An assessment of the Levenberg-Marquardt fitting algorithm on saturating exponential data sets.

Robert B. Hayes, Edwin H. Haskell and Gerry H. Kenner

University of Utah ,Radiobiology Div.. Center for Applied Dosimetry
729 Arapeen Dr. SLC, UT 84108-1218, U.S.A
email; r.b.hayes@m.cc.utah.edu . Phone; 801-585-4018

(Received 17 June 1998 ; in final form 20 October 1998)

Introduction

A common, flexible curve fitting algorithm has recently become available in a commercial graphical software package for both PC's and Mac's (Kaleidagraph©). The source code of the algorithm is also available from standard numerical recipes packages in FORTRAN, BASIC, Pascal and C (©Press et al. 1992). We assess the algorithm's ability to model saturating dose responses by comparing it to other established methods using the test data of Berger and Huntley (1989) for two intersecting and saturating curves constructed from the partial bleach method. The utility of the package for other applications is also discussed.

The Levenberg-Marquardt (LM) algorithm has some unique properties relative to other, more established algorithms devoted to modeling saturating exponential data sets (Berger et al. 1987, Brumby 1992, Grün and Macdonald 1989, Poljakov and Hütt 1990). It has been validated as an excellent tool for deconvoluting EPR spectra (Haskell et al. 1996A&B, Haskell et al. 1997A&B, Kenner et al. 1998, Jonas 1995 & 1997, Pilawa et al. 1995, Polyakov et al. 1995) and has also been used for, or in conjunction with, TL spectral deconvolution (Lucas and Akselrod 1996, Emfietzoglou and Moscovitch 1996). The algorithm can also fit standard OSL curves containing multiple centers scanned using either the conventional shine down method or the new linear modulation technique (Bulur 1996, Bulur and Göksu 1997, Figure 1). This allows the same software to be used for multiple applications, from first determining the measurement amplitude (via deconvolution) to fitting the resultant dose response.

The form of this dose response can vary from a linear response to multiple saturating exponentials (and virtually all other forms including sublinear, supralinear etc.) while still being modeled appropriately with the algorithm (assuming that the correct model is used). The algorithm also provides uncertainty estimates for all parameters of the fit. A tangential application where this becomes useful is in estimating the saturation dose for the material being studied. This is required to determine when a linear fit rather than a saturating exponential fit should be applied to the data (i.e., the maximum dose applied should be $\leq$ 1% of the saturation dose for linear fitting to be statistically more accurate, see Grün 1996). However, this criterion should only be applied if accurate values with reliable uncertainties can be assured. The importance of using the correct model, errors and fitting procedure is clearly laid out in the literature (Berger et al. 1987, Berger 1990, Grün 1996, Grün and Brumby 1994, Grün and Rhodes 1992, Grün and Packman 1993, Guibert et al. 1996, Lyons et al. 1992).

The LM method for nonlinear fitting is a numerical compromise between the Gauss-Newton method of linearization (used by Berger et al. 1987) and that of steepest descent. The LM method uses a least squares criterion for convergence and can therefore be expected to give results similar to the method of Poljakov and Hütt (1990) (which iteratively uses a strict least squares method to determine two parameters and a Newton-Raphson method to determine the third) for unweighted saturating exponential data. In generating the curve parameters, the LM algorithm also generates the covariance matrix for the model and data (see Press et al. 1992). It is from the diagonals of the covariance matrix that the individual variances of the fitted parameters are taken (Press et al. 1992). The LM method does require that the initial estimates be sufficiently near (typically within an order of magnitude, Motulsky 1997) the optimum values to attain a reliable convergence.

Aliquot curve parameter	GWB	S C	S-m	DOSE	LM
QNL84-2 unbleached I_O (x 10^{-4})	14.28	14.25 ± 0.46	14.25	-	14.25 ± 0.56
QNL84-2 unbleached D_X	122.7 ± 6.7	121.9 ± 6.7	121.9	121.9 ± 6.8	122.0 ± 8.3
QNL84-2 unbleached D_O	392	390 ± 31	389.9	-	390 ± 38
QNL84-2 bleached I_O (x 10^{-4})	9.64	9.7 ± 1.0	9.67	-	9.67 ± 0.86
QNL84-2 bleached D_X	193 ± 19	195 ± 20	195.2	195 ± 20	195 ± 17
QNL84-2 bleached D_O	762	770 ± 150	773	-	770 ± 130
STRB87-1 unbleached I_O (x 10^{-4})	21.21	21.15 ± 0.48	21.15	-	21.18 ± 0.45
STRB87-1 unbleached D_X	0.583 ± 0.018	0.583 ± 0.018	0.583	0.591 ± .020	0.591 ± 0.017
STRB87-1 unbleached D_O	5.96	5.95 ± 0.25	5.95	-	5.98 ± 0.23
STRB87-1 bleached I_O (x 10^{-4})	12.043	12.03 ± 0.32	12.03	-	12.03 ± 0.32
STRB87-1 bleached D_X	0.680 ± 0.023	0.682 ± 0.023	0.682	0.682 ± 0.023	0.682 ± 0.023
STRB87-1 bleached D_O	6.67	6.68 ± 0.31	6.68	-	6.68 ± 0.31
STRB87-1 intersection	0.485 ± 0.037	0.481 ± 0.080	-	-	0.481 ± 0.031
QNL84-2 intersection	86 ± 10	85 ± 21	-	-	85 ± 28

Table 1.
Curve parameters from modeling saturating exponential data. Corresponding error estimates (when available) are also given. Here I_O is the saturation intensity, D_X is the dose estimate and D_O is the characteristic saturation dose. The GWB method is the quasi-likelihood method of Berger et al. (1987), SC is the simplex algorithm (Berger and Huntley 1989), S-m is a weighted least squares method (Berger and Huntley 1989), DOSE is the simplex algorithm with quadratic convergence from Brumby (1992) and LM is the Levenberg-Marquardt method. The error estimators of the two curve intersections for the LM method were calculated using basic first order error propagation.

Like the methods of Brumby (1992), Grün and Macdonald (1989) or Poljakov and Hütt (1990), the LM fitting method does not have the ability to rigorously assess the relative error distribution of the data as does that of Berger et al. (1987). These errors must be independently determined if LM fitting is to be used (a very basic and simple method for estimating these uncertainties is described below). This is important because Grün and Rhodes (1992) verified the earlier derivation of Berger et al. (1987, appendix A) that curve fits to saturating exponential dosimetry data should be weighted by relative terms. It should be pointed out, however, that if a case were found where relative errors could be demonstrated to be negligible, then this would no longer be expected to be the case and equal weighting would be expected to give better results (Franklin 1986). The most common instance in which this will occur is when the high frequency noise from the measurement instrumentation (or other source) becomes large relative to the size of the dosimetric signals being evaluated. Examples fulfilling this criterion, and more precise descriptions of the effect, are given both by Hayes et al. (1997) and Scott and Sanderson (1988). In general, this is not expected to be observed in geological dating studies but should be restricted to more recent archeological dose reconstructions (Grün and Macdonald 1989). When the case of mixed errors is encountered (where neither the constant or relative errors are negligible), the appropriate weighting terms are given by Bluszcz (1988).

Using the saturating exponential test data of Berger and Huntley (1989), with the modified data point mentioned by Huntley (1996), we compare the results of the LM routine to the quasi-likelihood method of Berger et al. (1987), the simplex method (Berger and Huntley 1989), a weighted least squares method (Berger and Huntley 1989) and the quadratic convergance method of Brumby (1992).

Results

The LM method appears to be comparable in both parameter and error estimation to the other fitting methods based on the results given in Table 1 (error estimates were not available for all the parameters of some of the other fitting methods). Here, we have included the calculations of the intersections and the first order error estimates of the uncertainties associated with the two intersections. The data were weighted with either 4% or 3% relative errors as done in Berger and Huntley (1989) so that a direct comparison could be made to their data. This was done because the calculated errors (not the fitted parameters) using the LM fitting method for data weighted by inverse variance are directly proportional to the relative error selected (which is true for other methods also). If the weights chosen were increased or decreased by a factor of ten, then the error estimates would also be increased or decreased by a factor of ten respectively despite the identical fitted parameter values.

Because of the necessity to weight saturating data by relative errors, the magnitude of the relative error must be reasonably estimated and supplied to the LM algorithm prior to the fit. We offer a simple, model-dependent method. This is to initially fit the data with a weighted saturating exponential (here the weights are relative errors of arbitrary magnitude) to first estimate the residuals (the residual for each data point is the vertical distance from that point to the curve fit). We then divide each residual by the magnitude of the curve fit at the respective point. After taking the average of these numbers, we can equate this to the relative error magnitude which we should use in the LM method for a weighted fit to the data. This approach assumes both a correct model was employed in the weighted fit (e.g., single saturating exponential) and that the residuals are Gaussian distributed with variance dominated by relative values (constant errors, systematic errors etc. are assumed negligible). Because the fitting parameters will not be affected by a change in the magnitude of the weighting terms of the relative error (only their error estimators will be), iteration is not necessary for this approach. Application of this method resulted in average relative error amounts of 2.2% for the STRB87-1 measurements and 2.9% for the QNL84-2 measurements.

It should be pointed out that this simple approximation method is not offered as an optimal approach but rather as an easy to use approximation. It will not be acceptable in all cases. Indeed the main advantages of other established methods (Berger et al. 1987) over that of LM is their implicit ability to determine the respective weights for use in fitting each data set and to iteratively assess the uncertainty in the intersection of the two curve fits.

Some comments addressing convergence reliability and accuracy are now in order. Our comments are based on 4 years experience of modeling dose response data using the LM algorithm. When the maximum dose is less than about 10 percent of the saturation dose (D_O,) the estimates and associated uncertainties for both D_O and the saturation intensity I_O will converge to arbitrarily large values (with the uncertainties being typically at least a few orders of magnitude larger than their associated parameters). In the same situation however, the x-intercept of the fit with its estimated uncertainty is similar to that obtained using a linear fit. The algorithm otherwise will converge (with reasonable parameter and uncertainty estimates) for all data sets even when the relative errors are approaching 30% (which indicates the data set itself is not useful).

Example

A straightforward application for LM deconvolution is shown in Figure 1 which was done entirely using the software package Kaliedagraph©. Here we show a spectrum from the dose response of a K-feldspar sample. The sample was scanned using the new linear modulation method of OSL (Bulur 1996). The resultant dose response from the three signals used to model the spectrum are shown in Figure 2. The data for Figures 1 and 2 were supplied by Enver Bulur.

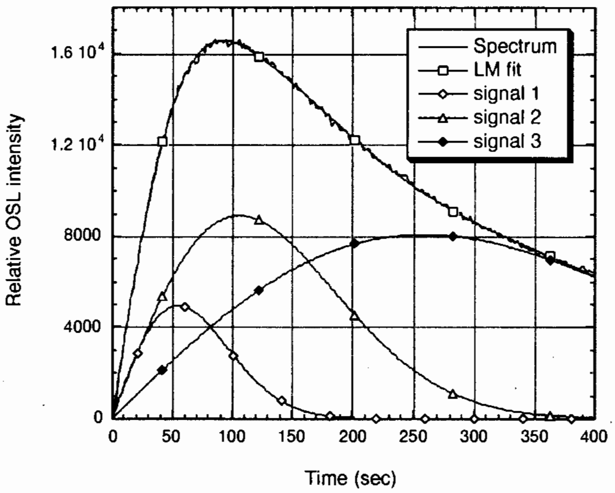

Figure 1.
Spectral deconvolution of an OSL curve recorded using the linear modulation technique (Bulur 1996). The curve fits used an unweighted model. This spectrum was chosen for the figure because it had the lowest correlation (it also had the lowest dose of 2.825 Gy) . The form of the curve used for the model assumed first order kinetics resulting in curve shapes of the form

$$Y=-m1*X*exp(-X^2/(2*m2^2))/m2^2.$$

This allowed us to consider only the dependence on the m1 parameter for the dose response although the m2 parameter was consistently stable to within 2% for each signal.

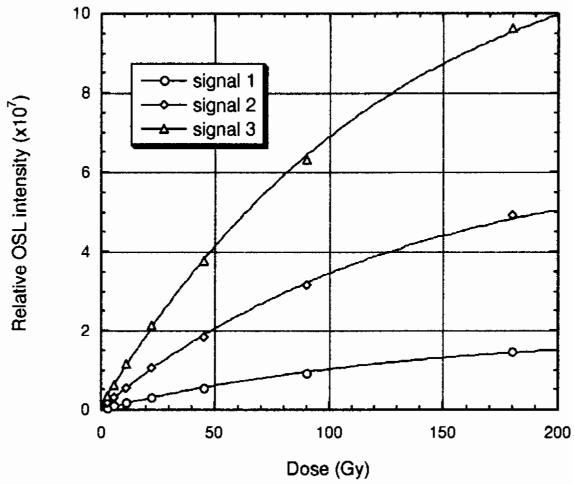

Figure 2.
Curve fit to the saturating exponential data set generated by the linear modulation OSL technique. The curve fit for each dose response from the three deconvoluted signals shown in Figure 1 are displayed. By plotting the m1 parameter (see caption to Figure 1) we effectively plot the integrated spectrum of each signal. Using the method described in the text for the LM algorithm, the fits were weighted with 2% relative errors. The applied doses were 2.825, 5.625, 11.25, 22.5, 45, 90 and 180 Gy. This resulted in reconstructed doses D_X of 420±95 mGy, 424±28 mGy and 686±106 mGy for signals 1, 2 and 3 respectively.

Conclusion

Because it is available as either a canned numerical routine (Press et al. 1992) or an integrated component of a commercial graphing software package (Kaleidagraph©), the LM fitting routine is readily available with little required effort of the user. Reasonable estimates may now be obtained in a user friendly graphing program using a viable algorithm for modeling many of the data sets encountered in dating and retrospective dosimetry.

It should be born in mind however that for the specific application considered here (determination of the intersection of two intersecting saturating exponentials), more accurate and precise algorithms can be found in the literature (Berger et al. 1987). This is due to both the lack of the LM algorithm to do any implicit and rigorous treatment of the dual curve intersection error analysis and also its inability to implicitly determine the error distribution of the data set it is modeling.

The method has some potential benefits not offered by certain other routines in use. In particular, it has the necessary flexibility to model saturating-exponential-plus-linear data (Hütt and Smirnov 1983, Berger 1990, Berger 1991), multiple saturating exponentials (Katzenberger and Willems 1988, Packman and Grün 1992), data with supralinearities (Valladas and Gillot 1978) or any number of other models used in TL/EPR dose responses (Berger 1985, Berger 1987, Guibert et al. 1996). The minimum requirements for using the algorithm is that both the fitting function and its gradient can be calculated (if these are not given in closed form then numerical calculation can be implemented at the cost of convergence speed).

Acknowledgments
We would like to thank Dr. Enver Bulur for graciously providing us with data he recorded using the linear modulation OSL technique. This work supported by the U.S. Department of Energy, Contract DE-FC03-97SF1354.

References

Berger G. W. (1985) Thermoluminescence dating applied to a thin winter varve of the Late Glacial Thompson Silt, south-central British Columbia. *Can. J. Earth Sci.*, **22**, 1736-1739.

Berger G. W. (1987) Thermoluminescence dating of the Pliestocene Old Crow tephra and adjacent loess, near Fairbanks, Alaska. *Can. J. Earth Sci.*, **24**, 1975-1984.

Berger G. W. (1990) Regression and error analysis for a saturating-exponential-plus-linear model. *Ancient TL*, **8**, 23-25.

Berger G. W. (1991) The use of glass for dating volcanic ash by thermoluminescence. *J. Geophys. Res.*, **96**, 19705-19720.

Berger G. W. and Huntley D. J. (1989) Test data for exponential fits. *Ancient TL*, **7**, 43-46.

Berger G. W., Lockhart R. A. and Kuo J. (1987) Regression and error analysis applied to the dose-response curves in thermoluminescence dating. *Nucl. Tracks Radiat. Meas.*, **13**, 177-184.

Bluszcz A (1988) The Monte-Carlo experiment with the least squares methods of line fitting. *Nucl. Tracks Radiat. Meas.* **14**, 355-360.

Brumby S. (1992) Regression analysis of ESR/TL dose-response data. *Nucl. Tracks Radiat. Meas.*, **20**, 595-599.

Bulur E. (1996) An alternative technique for optically stimulated luminescence (OSL) experiment *Radiat. Meas.* **26**, 701- 709.

Bulur E. and Göksu H. Y. (1997) IR stimulated luminescence from ZnS and SrS based storage phosphors: a re-examination with linear modulation technique. *Phys. Stat. Sol. (Rapid Research Notes)* **161**, R9-R10.

Emfietzoglou D. and Moscovitch M.. (1996) Phenomenological study of light-induced effects in a-Al_2O_3:C. *Radiat. Prot. Dos.*, **65**, 259-262.

Grün R. (1996) Errors in dose assessment introduced by the use of the "linear part" of a saturating dose response curve. *Radiat. Meas.*, **26**, 297-302.

Grün R. and Brumby S. (1994) The assessment of errors in past radiation doses extrapolated from ESR/TL dose-response data. *Radiat. Meas.*, **23**, 307-315.

Grün R. and Macdonald P. D. M. (1989) Non-linear fitting of TL/ESR dose-response curves. *Appl. Radiat. Isot.*, **40**, 1077-1080.

Grün R. and Packman S. C. (1993) Uncertainties involved in the measurement of TL intensities. *Ancient TL*, **11**, 14-20.

Grün R. and Rhodes E. J. (1992) Simulation of saturating exponential ESR/TL dose response curves - weighting of intensity values by inverse variance. *Ancient TL*, **10**, 50-56.

Guibert P., Vartanian E., Bechtel F. and Schvoerer M. (1996) Non linear approach of TL response to dose: polynomial approximation. *Ancient TL*, **14**, 7-14.

Haskell E., Kenner G., Hayes R., Chumak V. and Sholom S. (1996A) EPR dosimetry of teeth in past and future accidents: a prospective look at a retrospective method. *Effects of Low-Level Radiation for Residents Near Semipalatinsk Nuclear Test Site. Proceedings of the Second Hiroshima International Symposium.* M. Hoshi, J. Takada, R. Kim and Y. Nitta (Ed.). Research Institute for Radiation Biology and Medicine, Hiroshima University (Pub.). Hiroshima, Japan. 261-274.

Haskell E. H., Hayes R. B. and Kenner G. H. (1996B) Plasterboard as an emergency EPR dosimeter. *Health Phys.* **71**, 95.

Haskell E. H., Hayes R. B., Kenner G. H., Sholom S. V. and Chumak V. V. (1997A) EPR techniques and space biodosimetry. *Radiat. Res.* **148**, S51-S59.

Haskell E. H., Hayes R. B. and Kenner G. H. (1997B) Improved accuracy in EPR dosimetry using a constant rotation goniometer. *Radiat. Meas.* **27**, 325-329.

Hayes R. B., Haskell E. H. and Kenner G. H. (1997) A mathematical approach to optimal selection of dose values in the additive dose method of EPR dosimetry. *Radiat. Meas.*, **27**, 315-323.

Huntley D. J. (1996) Errata (Letters). *Ancient TL*, **14**, 31.

Hütt G. and Smirnov A. (1983) Thermoluminescence dating of sediments by means of the quartz and feldspar inclusion methods. *PACT*, **9**, 463-471.

Jonas M. (1995) Spectral deconvolution of the ESR dating signal in fossil tooth enamel. *Quat. Sci. Rev.*, **14**, 431-438.

Jonas M. (1997) *Electron Spin Resonance Dating of Tooth Enamel.* Ph.D. thesis, Cambridge University, Cambridge UK.

Katzenberger O. and Willems N. (1988) Interferences encountered in the determination of AD of mollusc samples. *Quat. Sci. Rev.*, **7**, 485-489.

Kenner G. H., Haskell E. H., Hayes R. B., Baig A. and Higuchi W. I., (1998) EPR properties of synthetic apatites, deorganified dentin and enamel. *Calcif. Tiss. Int.*, **62**, 443-446.

Lucas A. C. and Akselrod M. S. (1996) The determination of dose as a function of depth by deconvolution of the glow curve using thick TL dosimeters. *Radiat. Prot. Dosim.*, **66**, 57-62.

Lyons R. G., Brennan B. J. and Hosking P. L. (1992) Estimation of accumulated dose and its uncertainties: potential pitfalls in curve fitting. *Ancient TL*, **10**, 42-49.

Motulsky H. (1997) Fitting curves with nonlinear regression. *Scitech J.*, **7**, 20-23.

Pilawa B., Weickowski A. B. and Trzebicka B. (1995) Numerical analysis of EPR spectra of coal, trace macerals and extraction products. *Radiat. Phys. Chem.*, **45**, 899-908.

Packman S. C. and Grün R. (1992) TL analysis of loess samples from Achenheim. *Quat. Sci. Rev.*, **11**, 103-107.

Polyakov V., Haskell E., Kenner G., Huett G. and Hayes R. (1995) Effect of mechanically induced background signal on EPR dosimetry of tooth enamel. *Radiat. Meas.* **24**, 249-254.

Poljakov V. and Hütt G., (1990) Regression analysis of exponential palaeodose growth curves. *Ancient TL*, **8**, 1-2.

Press W. H., Teukolsky S. A., Vetterling W. T. and Flannery B. P. (1992) *Numerical Recipes in FORTRAN - The art of Scientific Computing. 2nd edn.* Cambridge University Press, New York. pp. 675-694.

Scott E. M. and Sanderson D. C. W. (1988) Statistics and the additive dose method in TL dating. *Nucl. Tracks. Radiat. Meas.*, **14**, 345-354.

Valladas G. and Gillot P. Y. (1978) Dating of the Olby lava flow using heated quartz pebbles. *PACT*, **2**, 141-150.

Reviewer
G. Berger

Comments

The authors deserve thanks for enduring my suggested iterative revisions. It is worth mentioning a few points here. All users of statistical packages should be alert to the assumptions and limitations of the methods used . The authors carefully point out that the LM method can be weak in error analysis (second last paragraph in Results section). Its strength appears to be in versatility and ease of use. Note that the « saturating exponential » model of Berger et al. (1987) can also be used successfully with most sublinear data, sometimes with supralinear data. However, as discussed by Berger et al. (1987), second-order polynomials (quadratics) may work better (as good approximation) than their exponential-model method for supralinear data, certainly for two intersecting curves. As well, the methods of Berger et al. (1987) and Berger (1990) probably may also be used to obtain estimates of uncertainties in certain fitting parameters (e.g. D_O and I_O), but have not yet bothered to extract this information from the covariance matrices. Also, it appears that the authors' approximate method for estimating relative errors in advance of solution by LM is equivalent to the use of equation 4 in Berger et al. (1987), except that the latter includes a weighting term in this equation (making it iterative).

Finally, for new workers in TL/OSL/ESR it is worth mentioning that the basic arguments for the use of weighting by inverse variance have deep historical roots. Several of these roots are cited in Berger et al. (1987). Among these are citations of Deming (1949), and the papers of York. I urge readers to consult York's papers on regression and error analysis from time to time. Although his papers deal with linear models, some of the concepts are still relevant to sublinear models.