Originally published as:

Stromeyer, D., Grünthal, G., Wahlström, R. (2004): Chi-square regression for seismic strength parameter relations, and their uncertainties, with applications to an M_w based earthquake catalogue for central, northern and northwestern Europe. - Journal of Seismology, 8, 1, 143-153

DOI: 10.1023/B:JOSE.00000009503.80673.51

Chi-square regression for seismic strength parameter relations, and their uncertainties, with applications to an M_w based earthquake catalogue for central, northern and northwestern Europe

Dietrich Stromeier¹, Gottfried Grünthal² & Rutger Wahlström³

GeoForschungsZentrum Potsdam, Department 5, Germany;

¹ e-mail: stro@gfz-potsdam.de; ² e-mail: ggrue@gfz-potsdam.de; ³ e-mail: rutger@gfz-potsdam.de

Received 23 January 2003; accepted in revised form 10 June 2003

Key words: chi-square regression, earthquake catalogue, intensity, magnitude, seismic hazard

Abstract

A generalized chi-square regression approach to establish empirical relations between different types of seismic strength parameters with uncertainties in all input data is presented and discussed in comparison with standard least-squares techniques. The chi-square technique can consider errors of individual entries but can also be applied when errors are not exactly known and only weaker quantitative constraints can be made. It can preserve the symmetry of the derived relations and is preferred for complex regression models. Results for three types of regression models are presented for (1) a linear relation between M_S and m_b for events in the North Atlantic Ocean; (2) a quadratic relation between M_w and M_L for events in central Europe; (3) linear relations between M_L and I_0 , with logarithmic dependency of the focal depth, for several regions in central and northern Europe.

Introduction

The quality of hazard or risk estimations of earthquake threatened regions is closely linked to the quality of the earthquake catalogues which are the base for these investigations. It is a goal to have catalogues which encompass as long a time period as possible and simultaneously contain only reliable earthquake data with specified statistical errors. A key measure in this connection is the strength of the quake. The practice of instrumental records of seismic events started at the end of the 19th century; from historical preinstrumental earthquakes we have only estimations of the intensities and other macroseismic data. The instrumentally based measure to characterize earthquake strength is the magnitude. The moment magnitude, M_w , was introduced more than 20 years ago and was preceded by and is in parallel use with a number of different magnitude concepts, such as local magnitude, M_L , body wave magnitude, m_b , and surface wave magnitude, M_S . The interconnections

between the magnitude scales and between the respective magnitude and the intensity are rather loose and the transformations usually follow empirical relations derived from the local conditions in each case. The task of this study is the use of a common chi-square regression model to derive a hierarchy of empirical relations between different magnitudes and between M_L magnitude and epicentral intensity I_0 . The present approach is particularly considering that all model parameters are affected by errors. From conditions of the statistics of errors of the input parameters, uncertainty limits for the calculated regression parameters and a criterion for the goodness-of-fit of the performed regressions can be defined. The advantages of the selected regression approach in comparison with other approaches (linear, multivariate and orthogonal regression, respectively) are discussed for three different applications: a linear relation $M_S = a_1 m_b + b_1$ for data in the North Atlantic Ocean, a quadratic relation $M_w = a_2 M_L^2 + b_2 M_L + c_2$ for

Central Europe and a relation $M_L = a_3 I_0 + b_3 \log(h) + c_3$ for data in Germany and Fenno-scandia, respectively. h in the logarithmic term denotes the focal depth and a_i , b_i and c_i are constants. The equations are important parts of the generation of an M_w based earthquake catalogue for central, northern and northwestern Europe (Grünthal and Wahlström, 2003).

General chi-square regression model

To establish an empirical relationship between different magnitudes, magnitudes and intensity or other seismological measurements, the observed data are summarized by fitting to a model that depends on adjustable parameters. Sometimes the model is only a simple class of functions. Sometimes it comes from an underlying theory that the data are supposed to satisfy. The basic technique to estimate the model parameters is usually the familiar least-squares approach

$$\min_{\mathbf{a}} \sum_{i=1}^N [y_i - f(\mathbf{x}_i; \mathbf{a})]^2 \quad (1)$$

where $y = f(\mathbf{x}; \mathbf{a})$ represents a relation between L independent observables $\mathbf{x} = (x_1, \dots, x_L)$, M adjustable parameters $\mathbf{a} = (a_1, \dots, a_M)$ and one dependent (measured) observable y . The squared difference between measurements and model is accumulated for a set of N data points $(\mathbf{x}_i, y_i)$, $i = 1, \dots, N$. It is well known (Mises von, 1964) that least-squares fitting by eq. (1) is a maximum likelihood estimation of $\mathbf{a}$ if the measurement errors of y are independent and normally distributed with constant variance σ^2

$$\text{var}(y_1) = \text{var}(y_2) = \dots = \text{var}(y_N) = \sigma^2$$

If each observable y_i has its own variance σ_i^2 , eq. (1) can be modified to a sum of N square quantities, each normalized to unit variance known as the ‘chi-square’ (Bevington, 1969)

$$\chi^2(\mathbf{a}) = \sum_{i=1}^N \frac{[y_i - f(\mathbf{x}_i; \mathbf{a})]^2}{\sigma_i^2} \quad (2)$$

For the problem of local magnitude regression, the

empirical data are subject to measurement errors not only in y_i , but also in $\mathbf{x}_i$. Assuming that the differences, $y_i - f(\mathbf{x}_i; \mathbf{a})$, between measurements and model are uncorrelated and normally distributed, the appropriate chi-square merit function to be minimized becomes

$$\chi^2(\mathbf{a}) = \sum_{i=1}^N \frac{[y_i - f(\mathbf{x}_i; \mathbf{a})]^2}{\text{var}[y_i - f(\mathbf{x}_i; \mathbf{a})]} \quad (3)$$

where the variances, $\text{var}[y_i - f(\mathbf{x}_i; \mathbf{a})]$, depend on the unknown parameter vector $\mathbf{a}$. In the neighbourhood of the optimal vector, $\hat{\mathbf{a}}$, where $\chi^2(\mathbf{a})$ takes its minimum, there is a quantitative measure Q for the goodness-of-fit associated with eq. (3). It is defined as the probability that the observed chi-square will exceed the value χ^2 by chance even for a correct model

$$Q(\nu, \chi^2) = 1 - P\left(\frac{\nu}{2}, \frac{\chi^2}{2}\right) \quad (4)$$

where $\nu = N - M$ is the degree of freedom of the chi-square probability function and P is the incomplete gamma function

$$P(p, z) = \frac{1}{(p-1)!} \int_0^z e^{-t} t^{p-1} dt$$

There is no simple way to give a threshold value Q_0 which should be exceeded by Q for a good fit, in particular if the differences, $y_i - f(\mathbf{x}_i; \mathbf{a})$, are not normally distributed where a parameter fit by minimizing the chi-square is often a useful means. A rule of thumb is that an acceptable value of χ^2 for a good fit is $\chi^2 \leq \nu$ (Press et al. (1992), §15) which leads to $Q_0 = Q(\nu, \nu)$.

The estimation of quantitative confidence limits for fitted parameters in the chi-square approach is straight-forward. If $\hat{\mathbf{a}}$ is the optimal parameter vector for a given chi-square problem defined by eq. (3), then

$$\Delta\chi^2 = \chi^2(\mathbf{a}) - \chi^2(\hat{\mathbf{a}}) \quad (5)$$

is a chi-square probability distribution with M degrees of freedom. The contours $\Delta\chi^2 = \text{const}$ are a convenient measure for the probability distribution of $\mathbf{a}$. For different probabilities p , the appropriate

Table 1. χ^2 as a function of confidence level and degree of freedom

confidence level p	degree of freedom μ		
	1	2	3
68.3%	1	2.30	3.53
90.0%	2.71	4.61	6.25
95.4%	4	6.17	8.02
99.0%	6.63	9.21	11.3

contour values are the roots of $Q(M, \Delta\chi^2) = 1 - p$. In practice, one is frequently interested not in the full M -dimensional confidence region, but in individual confidence limits for each single parameter a_i , $i = 1 \dots M$. Then the associated $\Delta\chi^2$ values of a confidence level p have to be calculated from $Q(1, \Delta\chi^2) = 1 - p$. Table 1 shows the roots $\Delta\chi^2$ of $Q(\mu, \Delta\chi^2) = 1 - p$ as a function of confidence level p and degree of freedom $\mu = 1, \dots, M$ (see, e.g. Press et al., 1992). For individual standard errors, $p = 68.3\%$, $\Delta\chi^2 = 1$ holds and limits of the model parameters are the projections of this closed region onto a_i .

From a Taylor series expansion of eq. (5)

$$\Delta\chi^2 \approx (\hat{\mathbf{a}} - \mathbf{a})^T \mathbf{C} (\hat{\mathbf{a}} - \mathbf{a}) = 1$$

where $\mathbf{C}$ is the matrix of second derivatives

$$c_{kl} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial x_k \partial x_l}, \quad k, l = 1 \dots N, \quad \text{at } \mathbf{a} = \hat{\mathbf{a}}$$

$\sigma(\mathbf{a})$, the standard errors of the model parameters, are obtained from

$$\sigma(a_i) = \sqrt{(\mathbf{C}^{-1})_{ii}}, \quad i = 1 \dots M. \quad (6)$$

While the estimation of $\mathbf{a}$ by minimizing the chi-squares of eq. (3) is a useful regression technique for a wide class of models, all statements about the goodness-of-fit and the resulting errors of $\mathbf{a}$ depend on the statistics of the input data. This approach assumes the normality of $y_i - f(\mathbf{x}_i; \mathbf{a})$. No explicit normal distribution of y_i and $\mathbf{x}_i$ is required, but the above assumption can only be proved after establishing $\mathbf{a}$. For normal distributed input data and models $f(\mathbf{x}_i; \mathbf{a})$, linear in $\mathbf{x}_i$, this is strongly related to the goodness-of-fit. Non-normal distributions of $y_i - f(\mathbf{x}_i; \mathbf{a})$ can in most cases be treated by time consuming Monte Carlo simulation for eq.

(3) in determining the expectancy and the variance of $\mathbf{a}$.

M_S vs. m_b

The regional relation between surface-wave magnitude, M_S , and body-wave magnitude, m_b , is usually expressed by a linear equation

$$M_S = a m_b + b \quad (7)$$

and its inverse

$$m_b = A M_S + B \quad (8)$$

where a and b (A and B) are estimated from a given data set of M_S and m_b . The determination of a and b (A and B) by a standard least-squares approach only considers errors of the dependent model variable M_S in eq. (7) and M_S in eq. (8), respectively. This implies that eq. (7) cannot be transformed into eq. (8) or reverse. Including independent input errors, $\sigma(M_S)$ and $\sigma(m_b)$, overcomes this ambiguity. From eq. (3), the appropriate chi-square function with respect to eq. (7) is

$$\chi^2(a, b) = \sum_{i=1}^N \frac{[M_{Si} - a m_{bi} - b]^2}{\sigma^2(M_{Si}) + a^2 \sigma^2(m_{bi})} \quad (9)$$

where the denominator is the approximation of $\text{var}[M_{Si} - a m_{bi} - b]$ by the error propagation law of Gauss (Press et al., 1992)

$$\begin{aligned} & \text{var}[M_{Si} - a m_{bi} - b] \\ & \approx \text{var}(M_{Si}) \left(\frac{\partial(M_{Si} - a m_{bi} - b)}{\partial M_{Si}} \right)^2 \\ & + \text{var}(M m_{bi}) \left(\frac{\partial(M_{Si} - a m_{bi} - b)}{\partial m_{bi}} \right)^2 \\ & = \sigma^2(M_{Si}) + a^2 \sigma^2(m_{bi}) \end{aligned}$$

assuming the independence of all combinations of the errors $\sigma(M_{Si})$ and $\sigma(m_{bi})$, $i, j = 1 \dots, N$. The corresponding expression with respect to eq. (8) is

$$\chi^2(A, B) = \sum_{i=1}^N \frac{[m_{bi} - A M_{Si} - B]^2}{\sigma^2(m_{bi}) + A^2 \sigma^2(M_{Si})} \quad (10)$$

Table 2. Chi-square and least-squares regression parameters, standard errors and mean fitting error σ of model (7) and model (8)

	$M_S = a m_b + b$				
	a	$\sigma(a)$	b	$\sigma(b)$	σ
chi-square	2.188	0.268	-6.264	1.334	0.256
least-squares	1.462	0.249	-2.661	1.442	0.524

	$m_b = A M_S + B$				
	A	$\sigma(A)$	B	$\sigma(B)$	σ
chi-square	0.457	0.056	2.863	0.261	0.256
least-squares	0.414	0.058	3.060	0.272	0.279

With $a = 1/A$ and $b = -B/A$, eq. (9) and eq. (10) can be converted into each other which shows the equivalence of model (7) and model (8) for this general approach. Standard least-squares solutions can be obtained from eq. (9) and eq. (10) by assuming $\sigma(m_b) = 0$ and $\sigma(M_S) = 0$, respectively. If the input data have an overall constant error $\sigma = \sigma(M_S) = \sigma(m_b)$, eq. (9) and eq. (10) define orthogonal regressions. While a standard least-squares fit with linear parameter dependency only requires the evaluation of an algebraic linear system, the minimization of the chi-squares of eq. (9) with respect to a and b is a nonlinear optimization problem. It can be solved by any appropriate iterative algorithm if the starting parameters are sufficiently close to the global minimum solution of eq. (9). For reasonable input data, a recommended starting point is the arithmetic mean of the standard least-squares solutions of model (7) and model (8). If the denominator of eq. (9) does not depend on index i , that is, $\sigma(M_{Si}) = \sigma(M_S)$ and $\sigma(m_{bi}) = \sigma(m_b)$, then the solution is connected to a linear eigenvalue problem (Appendix A) and can be found without iterations.

A set of $N = 42$ magnitude pairs for the North Atlantic Ocean will be used to demonstrate the chi-square regression procedure for the linear model (7). These are the available data for that region from the *Global Hypocenter Data Base* (1996). No detailed error information is available for this data set. Therefore, we suppose that all standard errors of M_S and m_b are equal to an unknown value $\sigma = \sigma(M_S) = \sigma(m_b)$. Then the chi-squares of eq. (9) are only scaled by a factor σ^2 , while the parameters a and b are not influenced and can be calculated for any assumed σ (e.g. $\sigma = 1$). As mentioned above, this reduces the chi-square ap-

proach to an ordinary orthogonal regression.

Statements about the goodness-of-fit and confidence limits of a and b require additional assumptions on the data. It seems reasonable to ask for a value σ for which the goodness-of-fit is greater than or equal to the threshold value $Q_0 = Q(\nu, \nu)$ (cf. previous chapter). This assumption fixes the value $\chi^2(a, b) = \nu = N - 2$ for the optimal parameters $a = \hat{a}$ and $b = \hat{b}$, and the unknown σ becomes

$$\sigma = \sigma(\hat{a}, \hat{b}) = \sqrt{\sum_{i=1}^N \frac{[M_{Si} - \hat{a} m_{bi} - \hat{b}]^2}{1 + \hat{a}^2}} / (N - 2) \quad (11)$$

Comparing this σ with an expected data error σ_{exp} for the studied region decides if the regression is well established, $\sigma \leq \sigma_{exp}$, or has to be rejected, $\sigma \gg \sigma_{exp}$. In terms of a standard least-squares ap-

Figure 1. Linear relation between M_S and m_b for the North Atlantic Ocean fitted by chi-square and least-squares techniques, respectively. The grey lines indicate the bilinear relation used by Grünthal and Wahlström (2003).

Figure 2. Mean fitting error $\sigma(a, b)$ with confidence regions of individual (dark gray) and joint (light gray) parameter errors for model (7).

proach, σ can be thought of as the mean fitting error.

From σ , defined by eq. (11), confidence limits for a and b can be estimated. The confidence ellipse for the parameters a and b jointly, $\mu = 2$, for a level of 68.3% corresponds to $\Delta\chi^2 = 2.30$, where the confidence interval for any individual parameter, $\sigma(a)$ and $\sigma(b)$, respectively, is connected with $\mu = 1$ and $\Delta\chi^2 = 1$ and is given by eq. (6).

Table 2 summarizes the solutions of the chi-square regressions for model (7) and model (8) in comparison with the standard least-squares results. The parameters of the chi-square approach are invertible for the two models, i.e., there is only one solution line (straight) in Figure 1, while the least-squares fits (long and short hatched lines) differ remarkably.

The distribution of $\chi^2(a, b)$ with respect to a and b in the neighbourhood of the optimal parameters $\hat{a}$ and $\hat{b}$ provides a quick overview of the regression errors. More clear is a representation of the mean fitting error $\sigma(a, b)$ in that neighbourhood shown in Figure 2. With our assumptions on constant input errors, $\sigma(a, b)$ can be expressed by

$$\sigma(a, b) = \sigma(\hat{a}, \hat{b}) \sqrt{\frac{\chi^2(\hat{a}, \hat{b}) + \Delta\chi^2(a, b)}{N - 2}} \quad (12)$$

where $\sigma(\hat{a}, \hat{b})$ is defined by eq. (11). For the confidence regions $\Delta\chi^2 = 1$ and $\Delta\chi^2 = 2.30$, the corresponding contour values are $\sigma(a, b) = 0.259$ and $\sigma(a, b) = 0.263$, respectively.

Obviously, the input data (Figure 1) do not reflect a uniform linear relation, with low magnitudes showing a different behaviour than high magnitudes. Therefore, Grünthal and Wahlström (2003) prefer a bilinear fit for this data set based on a modification of eq. (10) which takes into account that the point of intersection is initially unknown. The resulting relations (see Figure 1) are used to transform m_b magnitudes to M_w via M_S (the identity $M_w = M_S$ is assumed for the particular region) as a part of the generation of the M_w based catalogue by Grünthal and Wahlström (2003).

M_w vs. M_L

The relation between moment magnitude, M_w , and local magnitude, M_L , often cannot be approximated by a linear model. Several authors (e.g., Street et al., 1975; Hasegawa, 1983; Nuttli, 1983; Bollinger et al., 1993; Uhrhammer et al., 1996; Wahlström and Grünthal, 2000) prefer the quadratic approach

$$M_w = a M_L^2 + b M_L + c \quad (13)$$

which implies

$$M_L = A + \sqrt{B + C M_w} \quad (14)$$

The non-linearity has been ascribed to the intrinsic nature of the M_L concept. The corresponding merit functions with standard errors $\sigma(M_w)$ and $\sigma(M_L)$ are

$$\chi^2(a, b, c) = \sum_{i=1}^N \frac{[M_{wi} - a M_{Li}^2 - b M_{Li} - c]^2}{\sigma^2(M_{wi}) + (2a M_{Li} + b)^2 \sigma^2(M_{Li})} \quad (15)$$

and

$$\chi^2(A, B, C) = \sum_{i=1}^N \frac{[M_{Li} - A - \sqrt{B + C M_{wi}}]^2}{\sigma^2(M_{Li}) + C^2 / (4(B + C M_{wi})) \sigma^2(M_{wi})} \quad (16)$$

which can be reduced to the standard least-squares formulation supposing $\sigma(M_{Li}) = 0$ in eq. (15) or $\sigma^2(M_{wi}) = 0$ in eq. (16). It should be noted, that contrary to the linear case the chi-square approach

Figure 4. Quadratic relation between M_w and M_L for Central Europe fitted by chi-square and least-squares techniques, respectively.

Figure 4. Mean fitting error $\sigma(a, b, c)$ with confidence regions of individual (dark grey) and joint (light grey) parameter errors for model (13).

does not secure the equivalence of models (13) and (14). The different behaviour is due to the asymmetric appearance of M_w and M_L in eq. (13), but in most cases the difference remains negligible compared with the data errors. A fully symmetric approach for quadratic models is sketched in Appendix B.

Results of the chi-square technique for models (13) and (14) are presented for a data base of Central Europe. The input data ($N = 164$ magnitude pairs; Grünthal and Wahlström, 2003, Table 6) cover a wide range of M_w and M_L magnitudes (Figure 3) and form the basis for computation of M_w for all events in the study by Grünthal and

Wahlström (2003).

Similar to the North Atlantic data in the previous chapter there is no detailed information on the errors of the input data. As a first attempt we assume unknown but equal standard errors $\sigma = \sigma(M_w) = \sigma(M_L)$ of both magnitudes. For a reasonable value of goodness-of-fit $Q_0 = Q(\nu, \nu)$, the chi-squares become $\chi^2(a, b, c) = \chi^2(A, B, C) = N - 3 = \nu$. The appropriate error estimations are then

$$\sigma = \sigma(\hat{a}, \hat{b}, \hat{c}) = \sqrt{\sum_{i=1}^N \frac{[M_{wi} - \hat{a} M_{Li}^2 - \hat{b} M_{Li} - \hat{c}]^2}{1 + (2\hat{a} M_{Li} + \hat{b})^2}} / (N - 3)$$

for model (13) and

$$\sigma = \sigma(\hat{A}, \hat{B}, \hat{C}) = \sqrt{\sum_{i=1}^N \frac{[M_{Li} - \hat{A} - \sqrt{\hat{B} + \hat{C} M_{wi}}]^2}{1 + \hat{C}^2 / (4(\hat{B} + \hat{C} M_{wi}))}} / (N - 3)$$

for model (14), respectively. Table 3 gives the results of the chi-square regression including standard errors of the model parameters in comparison with the standard least-squares solutions. Figure 3 illustrates the difference between the two approaches, which is notable for small and large magnitudes. As mentioned above, the chi-square regression curves are different for model (13) and model (14), but the differences are too small to be distinguished in the figure. The quality of the fit is given in terms of the estimated magnitude error, $\sigma = 0.227$. The solution is acceptable if the real mean magnitude uncertainties are greater than σ . This is plausible for the selected data set. The parameter errors of model (13) are visualized in Figure 4 by means of the mean fitting error

$$\sigma(a, b, c) = \sigma(\hat{a}, \hat{b}, \hat{c}) \sqrt{\frac{\chi^2(\hat{a}, \hat{b}, \hat{c}) + \Delta\chi^2(a, b, c)}{N - 3}} \quad (17)$$

The confidence regions for the individual errors, $\Delta\chi^2 = 1$, and the joint errors, $\Delta\chi^2 = 3.53$ (see Table 1), are connected with values of $\sigma(a, b, c) = 0.228$ and $\sigma(a, b, c) = 0.229$, respectively.

Table 3. Chi-square and least-squares regression parameters, standard errors and mean fitting error σ of model (13) and model (14)

	$M_w = a M_L^2 + b M_L + c$						
	a	$\sigma(a)$	b	$\sigma(b)$	c	$\sigma(c)$	σ
chi-square	0.046	0.009	0.556	0.064	0.673	0.095	0.227
least-squares	0.059	0.009	0.450	0.064	0.840	0.097	0.284

	$M_L = A + \sqrt{B + C M_w}$						
	A	$\sigma(A)$	B	$\sigma(B)$	C	$\sigma(C)$	σ
chi-square	-5.948	0.318	20.57	3.596	21.55	0.933	0.227
least-squares	-13.37	5.286	164.2	139.5	36.88	11.76	0.360

Local magnitude M_L , intensity I_0 and focal depth h

Epicentral intensity, I_0 , is often the only available strength measure for historical earthquakes in local catalogues. The basic data used for the catalogue by Grünthal and Wahlström (2003) do not permit a reliable regression of M_w on I_0 . Instead, M_L is regressed on I_0 for each local catalogue for which such a relation does not already exist but where intensities need to be converted. Each of these relations is then combined with an M_w - M_L relation (see previous chapter).

A well established empirical model to estimate M_L from I_0 is the relation

$$M_L = a I_0 + b \log(h) + c \quad (18)$$

where the logarithmic term takes into account the influence of the focal depth h . Early compilations of several applications of model (18) are given by Sponheuer (1962) and Kárník (1969). Recently, Gutdeutsch et al. (2002) used this model to compare magnitude-intensity relations for Europe. The corresponding inverse model with respect to M_L and I_0 is

$$I_0 = A M_L + B \log(h) + C \quad (19)$$

With errors $\sigma(M_L)$, $\sigma(I_0)$ and $\sigma(h)$ for all three input parameters, the merit functions of a chi-square approach become

$$\chi^2(a, b, c) = \sum_{i=1}^N \frac{[M_{Li} - a I_{0i} - b \log h_i - c]^2}{\sigma^2(M_{Li}) + a^2 \sigma^2(I_{0i}) + b^2 \frac{\sigma^2(h_i)}{(\ln(10)h_i)^2}} \quad (20)$$

for model (18) and

$$\chi^2(A, B, C) = \sum_{i=1}^N \frac{[I_{0i} - A M_{Li} - B \log h_i - C]^2}{\sigma^2(I_{0i}) + A^2 \sigma^2(M_{Li}) + B^2 \frac{\sigma^2(h_i)}{(\ln(10)h_i)^2}} \quad (21)$$

for model (19), respectively. Standard least-squares solutions are obtained, setting $\sigma(I_0) = \sigma(h) = 0$ in eq. (20) or $\sigma(M_L) = \sigma(h) = 0$ in eq. (21). An often used orthogonal regression (Gutdeutsch et al., 2002) for model (18) implies $\sigma(M_L) = \sigma(I_0) = \sigma(\log(h)) = \text{constant}$, where the constant does not need to be specified in advance. Assuming a plausible error $\sigma(M_L) \approx \sigma(I_0) \approx 0.5$, this results in

$$\sigma(\log(h)) \approx \frac{\sigma(h)}{\ln(10)h} = 0.5 \quad (22)$$

or approximately $\sigma(h) \approx h$. This implies a strong overestimation of $\sigma(h)$ for large h and therefore an orthogonal regression is strongly determined by shallow events.

Obviously, eq. (18) and eq. (20) can be converted into eq. (19) and eq. (21), respectively, by the same transformation

$$a = 1/A \quad b = -B/A \quad c = -C/A$$

This means that the models are equivalent, a remarkable advantage compared to any least-squares procedure where the corresponding models produce different relations.

The numerical minimization of eq. (18) with respect to a , b and c requires one of the modern

Figure 5. 2-D representation of the chi-square regression surface of M_L , I_0 and h for a) Fennoscandia and b) Germany. The depth of the events is coded with different grey colours. The width of the stripe indicates the influence of the spread of focal depths, with the largest depth representing the upper limit and the smallest depth the lower limit.

Figure 6. Mean fitting error $\sigma(a, b, c)$ with confidence regions of individual (dark grey) and joint (light grey) parameter errors. a) Fennoscandia, $\sigma(a, b, c) = 0.331$ for individual errors and $\sigma(a, b, c) = 0.340$ for joint errors. b) Germany, $\sigma(a, b, c) = 0.388$ for individual errors and $\sigma(a, b, c) = 0.395$ for joint errors.

nonlinear optimization tools and appropriate starting values close to the unknown solution. We recommend starting values equal to the mean of the uniquely determined least-squares solutions of eq. (18) and eq. (19) from our experiences comparing nonlinear optimization and extensive Monte-Carlo simulation for several case studies.

Chi-square regression to estimate the parameters of models (18) and (19) were performed for six local catalogue data sets in the study by Grünthal and Wahlström (2003). Similar to the previous

regressions, there are no quantitative error estimations for the input data. Generally, it can be expected that the uncertainties of M_L and I_0 are in the same range of about 0.5. We suppose they are equal to an unknown value $\sigma = \sigma(M_L) = \sigma(I_0)$. Errors for the focal depth h are hard to estimate. For carefully selected data, the errors should be significantly smaller than the depth itself. Then $\sigma(h)/h \ll \ln(10)$ and the influence of $\sigma(h)$ in the denominator of eq. (20) and eq. (21) can be neglected. This assumption requires a well deter-

Table 4. Chi-square regression parameters, standard errors and mean fitting error σ of model (18) and model (19) for Fennoscandia and Germany

	$M_L = a I_0 + b \log(h) + c$						
	a	$\sigma(a)$	b	$\sigma(b)$	c	$\sigma(c)$	σ
Fennoscandia	0.880	0.089	0.640	0.246	-1.516	0.451	0.328
Germany	0.745	0.054	0.777	0.232	-0.840	0.362	0.385

	$I_0 = A M_L + bB \log(h) + C$						
	A	$\sigma(A)$	B	$\sigma(B)$	C	$\sigma(C)$	σ
Fennoscandia	1.136	0.111	-0.762	0.293	1.721	0.384	0.328
Germany	1.363	0.098	-1.044	0.316	1.174	0.424	0.385

mined focal depth for very shallow earthquakes and reduces the chi-square function (20) to

$$\chi^2(a, b, c) = \sum_{i=1}^N \frac{[M_{Li} - a I_{0i} - b \log h_i - c]^2}{\sigma^2(1 + a^2)} \quad (23)$$

Since the tectonic type of very shallow events is often questionable, the regressions by Grünthal and Wahlström (2003) include only events with $h \geq 5$ km, for which the criterion $\sigma(h)/h \ll \ln(10)$ is fulfilled. The approximation (23) can be geometrically understood as a partial orthogonal regression in the (M_L, I_0) plane of the projected differences between model and data. In Appendix C we reformulate the nonlinear minimization problem connected with the determination of a , b and c in eq. (23) to a simple eigenvalue problem. This avoids the application of iterative optimization algorithms and ensures a correct solution independent of any starting values. The estimation of the unknown mean fitting error σ in eq. (23) for a reasonable value of goodness-of-fit $Q_0 = Q(v, v)$ expressed by the optimal model parameters $\hat{a}$, $\hat{b}$ and $\hat{c}$ is

$$\begin{aligned} \sigma &= \sigma(\hat{a}, \hat{b}, \hat{c}) \\ &= \sqrt{\sum_{i=1}^N \frac{[M_{Li} - \hat{a} I_{0i} - \hat{b} \log h_i - \hat{c}]^2}{1 + \hat{a}^2}} / (N - 3) \end{aligned}$$

In Table 4, the results of the chi-square regression are given for two of the studied regions, Fennoscandia with 101 entries and Germany with 145 entries. Figure 5 visualizes the input data and the dominant linear part of the two regression surfaces by a 2D-plot, in which the width of the stripes indicates the influence of the focal depth h in model (18).

The confidence regions of the model parameters for Fennoscandia and Germany in terms of the mean fitting errors are shown in Figure 6.

Conclusions

The compilation of earthquake catalogues often require empirical relations for the conversion of macroseismic data to magnitude and for conversions between different types of magnitudes. The errors of the data should be considered when such relations are established. Standard least-squares techniques are restricted to include errors of only one (dependent) parameter and therefore these regressions result in asymmetrical relations.

A generalized chi-square regression approach is adapted for magnitude relations with uncertainties in all input data. Individual error distributions for each data entry can be taken into account. This approach fulfils all necessary criteria of a fitting procedure. It provides regression parameters, error estimates on the parameters and a statistical measure of goodness-of-fit. In most cases, a chi-square approach can preserve the symmetry of the empirical relations, which ensures the path independent determination of magnitudes in a hierarchy of conversions.

The presented chi-square technique can also be utilized if the errors of input data are not exactly known, a typical situation in establishing magnitude relations where at most error constraints are available. An orthogonal regression approach, often suggested in such cases, is only reasonable if the regression model is linear and the unknown input errors are equal for all data. For more complex regression models and differentiated errors, the described chi-square technique is preferred.

Acknowledgements

We are indebted to D. Mayer-Rosa and an anonymous reviewer for valuable suggestions which have improved the text.

References

- ‘Global Hypocenter Data Base’, 1996, U.S. Geological Survey, National Earthquake Information Center, World Data Center A for Seismology, Boulder, Colorado. CD version 2.0, 1996.
- Bevington, P.R.: 1969, *Data Reduction and Error Analysis for the Physical Sciences*, McHill, New York.
- Björk, Å.: 1996, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia.
- Bollinger, G.A., Chapman, M.C. and Sibol, M.S., 1993, A comparison of earthquake damage areas as a function of magnitude across the United States, *Bull. Seism. Soc. Am.* **83**, 1064-1080.
- Grünthal, G. and Wahlström, R., 2003, An M_w based earthquake catalogue for central, northern and northwestern Europe using a hierarchy of magnitude conversions, *J. Seismology* **7**, 507-531.
- Gutdeutsch, R., Kaiser, D. and Jentzsch, G., 2002, Estimation of earthquake magnitudes from epicentral intensities and other focal parameters in Central and Southern Europe, *Geophy. J. Int.* **151**, 824-834.
- Hasegawa, H.S., 1983, Lg spectra of local earthquake recorded by the Eastern Canada Telemetered Network and spectral scaling, *Bull. Seism. Soc. Am.* **73**, 1041-1061.
- Kárník, V., 1969, *Seismicity of the European area, Part 1*, Reidel Publishing Company, Dordrecht, Holland.
- Mises von, R., 1964, *Mathematical Theory of Probability and Statistics*, Academic Press, New York.
- Nuttli, O.W., 1983, Average seismic source parameter relations for mid-plate earthquakes, *Bull. Seism. Soc. Am.* **73**, 519-535.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. and F.B.P., 1992, *Numerical Recipes in C The Art of Scientific Computing*, University Press, Cambridge.
- Sponheuer, W., 1962, Untersuchungen zur Seismizität von Deutschland, *Veröffentlichungen des Instituts für Bodenmechanik und Erdbebenforschung in Jena, Heft 72*, pp. 23-52.
- Street, R.L., Hermann, R.B. and Nuttli, O.W., 1975, Spectral characteristics of the Lg wave generated by central United States earthquakes, *Geophys. J. R. Astron. Soc.* **41**, 51-63.
- Uhrhammer, R.A., Lopez, S.J. and Romanowicz, B., 1996, Determination of local magnitude using BDSN broadband records, *Bull. Seism. Soc. Am.* **86**, 1314-1330.
- Varga, R.S., 1962, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Wahlström, R. and Grünthal, G., 2000, Probabilistic seismic hazard assessment (horizontal PGA) for Sweden, Finland and Denmark using different logic tree approaches, *Soil Dyn. Earthquake Engineering* **20**, 45-58.

Appendix A

Rewriting the minimization problem of the chi-squares (9) with constant errors $\sigma(M_s)$ and $\sigma(m_b)$ in a compact matrix form

$$\min_{a,b} \chi^2(a,b) = \min_{\mathbf{u},b} \frac{\|\mathbf{P}\mathbf{u} - b\mathbf{1}\|^2}{\mathbf{u}^T\mathbf{u}} \quad (\text{A1})$$

is obtained, where vector $\mathbf{u}^T = (\sigma(M_s), a \sigma(m_b))$, matrix

$$\mathbf{P}^T = \begin{pmatrix} \frac{M_{S1}}{\sigma(M_s)} & \frac{M_{S2}}{\sigma(M_s)} & \cdots & \frac{M_{SN}}{\sigma(M_s)} \\ \frac{-m_{b1}}{\sigma(m_b)} & \frac{-m_{b2}}{\sigma(m_b)} & \cdots & \frac{-m_{bN}}{\sigma(m_b)} \end{pmatrix}$$

and $\mathbf{1}$ is a column vector of N ones. At the minimum of (A1), $\frac{\partial \chi^2}{\partial b} = 0$. This results in a linear relation $b = \mathbf{q}^T\mathbf{u}$ between b and $\mathbf{u}$. Here, $\mathbf{q}$ is a vector whose components are the averages of each column of $\mathbf{P}$

$$\mathbf{q}^T = \left(\sum_{i=1}^N \frac{M_{Si}}{N \sigma(M_s)}, \sum_{i=1}^N \frac{-m_{bi}}{N \sigma(m_b)} \right)$$

With $\mathbf{R} = \mathbf{P} - \mathbf{1}\mathbf{q}^T$ eq. (A1) becomes

$$\min_{\mathbf{u}} \frac{\|\mathbf{R}\mathbf{u}\|^2}{\mathbf{u}^T\mathbf{u}} = \min_{\mathbf{u}} \frac{\mathbf{u}^T\mathbf{R}^T\mathbf{R}\mathbf{u}}{\mathbf{u}^T\mathbf{u}} \quad (\text{A2})$$

The global minimum of the *Rayleigh quotient* (A2) is the square of the smallest eigenvalue λ of the eigenvalue problem $\mathbf{R}^T\mathbf{R}\mathbf{u} - \lambda\mathbf{u} = 0$ with a positive symmetric matrix $\mathbf{R}^T\mathbf{R}$ (Varga, 1962). Because the total length of the associated eigenvector $\mathbf{u}^T = (u_1, u_2)$ does not change the value of eq. (A2), $\mathbf{u}$ can be normalized in such a way that $u_1 = \sigma(M_s)$ which implies

$$a = \frac{u_2}{\sigma(m_b)} \quad \text{and} \quad b = \mathbf{q}^T\mathbf{u}$$

From the numerical point of view it is recommended to calculate λ and $\mathbf{u}$ by a singular value decomposition of $\mathbf{R}$ instead of determining the eigenvalues and eigenvectors $\mathbf{R}^T\mathbf{R}$ (Björk, 1996).

Let $\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, with $\mathbf{U}$ and $\mathbf{V}$ orthogonal and $\mathbf{S}$ diagonal with positive, decreasing elements

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Then the minimum value of eq. (A2) is $\lambda^2 = \lambda_N^2$ and $\mathbf{u}$ is the N th column of $\mathbf{V}$.

Appendix B

A symmetric second order relation between M_w and M_L requires a five parameter model. All linear and quadratic terms in both variables have to be considered while the constant term can be scaled to 1.

$$a M_w^2 + b M_w M_L + c M_L^2 + d M_w + e M_L + 1 = 0 \quad (\text{B1})$$

The corresponding chi-square merit function reads as follows

$$\chi^2(a, b, c, d, e) = \sum_{i=1}^N \frac{[a M_w^2 + b M_w M_L + c M_L^2 + d M_w + e M_L + 1]^2}{A^2 \sigma^2(M_w) + B^2 \sigma^2(M_L)}$$

With $A = 2a M_w + b M_L + d$ and $B = 2c M_L + b M_w + e$.

Depending on the solution, the implicit representation (B1) can be resolved explicitly for each of M_w and M_L .

Appendix C

Assuming constant errors $\sigma(M_L)$ and $\sigma(I_0)$ and neglecting $\sigma(h)$ in eq. (20), the parameters a , b and c are the solution of the minimization problem

$$\min_{a,b,c} \chi^2(a, b, c) = \min_{\mathbf{u}, \mathbf{b}, \mathbf{c}} \frac{\|\mathbf{P}\mathbf{u} - \mathbf{b}\mathbf{v} + \mathbf{c}\mathbf{1}\|^2}{\mathbf{u}^T \mathbf{u}} \quad (\text{C1})$$

with $\mathbf{u}^T = (\sigma(M_L), a \sigma(I_0))$ and $\mathbf{v}^T = (\log(h_1) \dots \log(h_N))$. $\mathbf{1}$ is a column vector of N ones and $\mathbf{P}$ is the matrix

$$\mathbf{P}^T = \begin{pmatrix} \frac{M_{L1}}{\sigma(M_L)} & \frac{M_{L2}}{\sigma(M_L)} & \dots & \frac{M_{LN}}{\sigma(M_L)} \\ \frac{-I_{01}}{\sigma(I_0)} & \frac{-I_{02}}{\sigma(I_0)} & \dots & \frac{-I_{0N}}{\sigma(I_0)} \end{pmatrix}$$

At the minimum of eq. (C1), $\frac{\partial \chi^2}{\partial b} = 0$ and $\frac{\partial \chi^2}{\partial c} = 0$. This results in a linear system

$$b = \mathbf{q}^T \mathbf{u} + s \quad c = \mathbf{r}^T \mathbf{u} + t$$

between b , c and $\mathbf{u}$ with vectors

$$\mathbf{r}^T = \left(\sum_{i=1}^N \frac{M_{Li}}{N \sigma(M_L)}, \sum_{i=1}^N \frac{-I_{0i}}{N \sigma(I_0)} \right)$$

$$\mathbf{q}^T = \left(\sum_{i=1}^N \frac{M_{Li} \log(h_i)}{\sigma(M_L)}, \sum_{i=1}^N \frac{-I_{0i} \log(h_i)}{\sigma(I_0)} \right) / \sum_{k=1}^N \log(h_k)$$

and scalars $t = -\sum_{i=1}^N \log(h_i)/N$ and $s = -N$. Inserting the solution

$$b = \frac{(\mathbf{q}^T + s \mathbf{r}^T)}{1 - s t} \quad c = \frac{(\mathbf{r}^T + t \mathbf{q}^T)}{1 - s t} \quad (\text{C2})$$

into eq. (C1) gives

$$\min_{\mathbf{u}} \frac{\|\mathbf{R}\mathbf{u}\|^2}{\mathbf{u}^T \mathbf{u}} = \min_{\mathbf{u}} \frac{\mathbf{u}^T \mathbf{R}^T \mathbf{R} \mathbf{u}}{\mathbf{u}^T \mathbf{u}}$$

with

$$\mathbf{R} = \mathbf{P} - \frac{(\mathbf{q}^T + s \mathbf{r}^T) \mathbf{u}}{1 - s t} \mathbf{v} - \frac{(\mathbf{r}^T + t \mathbf{q}^T) \mathbf{u}}{1 - s t}$$

Following the same arguments as in Appendix A, u_1 can be scaled to $u_1 = \sigma(M_L)$ which implies $a = u_2/\sigma(I_0)$. Finally, b and c can be evaluated from eq. (C2).