et al.*, 2007; Yamazaki, 2009; Egli *et al.*, 2010]. Confidence intervals can be assigned to each point in a profile enabling a realistic model to be formulated that takes into account the uncertainty in the data. Finally, averaging multiple FORC runs is considered to be an effective means with which to increase the signal-to-noise ratio of a FORC distribution [Egli *et al.*, 2010]. Significance levels and confidence intervals can guide the decision of how many runs must be made to produce an averaged FORC distribution with sufficient fidelity for the task at hand. For example, if the aim of the analysis is to produce profiles that elucidate the nature of a mixed magnetic mineral assemblage, runs can be repeated until the confidence intervals associated with the profile are sufficiently narrow to allow construction of a meaningful mixing model. We demonstrate our approach with a number of examples.

2. Estimation of a FORC Distribution

[6] Estimation of a FORC distribution is based on measurement of a collection of FORCs. An individual FORC is measured by first saturating a sample in a positive magnetic field and then decreasing the field to a pre-defined reversal field, B_a . From this reversal point, the field is once again increased and the magnetization, M , is measured at each applied field, B_b (Figure 1). The magnetization at any given point along a FORC is therefore a function of both the reversal field and the applied fields. A FORC diagram is constructed by measuring a suite of FORCs across a range of B_a values, which yields a grid of magnetization data: $M(B_a, B_b)$.

[7] A FORC distribution is given by the mixed second derivative of the magnetization data with respect to B_a and B_b [Mayergoyz, 1986]:

$$\rho(B_a, B_b) = -\frac{1}{2} \frac{\partial^2 M(B_a, B_b)}{\partial B_a \partial B_b}, \quad (1)$$

where ρ is only well defined for $B_b > B_a$ and is displayed in a rotated $\{B_c, B_u\}$ coordinate system given by $B_c = (B_a - B_b)/2$ and $B_u = (B_a + B_b)/2$ [Pike *et al.*, 1999]. When estimated directly from a measured data set, ρ will usually have a low signal-to-noise ratio, thereby hindering interpretation of the FORC distribution. To overcome this issue, Pike *et al.* [1999] employed a local second-order polynomial surface that is fitted to a regular grid of points surrounding a point of interest using ordinary least squares (OLS). The polynomial surface takes the form $a_1 + a_2 B_a + a_3 B_a^2 + a_4 B_b + a_5 B_b^2 + a_6 B_a B_b$; multiplication of the a_6 coefficient (associated with the mixed $B_a B_b$ term) by -0.5 provides an estimate of ρ . The size of the local grid is controlled by a smoothing factor (SF), with larger values yielding a smoother FORC distribution [Roberts *et al.*, 2000]. The optimal value of SF provides a balance between removing noise from the FORC distribution while ensuring that the structure of the distribution is not distorted by excessive smoothing [Heslop and Muxworthy, 2005; Harrison and Feinberg, 2008]. The FORCIT software package of Acton *et al.* [2007b] provides a full implementation of the Pike *et al.* [1999] approach.

[8] An alternative locally weighted polynomial regression (LOESS) approach to estimate ρ was presented by Harrison and Feinberg [2008]. While Pike *et al.* [1999] used a regular grid of points, Harrison and Feinberg [2008] selected points around a location of interest on a nearest-neighbor basis. A second-order polynomial surface is then fitted to the selected points using weighted least squares (WLS), where the weight of each point decreases with increasing distance from the location of interest. The approaches of Pike *et al.* [1999] and Harrison and Feinberg [2008] are used routinely to estimate FORC distributions.

[9] Additional consideration has been given to the measurement resolution required to identify specific features in a FORC distribution [Egli *et al.*, 2010] and methods to select an optimal value of SF for a given data set [Heslop and Muxworthy, 2005; Harrison and Feinberg, 2008]. Little attention, however, has been given to the statistical significance of features within a FORC distribution. As interpretation of FORC distributions becomes

more quantitative, it is essential that subtle features can be demonstrated to be statistically significant and not simply an artifact of measurement noise or the algorithm employed to estimate ρ . In this paper we demonstrate how significance levels and confidence intervals can be calculated for estimated values of ρ from the approaches of Pike *et al.* [1999] and Harrison and Feinberg [2008]. The significance levels can be mapped in the $\{B_c, B_u\}$ plane to illustrate which parts of a calculated FORC distribution are significantly non-zero, while the confidence intervals provide crucial information when profiles through a distribution are analyzed.

3. Estimating $\rho(B_a, B_b)$ and Its Significance Level

[10] We consider the regression problem of fitting a second-order polynomial surface to a collection of $M(B_a, B_b)$ points using the OLS approach of Pike *et al.* [1999] and WLS approach of Harrison and Feinberg [2008]. The initial step in both techniques is to select n points around a location of interest in the $\{B_a, B_b\}$ plane (where n is a function of SF). A design matrix, \mathbf{X} , is then constructed to contain the regressors, with each row of the matrix containing a constant and field terms ($[1, B_a, B_a^2, B_b, B_b^2, B_a B_b]$) for one of the n selected points. The magnetization values of the n selected points are then placed in the column vector, \mathbf{y} . Using matrix notation, the magnetization values are related to the design matrix by the linear model:

$$\mathbf{y} = \mathbf{X} \mathbf{a} + \mathbf{e}, \quad (2)$$

where \mathbf{a} is a vector containing the coefficients a_1 to a_6 and \mathbf{e} is a n -by-1 vector of errors. An estimate of the vector of regression coefficients, $\hat{\mathbf{a}}$, then provides an estimate of $\rho(B_a, B_b)$ from \hat{a}_6 . Because \hat{a}_6 is only an estimate of the true value of a_6 and thus of $\rho(B_a, B_b)$, it is essential to examine the result statistically. Specifically, this requires calculation of a p -value for the null hypothesis that $a_6 = 0$ and thus that $\rho(B_a, B_b) = 0$. To test the null hypothesis, an estimate of the error on \hat{a}_6 must be made.

[11] It is important to consider the ability of a local second-order polynomial surface to represent the FORC function of a collection of sampled points. Generally such a trend-surface should perform well in most regions of the FORC diagram providing that data have been collected at sufficient resolution. In such regions the \mathbf{e} term in the linear model (equation (2)) should correspond to the measurement errors. The central ridge of a

FORC distribution (produced by noninteracting single domain particles) is a more complicated case. Because the ridge results from a discontinuous FORC function, its form cannot be represented using a second-order polynomial surface [Egli *et al.*, 2010]. Therefore part of the FORC signal will be incorrectly assigned to \mathbf{e} because the fitted function provides an inadequate representation of the true function. The implications of this issue will be considered in more detail below.

[12] The linear model presented in equation (2) could be extended easily to fit FORC diagrams using higher-order polynomial functions with more flexible surfaces. This would simply require addition of appropriate terms to the design matrix and a corresponding adjustment to the degrees-of-freedom in the statistical analysis presented below.

3.1. Ordinary Least Squares

[13] Using the OLS approach, a solution to equation (2) is found by minimizing the sum of squared errors:

$$\mathbf{e}^T \mathbf{e} = (\mathbf{y} - \mathbf{X}\hat{\mathbf{a}})^T (\mathbf{y} - \mathbf{X}\hat{\mathbf{a}}), \quad (3)$$

which yields:

$$\hat{\mathbf{a}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (4)$$

The measured magnetizations of the selected points, \mathbf{y} , and their estimates from the regression model, $\hat{\mathbf{y}}$, can be related by the so-called ‘‘hat’’ matrix, \mathbf{H} . If $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$ then:

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T. \quad (5)$$

The mean squared error, MSE, of the fitted regression model can be found using the hat matrix:

$$\text{MSE} = \frac{\mathbf{y}^T (\mathbf{I} - \mathbf{H}) \mathbf{y}}{n - 6}, \quad (6)$$

where \mathbf{I} is a n -by- n identity matrix. Once the MSE has been found, the variance-covariance matrix of the regression coefficients can be estimated:

$$\hat{\Sigma}_{\hat{\mathbf{a}}} = (\mathbf{X}^T \mathbf{X})^{-1} \text{MSE}. \quad (7)$$

The estimated standard error, SE, on \hat{a}_6 is then given by the square root of the sixth diagonal element of $\hat{\Sigma}_{\hat{\mathbf{a}}}$. To obtain a p -value with which to test the null hypothesis (i.e., that $a_6 = 0$), the ratio of \hat{a}_6 to the SE is compared to a Student’s t -distribution in a two-sided manner (thus taking into consideration that \hat{a}_6

can be different from 0 in two ways; $\hat{a}_6 > 0$ and $\hat{a}_6 < 0$). The p -value is given by:

$$p = 2(1 - t_{\alpha}(|\hat{a}_6|/\text{SE})), \quad (8)$$

where t represents the value of the Student’s t -distribution with $n-6$ degrees of freedom at a specified α level. The p -value gives the probability that the null hypothesis is true given the observations. Thus, if p falls below a pre-defined α significance level, the null hypothesis can be rejected and the alternative hypothesis ($a_6 \neq 0$) will be accepted. Typically, α would be set at a value of 0.05, which implies that there is a 5% chance that the test will incorrectly reject a true null hypothesis (i.e., the test returns $a_6 \neq 0$ incorrectly).

[14] By applying the OLS method in the above manner the measurement errors of the magnetization in \mathbf{y} are assumed to be uncorrelated random variables with zero means and constant variance. The assumption of errors with constant variance is termed homoscedasticity and bias can be introduced into $\hat{\Sigma}_{\hat{\mathbf{a}}}$ and in turn into SE and p , if the assumption is not met.

[15] Errors that do not have a constant variance are termed heteroscedastic and require special consideration. For a small local grid of points over which the linear model (equation (2)) is fitted, errors in the measurement of M and those related to the applied field can be assumed to be effectively homoscedastic. Instrument drift during FORC measurement will potentially induce different noise characteristics in each of the curves included in a local grid [Jackson and Solheid, 2010]. Such drift related errors must be considered to be heteroscedastic.

[16] A heteroscedasticity-consistent standard error (HCSE) estimator is adopted, which should allow robust inferences to be drawn even if the assumption of homoscedasticity is violated. The HC3 estimator [MacKinnon and White, 1985] provides a heteroscedasticity robust estimator of $\Sigma_{\hat{\mathbf{a}}}$:

$$\text{HC3} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{diag} \left[\frac{e_i^2}{(1 - h_{ii})^2} \right] \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}, \quad (9)$$

where $\mathbf{e} = \mathbf{y} - \mathbf{H}\mathbf{y}$ and h_{ii} are the diagonal elements of \mathbf{H} . Once the HC3 estimate of $\Sigma_{\hat{\mathbf{a}}}$ is made, equation (8) can be solved to obtain p . Although a variety of HCSE estimators exist, empirical investigations demonstrate that HC3 performs well when using t -tests to assess regression coefficients in both homoscedastic and heteroscedastic cases [Long and Ervin, 2000].

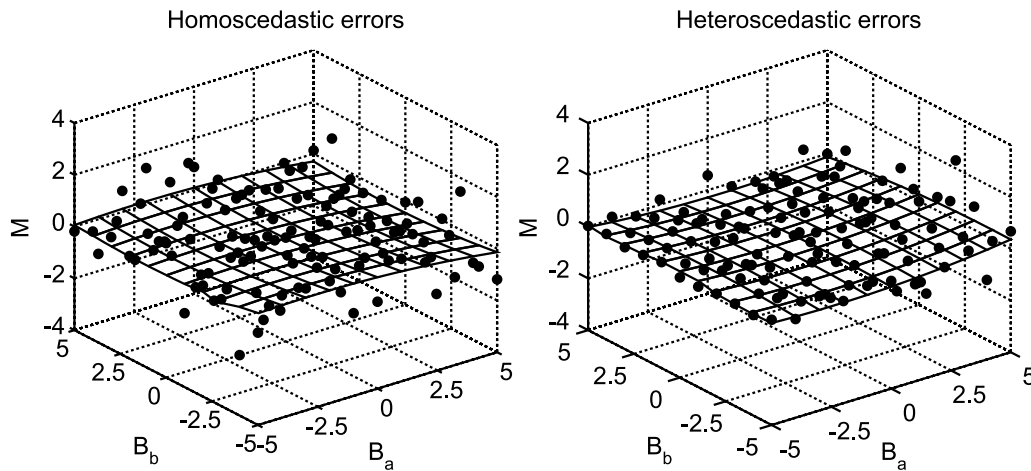


Figure 2. Examples of (left) homoscedastic and (right) heteroscedastic errors used to test for bias in the regression analysis (units of the field grids are arbitrary). The homoscedastic errors are drawn from a normal distribution defined by $\mathcal{N}(0, 1)$ and are thus independent of the applied field. The heteroscedastic errors are drawn from normal distributions, the variance of which increases as a linear function of B_a . The spatial correlation of the errors in the latter case violates the assumptions of the least squares approach and a heteroscedastic-consistent method must be adopted.

3.1.1. Testing for Bias in OLS

[17] To test the OLS approach outlined above, a numerical experiment was performed that mimicked the calculation of a FORC distribution. A regular 11×11 grid of points in the interval $[-5, 5]$ was constructed to simulate calculation of $\rho(0, 0)$ with $SF = 5$. The true magnetization value of each point on the grid was set to zero, therefore it is known that the second-order polynomial surface defined by the error free magnetizations is given by $a_1 = \dots = a_6 = 0$. Random numbers were then added to the magnetizations to simulate measurement noise. These random numbers took two forms to investigate the homoscedastic and heteroscedastic scenarios, respectively.

[18] In the homoscedastic case, random numbers were drawn from a normal distribution with a mean of 0 and a variance of 1, i.e., $\mathcal{N}(0, 1)$ (Figure 2). To represent heteroscedastic errors a simple model of instrument drift was developed. It was assumed that drift within the segment of any given FORC included in a local grid is negligible, but the influence of drift between the different FORCs is important. To represent this effect the variance of the errors associated with each FORC was increased linearly between 0.01 and 1 as a function of B_a (Figure 2). Regression coefficients were estimated for both the homoscedastic and heteroscedastic cases and their corresponding p -values for the null hypothesis ($a_6 = 0$) were determined. This procedure was repeated for 10^4 realizations

and the resulting distributions of p examined (Figure 3).

[19] If the assumptions of OLS are met then the observed p -values should be distributed uniformly in the interval $[0, 1]$, therefore their empirical cumulative distribution function, $\hat{F}(p)$, will follow a line of unity between 0 and 1. In the case of the homoscedastic errors, both the traditional OLS and the HCSE methods perform well. Bias in the p -values, estimated simply as $F(p) - \hat{F}(p)$, where $F(p)$ is a uniform cumulative distribution function, is small. In contrast, the heteroscedastic case has a positive bias in p based on calculation of $\hat{\Sigma}_{\hat{a}}$ using equation (7). Such a bias would lead to the null hypothesis $a_6 = 0$ being rejected too infrequently. The use of HC3 (equation (9)) yields a negligible bias compared to the traditional estimate (equation (7)). The good performance of HC3 regardless of the presence or absence of heteroscedasticity supports the recommendation of Long and Ervin [2000] that HCSE estimators should be used in OLS unless homoscedasticity can be guaranteed.

3.2. Weighted Least Squares

[20] In the method of Harrison and Feinberg [2008], a vector of associated weights, \mathbf{w} , form a n -by- n diagonal matrix, $\mathbf{W} = \text{diag}(\mathbf{w})$ (i.e., all the elements of \mathbf{W} are zero except along the main diagonal, which is filled with the values of \mathbf{w}). The weight, w_i , of a point, i , is given by the tri-cube

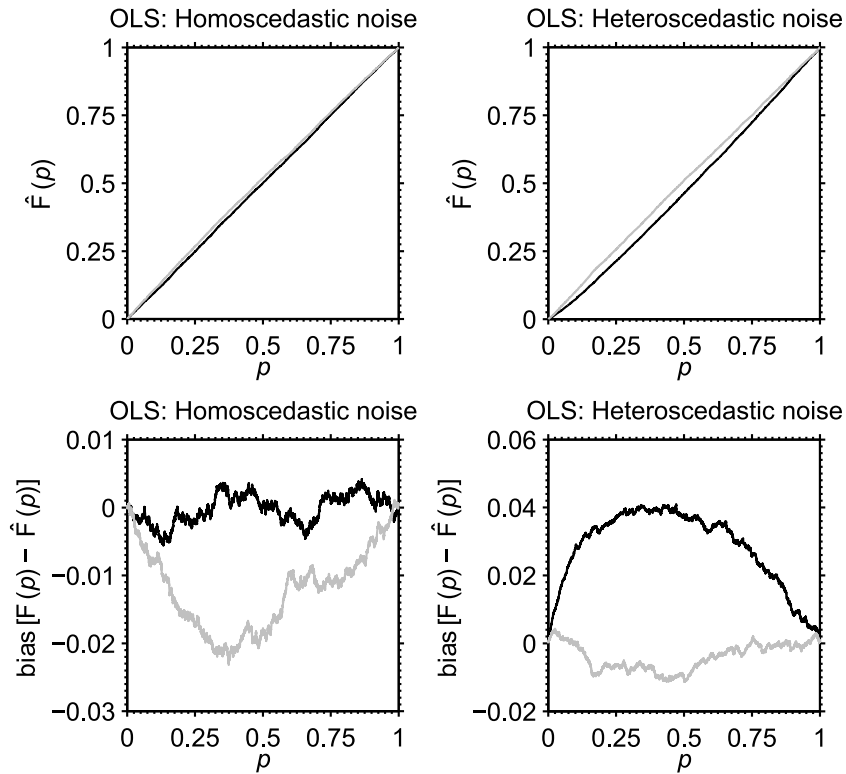


Figure 3. Empirical cumulative distributions, \hat{F} , of the p -values calculated using an OLS approach under the assumption of homoscedastic errors (black) and using the heteroscedasticity-consistent estimator, HC3 (gray). The bias of each approach can be estimated by finding the difference between the empirical distribution, $\hat{F}(p)$, and a uniform distribution, $F(p)$. The heteroscedasticity-consistent estimator performs well in both the homoscedastic and heteroscedastic cases, whereas the standard OLS approach has a bias in the presence of heteroscedasticity.

function [Cleveland, 1988; Harrison and Feinberg, 2008]:

$$w_i = \left(1 - \left[\frac{\|\bar{r} - \bar{r}_i\|}{\max\|\bar{r} - \bar{r}_i\|} \right]^3 \right)^3, \quad (10)$$

where \bar{r} defines the position in the $\{B_a, B_b\}$ plane for which ρ is to be estimated, \bar{r}_i defines the position of the point for which the weight is being calculated and $\max\|\bar{r} - \bar{r}_i\|$ is the maximum Euclidean distance between \bar{r} and the points selected in the local grid. In the WLS approach, a solution to equation (2) is found by minimizing the weighted sum of squared errors:

$$e^T W e = (y - X\hat{a})^T W (y - X\hat{a}). \quad (11)$$

The weighted least squares estimate, \hat{a} , of the vector a is given by:

$$\hat{a} = (X^T W X)^{-1} X^T W y, \quad (12)$$

and the corresponding hat matrix is:

$$H = X(X^T W X)^{-1} X^T W. \quad (13)$$

In WLS the mean squared error is:

$$\text{MSE} = \frac{y^T (W - WH) y}{n - 6}, \quad (14)$$

and the estimate of the variance-covariance matrix, $\hat{\Sigma}_{\hat{a}}$, is given by:

$$\hat{\Sigma}_{\hat{a}} = (X^T W X)^{-1} \text{MSE}. \quad (15)$$

The standard error on the WLS estimate of a_6 is given by the square root of the sixth element on the main diagonal of $\hat{\Sigma}_{\hat{a}}$ and in turn the p -value of the null hypothesis (i.e., $a_6 = 0$), can be obtained from equation (8).

[21] As with the OLS approach discussed in section 3.1, the statistics associated with the WLS method may be biased if certain assumptions are not met. In the formulation of the WLS method, the estimate \hat{a} will only be the best linear unbiased estimator (BLUE) when the weight of each point is equal to the reciprocal of the variance of the point (i.e., the reciprocal of the variance of the error on the magnetization). The tri-cube based weights do not meet this requirement and,

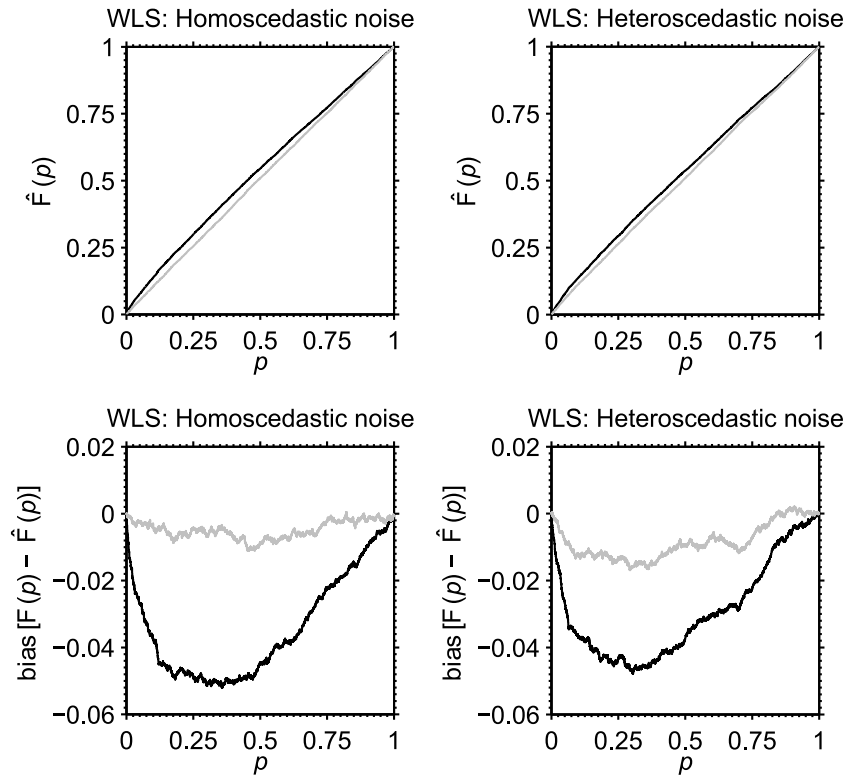


Figure 4. Empirical cumulative distributions, \hat{F} , of the p -values calculated using a WLS approach under the assumption of homoscedastic errors (black) and using the heteroscedasticity-consistent estimator, $HC3_W$ (gray). The heteroscedasticity-consistent estimator performs well in both the homoscedastic and heteroscedastic cases because of the heteroscedasticity introduced by the tri-cube weighting scheme. We conclude that it is advisable to use the $HC3_W$ estimator (equation (16)) when performing weighted least squares to calculate FORC distributions.

given their spatial autocorrelation, will induce heteroscedasticity in the estimated error terms. As with the OLS approach a HCSE can be adopted, this time in a weighted form:

$$HC3_W = (X^T W X)^{-1} X^T W \text{diag} \left[\frac{w_i e_i^2}{(1 - h_{ii})^2} \right] X (X^T W X)^{-1}. \quad (16)$$

3.2.1. Testing for Bias in WLS

[22] The numerical experiment involving homoscedastic and heteroscedastic errors (section 3.1.1) was repeated for the outlined WLS procedure with sample weights calculated with respect to the center of the grid according to equation (10). In the case of homoscedastic errors, the WLS approach using the calculation of $\hat{\Sigma}_{\hat{a}}$ in equation (15) produces a slight bias in p (Figure 4). As discussed above this bias results from use of the tri-cube based weights that tend to induce heteroscedasticity, which means that \hat{a}_6 is not BLUE. In contrast, use of $HC3_W$ reduces much of the bias.

[23] For the heteroscedastic case, the values of p obtained from $\hat{\Sigma}_{\hat{a}}$ in equation (15) exhibit a similar bias to that observed for homoscedastic errors. This correspondence is a result of the weighting procedure employed in the WLS approach. The extremes of the grid where the relative differences in the drift are greatest are down-weighted, while the center of the grid, which exhibits smaller differences in the drift, are up-weighted. For the case of heteroscedasticity the bias in the traditional WLS approach can again be largely suppressed using $HC3_W$. As with the analysis of OLS and $HC3$, it is advisable to employ $HC3_W$ in both the homoscedastic and heteroscedastic cases when performing WLS to estimate FORC distributions.

4. Statistically Significant Regions in FORC Distributions

[24] If the p -value for each point in a FORC distribution has been determined it becomes possible to map the statistically significant regions of ρ in the

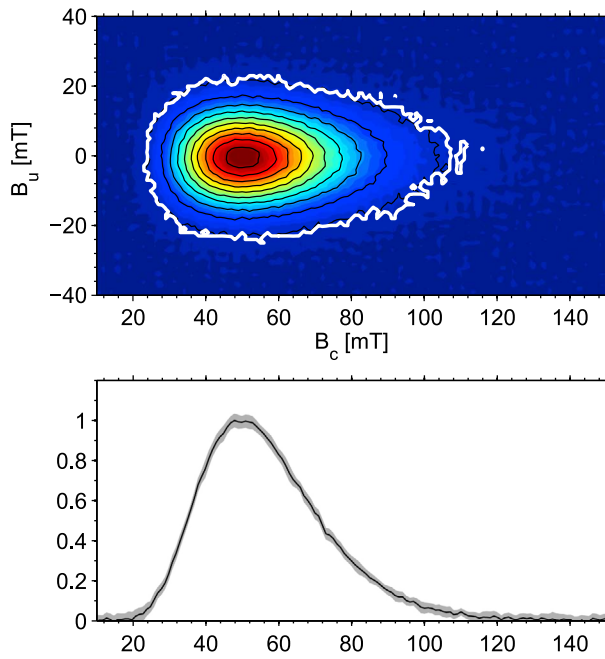


Figure 5. (top) FORC distribution calculated from an analytical Preisach model [from *Pike et al.*, 1999] with a simulated noise contribution. The thick contour line indicates the regions of the FORC distribution that are significant at the 0.05 level once the Bonferroni correction is taken into consideration. (bottom) The gray band represents the 95% confidence interval for the ρ values along $B_u = 0$ mT for the FORC distribution in Figure 5 (top). The black line represents the ρ values estimated from the data.

$\{B_c, B_u\}$ plane. Such regions should correspond to larger-scale features of the distribution and significance levels can be used to show their form and extent in terms of both B_c and B_u .

[25] Once a FORC distribution is calculated on the basis of a predetermined optimal SF, locations in the $\{B_a, B_b\}$ plane where $p < \alpha$ can be identified. As discussed in section 3.1, the value of α implies that there is a probability of α that the test will incorrectly reject a true null hypothesis (i.e., the test incorrectly returns $a_6 \neq 0$). Calculation of a FORC distribution requires multiple tests to be performed, one for each considered location in the $\{B_a, B_b\}$ plane and, by definition, a fraction of α of those tests will incorrectly reject the null hypothesis. This will result in a FORC distribution where the number of locations specified as $a_6 \neq 0$ is artificially high. To control the error-rate during the calculation, the so-called Bonferroni correction is adopted [Miller, 1981]. The correction states that if a

collection of g tests is to be made on a data set, then the significance level for each individual test should be set at α/g rather than α . With the Bonferroni correction in place, the significance level for the FORC distribution as a whole, rather than at any single location, should be α . Therefore, the region of the FORC distribution that is statistically significant at the α level can be identified by mapping $p < \alpha/g$ in the $\{B_c, B_u\}$ plane.

[26] In section 3 the inability of a local second-order polynomial surface to fit the central ridge of a FORC function was considered. If part of the FORC signal is assigned incorrectly to the model errors, the MSE will increase. In turn, the corresponding p -value will be elevated, increasing the probability that the null hypothesis ($a_6 = 0$) will be accepted incorrectly. This implies that significance levels around a central ridge must be considered carefully. However, the high values of ρ associated with central ridges suggest that the regions will remain statistically significant even when the errors are artificially high. An example of this is shown below.

[27] Finally, it is important to note that the significant regions of a FORC distribution will depend on the SF employed in the calculation of ρ . It would therefore be inappropriate to tune the SF to ensure that certain parts of the distribution become significant. Instead, an optimal SF should be preselected on the basis of an independent method [Heslop and Muxworthy, 2005; Harrison and Feinberg, 2008].

4.1. Calculation of Confidence Intervals

[28] Recent studies have shown that fitting parametric functions to profiles taken through FORC distributions, for example, along the line $B_u = 0$, provides information concerning the composition and grain size of the magnetic mineral assemblage [Egli, 2006; Chen et al., 2007; Egli et al., 2010; Roberts et al., 2011]. By providing confidence intervals on such profiles robust assessment of the level of mismatch between the data and fitted model can be achieved. For example, if confidence intervals on a profile are large, it may demonstrate that a close data-model match cannot be expected and inferences drawn from the modeling results should be treated with caution.

[29] Once the SE on \hat{a}_6 has been found, it is a simple task to calculate a $100(1 - \alpha)\%$ confidence interval that incorporates the Bonferroni correction for the estimated value of ρ :

$$\rho = -0.5(\hat{a}_6 \pm t^{-1}(\alpha/2g)SE), \quad (17)$$

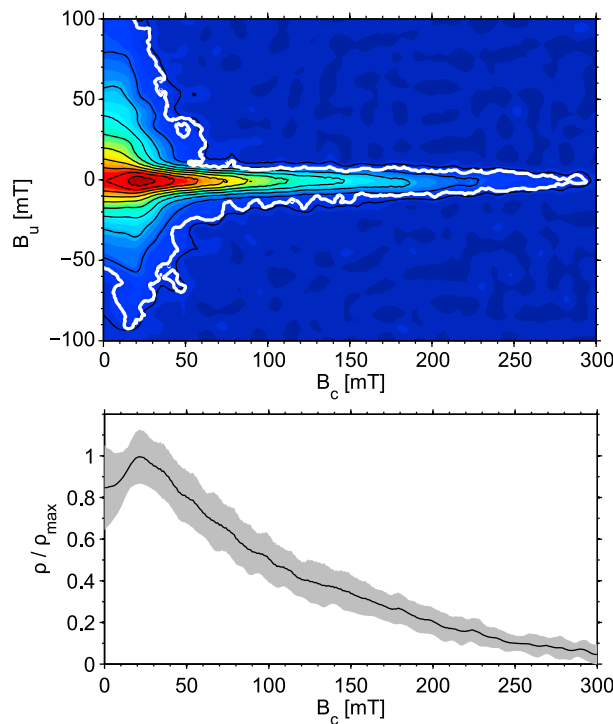


Figure 6. (top) FORC distribution (SF = 3) for the magnetite-bearing plagioclase sample reported by *Harrison and Feinberg* [2008]. The thick contour line indicates the regions of the FORC distribution that are significant at the 0.05 level once the Bonferroni correction is taken into consideration. (bottom) The gray band represents the 95% confidence interval for ρ values along $B_u = 0$ mT for the FORC distribution in Figure 6 (top). The black line represents the ρ values estimated from the data.

where g is the number of points included in the profile and t^{-1} is the inverse of the Student's t -distribution with $n - 6$ degrees of freedom.

[30] As with the discussion of significance levels it is important to consider the behavior of confidence intervals around a central ridge. Because a second-order polynomial surface is inadequate to represent the central ridge, artificially high errors will increase the confidence interval associated with a given point. In this way the calculated confidence interval is a conservative estimate of the true confidence interval.

5. Examples

[31] To demonstrate our approach, a number of brief examples are provided below. All FORC distributions were calculated using the LOESS approach of *Harrison and Feinberg* [2008], which allows estimation to $B_c = 0$ without employing the extended FORC protocol of *Pike* [2003]. Additionally,

standard errors on the calculated values of \hat{a}_6 were estimated using the weighted HCSE approach to limit the influence of heteroscedasticity.

5.1. A Basic Preisach Model

[32] A smooth FORC diagram was constructed analytically using a basic Preisach model [*Preisach*, 1935] as employed by *Pike et al.* [1999]. The diagram was constructed from hysterons with a normal distribution of local interaction fields with zero mean and a standard deviation of 9 mT and a log-normal distribution of switching fields with a mean of $\log(60$ mT) and a standard deviation of 0.3. To simulate measurement noise, random numbers were

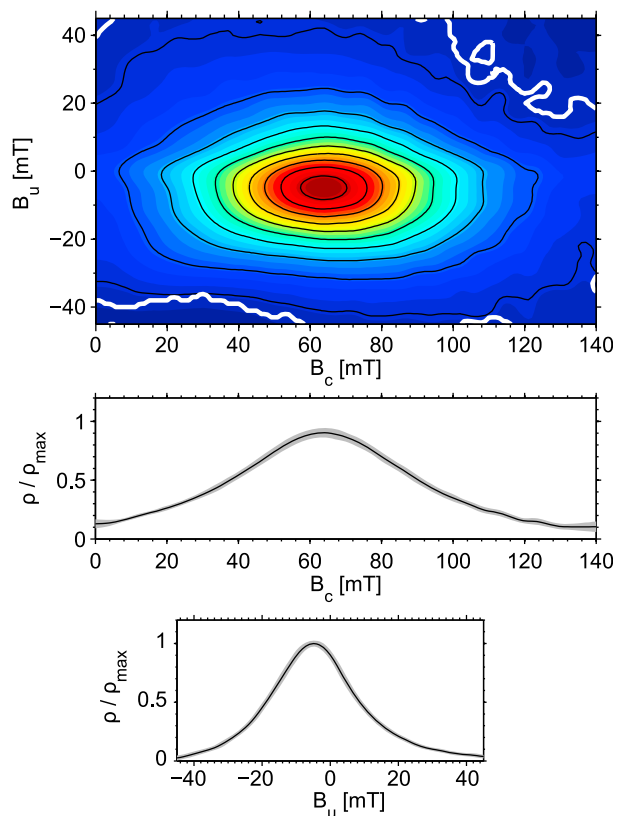


Figure 7. (top) FORC distribution (SF = 5) for a greigite-bearing sediment from the Crostolo River, Italy [*Tric et al.*, 1991; *Roberts et al.*, 2005]. The strong magnetostatic interactions within the magnetic assemblage produce the broad vertical spread of the FORC distribution, which has ρ values that are significant at the 0.05 level almost throughout the entire diagram. (middle) The gray band represents the 95% confidence interval for ρ along $B_u = -4$ mT for the FORC distribution in Figure 7 (top). The black line represents the ρ values estimated from the data. (bottom) A vertical profile of the magnetic interaction field distribution through the FORC distribution at $B_c = 65$ mT.

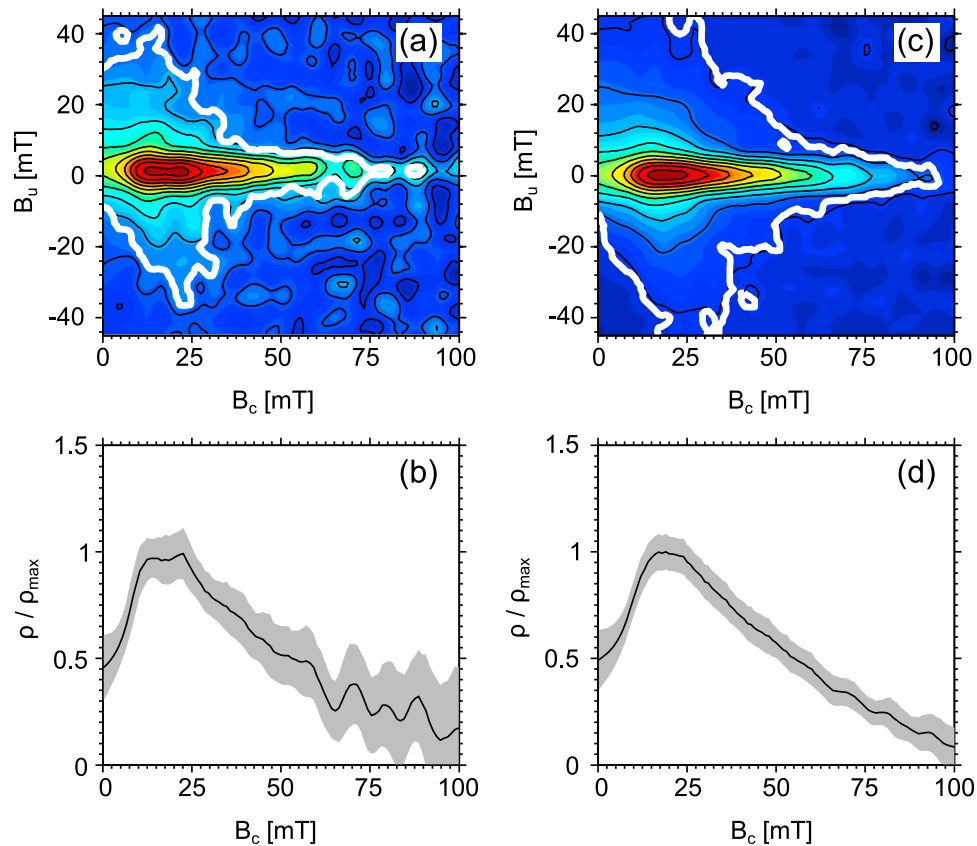


Figure 8. (a) FORC distribution (SF = 4) for a magnetically weak sample from marine sediment core MD002361. The thick contour line indicates the regions of the FORC distribution that are significant at the 0.05 level once the Bonferroni correction is taken into consideration. (b) The gray band represents the 95% confidence interval for ρ values along $B_u = 0$ mT for the FORC distribution in Figure 8a. At high values of B_c the large confidence intervals demonstrate that ρ is poorly defined and the apparent central ridge structure cannot be quantified properly. (c) FORC distribution (SF = 4) obtained by averaging nine repeat measurement runs of the same sample. The FORC distribution is defined more clearly with the statistically significant region extending further in both the horizontal and vertical directions. (d) Averaging nine FORC runs reduces the confidence interval along the $B_u = 0$ mT profile and the central ridge structure remains significant until $B_c \approx 95$ mT.

drawn from the normal distribution $\mathcal{N}(0, 10^{-8})$ and were added to the magnetization values. The calculated FORC distribution (SF = 2) is shown in Figure 5 with a contour indicating the $\alpha = 0.05$ significance level. A profile through the FORC distribution at $B_u = 0$ mT is also shown in Figure 5 with the associated 95% confidence interval for each estimated value of ρ . The lognormal distribution of switching fields used to construct the model consistently lies within the 95% confidence interval of the estimated ρ values along the profile.

5.2. Magnetite-Bearing Plagioclase

[33] In their description of the FORCInel algorithm, *Harrison and Feinberg* [2008] gave an example of a magnetite-bearing plagioclase sample with an optimal SF of ~ 3 . The calculated FORC distribution and $\alpha = 0.05$ significance levels for this sample

are shown in Figure 6. The $\alpha = 0.05$ contour follows the main body of the FORC distribution and demonstrates how far the central ridge propagates along the B_c axis. The profile of normalized ρ values along $B_u = 0$ mT has relatively wide 95% confidence intervals. This implies that any unmixing analysis based on fitting appropriate distributions to the profile should be performed with caution while taking uncertainty in the ρ values into account.

5.3. Crostolo River Section, Italy

[34] Sediments from the Crostolo River section contain a magnetic assemblage that is dominated by authigenic greigite [*Tric et al.*, 1991; *Roberts et al.*, 2005], with strong magnetic interactions (Figure 7). When calculated with an optimal SF of 5, the interacting SD greigite particles produce

a large vertical spread in the FORC distribution; even away from the main body of the distribution values remain significant at the 0.05 level (Figure 7). Profiles through the distribution along $B_u = -4$ mT and $B_c = 65$ mT reveal the extent of the distribution, the magnitude of the magnetic interactions, and the relatively small confidence intervals associated with the individual values of ρ .

5.4. Marine Sediments, Eastern Indian Ocean

[35] A sample of carbonate-rich sediment was analyzed from core MD002361, which was recovered offshore of northwestern Australia at a water depth of ~ 1800 m [Spooner *et al.*, 2011]. An age model based on correlation of $\delta^{18}\text{O}$ to a pre-existing orbital chronology suggests that the studied sediment sample originates from marine isotope stage 8 [Spooner *et al.*, 2011]. A FORC distribution for this sample (Figure 8a, SF = 4) has vertical spreading indicative of pseudosingle domain and interacting single domain particles [Roberts *et al.*, 2000; Muxworthy and Dunlop, 2002] and a central ridge that may correspond to noninteracting single domain particles [Roberts *et al.*, 2000; Egli *et al.*, 2010]. The relatively large noise contribution means that the apparent ridge structure is poorly defined at B_c values above ~ 60 mT, but ρ remains significant until $B_c \approx 80$ mT (Figure 8a). A profile along $B_u = 0$ mT confirms this observation, with large uncertainties indicating that ρ is poorly defined at higher values of B_c (Figure 8b). Additionally, a small significant region at $B_c \approx 90$ mT indicates that, if a ridge is present, it may extend further across the diagram than suggested by the main significant region.

[36] As recommended by Egli *et al.* [2010], repeat measurements of the sediment sample were undertaken to determine if the high field region of the central ridge structure was being suppressed by noise. After averaging nine measurement runs, values are clearly significant at higher B_c values and the high field part of the ridge has been recovered from the noise (Figure 8c). The averaged FORC distribution and the confidence intervals on the corresponding $B_u = 0$ mT profile (Figure 8d) indicate that the central ridge is statistically significant until $B_c \approx 95$ mT, which is in good agreement with earlier work [Egli *et al.*, 2010]. This example demonstrates that parts of a FORC distribution can be rendered ambiguous by large noise contributions. Determination of statistical significance and confidence intervals can not only identify these

regions, but also reveal if the regions become significant after repeat measurement.

6. Conclusions

[37] First-order reversal curve distributions have become a popular tool with which to characterize complex magnetic mineral assemblages. We have presented simple extensions to existing FORC processing algorithms whereby significance levels and confidence intervals can be estimated for experimentally and numerically determined distributions. This information aids interpretation and quantification of FORC distributions because it enables assessment of whether given features are statistically significant. Additionally, where parametric functions are to be fitted to profiles through FORC distributions, calculation of confidence intervals quantifies the level of uncertainty in the distribution and therefore the level of uncertainty that can be tolerated by any attempt to fit components to such distributions. We recommend use of a heteroscedasticity-consistent standard error estimator when calculating FORC distributions. Such estimators remove the necessity for an assumption of errors with constant variance that, if violated, can introduce large biases into calculated significance levels.

Acknowledgments

[38] This work was supported by the Australian Research Council (grant DP110105419). The authors are grateful to the Associate Editor and two reviewers for constructive comments that helped to improve the final manuscript.

References

- Acton, G., Q.-Z. Yin, K. L. Verosub, L. Jovane, A. Roth, B. Jacobsen, and D. S. Ebel (2007a), Micromagnetic coercivity distributions and interactions in chondrules with implications for paleointensities of the early solar system, *J. Geophys. Res.*, *112*, B03S90, doi:10.1029/2006JB004655.
- Acton, G., A. Roth, and K. L. Verosub (2007b), Analyzing micromagnetic properties with FORCIT software, *Eos Trans. AGU*, *88*, 230.
- Carvallo, C., A. R. Muxworthy, and D. J. Dunlop (2006a), First-order reversal curve (FORC) diagrams of magnetic mixtures: Micromagnetic models and measurements, *Phys. Earth Planet. Inter.*, *154*, 308–322.
- Carvallo, C., A. P. Roberts, R. Leonhardt, C. Laj, C. Kissel, M. Perrin, and P. Camps (2006b), Increasing the efficiency of paleointensity analyses by selection of samples using first-order reversal curve diagrams, *J. Geophys. Res.*, *111*, B12103, doi:10.1029/2005JB004126.
- Carvallo, C., S. Hickey, D. Faivre, and N. Menguy (2009), Formation of magnetite in *Magnetospirillum gryphwaldense*

- studied with FORC diagrams, *Earth Planets Space*, *61*, 143–150.
- Chen, A. P., R. Egli, and B. M. Moskowitz (2007), First-order reversal curve (FORC) diagrams of natural and cultured biogenic magnetite particles, *J. Geophys. Res.*, *112*, B08S90, doi:10.1029/2006JB004575.
- Cleveland, W. S. (1988), Robust locally weighted regression and smoothing scatterplots, *J. Am. Stat. Assoc.*, *74*, 829–836, doi:10.2307/2286407.
- Egli, R. (2006), Theoretical aspects of dipolar interactions and their appearance in first-order reversal curves of thermally activated single-domain particles, *J. Geophys. Res.*, *111*, B12S17, doi:10.1029/2006JB004567.
- Egli, R., A. P. Chen, M. Winklhofer, K. P. Kodama, and C.-S. Horng (2010), Detection of noninteracting single domain particles using first-order reversal curve diagrams, *Geochem. Geophys. Geosyst.*, *11*, Q01Z11, doi:10.1029/2009GC002916.
- Harrison, R. J., and J. M. Feinberg (2008), FORCinel: An improved algorithm for calculating first-order reversal curve distributions using locally weighted regression smoothing, *Geochem. Geophys. Geosyst.*, *9*, Q05016, doi:10.1029/2008GC001987.
- Heslop, D., and A. R. Muxworthy (2005), Aspects of calculating first-order reversal curve distributions, *J. Magn. Magn. Mater.*, *288*, 155–167, doi:10.1016/j.jmmm.2004.09.002.
- Jackson, M., and P. Solheid (2010), On the quantitative analysis and evaluation of magnetic hysteresis data, *Geochem. Geophys. Geosyst.*, *11*, Q04Z15, doi:10.1029/2009GC002932.
- Krásá, D., C. D. W. Wilkinson, N. Gadegaard, X. Kong, H. Zhou, A. P. Roberts, A. R. Muxworthy, and W. Williams (2009), Nanofabrication of two-dimensional arrays of magnetite particles for fundamental rock magnetic studies, *J. Geophys. Res.*, *114*, B02104, doi:10.1029/2008JB006017.
- Krásá, D., A. R. Muxworthy, and W. Williams (2011), Room- and low-temperature magnetic properties of 2-D magnetite particle arrays, *Geophys. J. Int.*, *185*, 167–180, doi:10.1111/j.1365-246X.2011.04956.x.
- Long, J. S., and L. H. Ervin (2000), Using heteroscedasticity consistent standard errors in the linear regression model, *Am. Stat.*, *54*, 217–224.
- MacKinnon, J. G., and H. White (1985), Some heteroscedasticity-consistent covariance matrix estimators with improved finite sample properties, *J. Econometr.*, *29*, 305–325.
- Mayergoyz, I. D. (1986), Mathematical models of hysteresis, *Phys. Rev. Lett.*, *56*, 1518–1561.
- Miller, R. G., Jr. (1981), *Simultaneous Statistical Inference*, Springer, New York.
- Muxworthy, A. R., and D. J. Dunlop (2002), First-order reversal curve (FORC) diagrams for pseudo-single-domain magnetites at high temperature, *Earth Planet. Sci. Lett.*, *203*, 369–382.
- Muxworthy, A. R., and A. P. Roberts (2007), First-order reversal curve (FORC) diagrams, in *Encyclopedia of Geomagnetism and Paleomagnetism*, pp. 266–272, Springer, Dordrecht, Netherlands.
- Muxworthy, A. R., D. Heslop, and W. Williams (2004), Influence of magnetostatic interactions on first-order reversal curve (FORC) diagrams: A micromagnetic approach, *Geophys. J. Int.*, *158*, 888–897.
- Muxworthy, A. R., J. G. King, and D. Heslop (2005), Assessing the ability of first-order reversal curve (FORC) diagrams to unravel complex magnetic signals, *J. Geophys. Res.*, *110*, B01105, doi: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, *Geochem. Geophys. Geosyst.*, *6*, Q05010, doi:10.1029/2004GC000877.
- Pan, Y., N. Petersen, M. Winklhofer, A. F. Davila, Q. Liu, T. Frederichs, M. Hanzlik, and R. Zhu (2005), Rock magnetic properties of uncultured magnetotactic bacteria, *Earth Planet. Sci. Lett.*, *237*, 311–325, doi:10.1016/j.epsl.2005.06.029.
- Pike, C. R. (2003), First-order reversal-curve diagrams and reversible magnetization, *Phys. Rev. B*, *68*, 104424, doi:10.1103/PhysRevB.68.104424.
- Pike, C. R., A. P. Roberts, and K. L. Verosub (1999), Characterizing interactions in fine magnetic particle systems using first order reversal curves, *J. Appl. Phys.*, *85*, 6660–6667, doi:10.1063/1.370176.
- Pike, C. R., A. P. Roberts, and K. L. Verosub (2001a), First-order reversal curve diagrams and thermal relaxation effects in magnetic particles, *Geophys. J. Int.*, *145*, 721–730, doi:10.1046/j.0956-540x.2001.01419.x.
- Pike, C. R., A. P. Roberts, M. J. Dekkers, and K. L. Verosub (2001b), An investigation of multi-domain hysteresis mechanisms using FORC diagrams, *Phys. Earth Planet. Inter.*, *126*, 11–25, doi:10.1016/S0031-9201(01)00241-2.
- Preisach, F. (1935), Über die magnetische Nachwirkung, *Z. Phys.*, *94*, 277–302.
- Roberts, A. P., C. R. Pike, and K. L. Verosub (2000), First-order reversal curve diagrams: A new tool for characterizing the magnetic properties of natural samples, *J. Geophys. Res.*, *105*, 28,461–28,475, doi:10.1029/2000JB900326.
- Roberts, A. P., W. T. Jiang, F. Florindo, C. S. Horng, and C. Laj (2005), Assessing the timing of greigite formation and the reliability of the Upper Olduvai polarity transition record from the Crostolo River, Italy, *Geophys. Res. Lett.*, *32*, L05307, doi:10.1029/2004GL022137.
- Roberts, A. P., Q. Liu, C. J. Rowan, L. Chang, C. Carvallo, J. Torrent, and C.-S. Horng (2006), Characterization of hematite (α -Fe₂O₃), goethite (α -FeOOH), greigite (Fe₃S₄), and pyrrhotite (Fe₇S₈) using first-order reversal curve diagrams, *J. Geophys. Res.*, *111*, B12S35, doi:10.1029/2006JB004715.
- Roberts, A. P., F. Florindo, G. Villa, L. Chang, L. Jovane, S. M. Bohaty, J. C. Larrasoána, D. Heslop, and J. D. Fitz Gerald (2011), Magnetotactic bacterial abundance in pelagic marine environments is limited by organic carbon flux and availability of dissolved iron, *Earth Planet. Sci. Lett.*, *310*, 441–452.
- Spooner, M. I., P. De Deckker, T. T. Barrows, and L. K. Fifield (2011), The behaviour of the Leeuwin Current offshore NW Australia during the last five glacial–interglacial cycles, *Global Planet. Change*, *75*, 119–132.
- Tric, E., C. Laj, C. Jehanno, J.-P. Valet, C. Kissel, A. Mazaud, and S. Iaccarino (1991), High-resolution record of the Upper Olduvai transition from Po Valley (Italy) sediments: Support for dipolar transition geometry?, *Phys. Earth Planet. Inter.*, *65*, 319–336, doi:10.1016/0031-9201(91)90138-8.
- Winklhofer, M., and G. T. Zimanyi (2006), Extracting the intrinsic switching field distribution in perpendicular media: A comparative analysis, *J. Appl. Phys.*, *99*, 08E710, doi:10.1063/1.2176598.
- Winklhofer, M., R. K. Dumas, and K. Liu (2008), Identifying reversible and irreversible magnetization changes in prototype patterned media using first- and second-order reversal curves, *J. Appl. Phys.*, *103*, 07C518.
- Yamazaki, T. (2009), Environmental magnetism of Pleistocene sediments in the North Pacific and Ontong-Java Plateau: Temporal variations of detrital and biogenic components, *Geochem. Geophys. Geosyst.*, *10*, Q07Z04, doi:10.1029/2009GC002413.