Automatic computation of moment magnitudes for small earthquakes and the scaling of local to moment magnitude

Benjamin Edwards, Bettina Allmann, Donat Fäh and John Clinton

Schweiz. Erdbebendienst, Sonneggstrasse 5, 8092 Zürich, Switzerland. E-mail: edwards@sed.ethz.ch

Accepted 2010 July 16. Received 2010 July 15; in original form 2009 July 16

SUMMARY

Moment magnitudes ($M_W$) are computed for small and moderate earthquakes using a spectral fitting method. 40 of the resulting values are compared with those from broadband moment tensor solutions and found to match with negligible offset and scatter for available $M_W$ values of between 2.8 and 5.0. Using the presented method, $M_W$ are computed for 679 earthquakes in Switzerland with a minimum $M_L = 1.3$. A combined bootstrap and orthogonal L1 minimization is then used to produce a scaling relation between $M_L$ and $M_W$. The scaling relation has a polynomial form and is shown to reduce the dependence of the predicted $M_W$ residual on magnitude relative to an existing linear scaling relation. The computation of $M_W$ using the presented spectral technique is fully automated at the Swiss Seismological Service, providing real-time solutions within 10 minutes of an event through a web-based XML database. The scaling between $M_L$ and $M_W$ is explored using synthetic data computed with a stochastic simulation method. It is shown that the scaling relation can be explained by the interaction of attenuation, the stress-drop and the Wood–Anderson filter. For instance, it is shown that the stress-drop controls the saturation of the $M_L$ scale, with low-stress drops (e.g. 0.1–1.0 MPa) leading to saturation at magnitudes as low as $M_L = 4$.

Key words: Fourier analysis; Inverse theory; Earthquake ground motions; Earthquake source observations; Seismic attenuation.

1 INTRODUCTION AND BACKGROUND

The computation of moment magnitudes ($M_W$) of small earthquakes ($M_W < 3.5$) is not a routine procedure at observatories such as the Swiss Seismological Service (SED). Classical determination of $M_W$ using moment tensor solutions (e.g. Bernardi et al. 2004; Clinton et al. 2006), hereinafter referred to as $M_W$(MT), relies upon the availability of long period data (e.g. greater than 5 s). As such it is limited by a minimum magnitude threshold, above which the recorded ground motions at these long periods are routinely stronger than the background noise. This minimum magnitude is typically in the range of $M_W$ 3.5–4 for local events, but depends on the density of stations and the event location.

For regions of low-to-moderate seismicity, the determination of $M_W$ for small events is desirable to obtain robust $a$ and $b$ values in a recurrence relationship and the magnitude of completeness of the earthquake catalogue (Woessner & Wiemer 2005). Typically, $M_L$ is converted to $M_W$ using a simple scaling relation derived from the larger events (where both magnitudes can be independently determined). In addition, although the use of small earthquakes in deriving empirical ground–motion relationships is strongly debated, their use may lead to more realistic predictive equations that capture features that scale with magnitude, for instance, the magnitude dependence of geometrical spreading (Cotton et al. 2008). The use of poor meta-data, such as $M_W$ inappropriately scaled from $M_L$, for small-magnitude events introduces a significant uncertainty in the prediction of ground–motion at low magnitudes (e.g. Bommer et al. 2007). To extend the available magnitude range for empirical ground–motion relationships, the direct determination (or appropriate conversion) of $M_W$ for small earthquakes is therefore vital. The work detailed in this article forms the basis of the scaling relation used to create a homogenized magnitude scale for the earthquake catalogue of Switzerland 2009 (ECOS09), which is an update of the previous catalogue, ECOS02 (Fäh et al. 2003a). Extensions of the scaling relation to lower and higher magnitudes than those covered in the scope of this paper will be published elsewhere.

The Swiss Seismological Service (SED) is responsible for the operation and maintenance of a large digital broadband seismic network in Switzerland (SDSNet) in addition to a large strong-motion network (SSMNet). Collectively, these two networks are referred to as CHNet. By the end of 2006, CHNet included 52 sites with 64 modern high-quality sensors: 27 three-component Streckeisen STS-2 broadband seismometers; eight three-component Lennertz 5-s LE-3D/5s sensors; one three-component Lennartz 1-s LE-3-D/1s sensor; and 28 three-component Kinematics Epsensor strong-motion sensors (Baer et al. 2007). The network is however, rapidly expanding. In addition to CHNet, several digital broadband instruments of foreign networks operating near the Swiss border region are available for use in real-time. These include instruments of the Italian National Seismic Network, the German Regional Seismic
Network, the Landes Erdbebenendienst Baden-Württemberg and the Austrian Broadband Seismic Network. In total, around 80 sites with broadband digital recording instruments are routinely available in Switzerland and its border regions.

Data from the stations of the CHNet and selected stations of the bordering networks are digitally transmitted to a central processing server in real time. Automatic phase picking and hypocentre location is performed as soon as an earthquake is detected. Estimated location, $M_L$, and pick times are then available within minutes of an earthquake. $M_L$ is calculated following a modification of the original formulation of Gutenberg & Richter's (1954) local magnitude scale:

$$M_L(\text{SED}) = \log A_0 + C_d + C_e,$$

where $A_0$ is the maximum mean-to-peak horizontal amplitude in millimetres on a synthesized 2800× Wood–Anderson torsion seismometer and $C_d$ is a distance correction, given by

$$C_d = 0.0180 R + 1.77 \quad \text{for } R = 60 \text{ km}$$
$$C_d = 0.0038 R + 2.62 \quad \text{for } R > 60 \text{ km}$$

and where $R$ is the epicentral distance in kilometres (Kradolfer 1984). This scaling relation was derived using vertical recordings on short-period instrumentation with analogue telemetry. $C_e = 0.1$ is an empirically derived correction to the scaling relation that was implemented to maintain consistency in the $M_L(\text{SED})$ scale after:

(a) the update of the Swiss seismic network from the original analogue short-period to a broadband digital system. Contribution to $C_e$ is 0.4;

(b) a change to the use of the horizontal component of motion rather than vertical. Contribution to $C_e$ is $-0.3$.

It is well known that many agencies report different systematic trends in the relation of $M_L$ to $M_w$. One disadvantage of the $M_L$ scale is that the effects of frequency-dependent attenuation (e.g. due to scattering and intrinsic $Q$) are only loosely considered by the distance correction. In addition, the distance term is often calibrated to southern California (although not in the case of Switzerland) and does not consider effects that could be regionally variable, such as $Q$ or stress-drop (Deichmann 2006).

$M_w$ is defined to be proportional to the seismic moment ($M_0$) measured in Nm. This is itself related to the fault area, $A$, the average slip across the fault, $d$, and the shear strength of the rock, $\mu$, by

$$M_w = 2/3 \log(M_0) - 6.03 = 2/3 \log(\mu Ad) - 6.03. \quad (3)$$

This measure of magnitude is independent of the attenuation effects previously described (Hanks & Kanamori 1979). It is therefore useful to provide this measure of magnitude alongside the traditional $M_L$. Currently, the SED operates an automatic real-time moment tensor inversion for local earthquakes (Clinton et al. 2006), which is based on the method of Dreger (2003). The automatic $M_w(\text{MT})$ solution is typically available within 15 minutes of the event. However, they are only available for larger earthquakes (i.e. greater than $M_L 3.5–4$) due to the presence of noise at low frequencies for smaller events.

It has long been known that the seismic moment, $M_w$, of an earthquake can be related to the long-period spectral plateau of earthquake spectra (e.g. Brune 1970; 1971; Eshelby 1957; Madagaria 1976). Although the seismic moment is related to the long-period spectral level, for small earthquakes (i.e. $M_w < 3$) the ‘long-period’ plateau is actually present at fairly high frequencies: for an earthquake with $M_w = 3$ and stress-drop of 10 MPa the theoretical corner-frequency is around 10 Hz. Furthermore, the corner-frequency increases inversely proportional to the cube of the seismic moment for a constant stress-drop, such that even though smaller events suffer more due to low-frequency noise, the plateau is present at higher and higher frequencies. As a result, the robustness of the determination of the long-period spectral level is good even for small earthquakes where the background noise dominates the recording at low frequencies. Following previous work of Ottemöller & Havskov (2003), Edwards et al. (2008), Edwards (2008) and Edwards & Rietbrock (2009) it is shown that a robust determination of $M_w$ is possible using spectral fitting methods and the instantaneous slip model of Brune (1970, 1971). The method has been improved and targeted specifically at computation of $M_w$, hereinafter referred to as $M_w(\text{SP})$, at the SED. In addition, the process is fully automated by checking for new earthquake data (automatic pick and location) and updating an XML database with any solutions within five minutes.

This paper is presented in three main parts. The first section describes the derivation of $M_w$ for small and moderate earthquakes in Switzerland using a spectral fitting method, including a comparison with existing $M_w(\text{MT})$ values computed using a moment tensor inversions. In the second section, we use the new $M_w(\text{SP})$ to define a $M_L$ to $M_w$ scaling relation over the range of $M_L$ values from 1.3 to 5.4. In the final section, we try to understand the observed scaling relation in terms of source and path effects using synthetic data.

## 2 COMPUTATION OF $M_w$ FROM EARTHQUAKE SPECTRA

### 2.1 Data processing

Data processing and quality control play a crucial role in the determination of $M_w$ using spectral fitting techniques. It is particularly important for small earthquakes, where the background seismic noise, rather than the earthquake, dominates the low-frequency spectral amplitudes. The data collection and processing approach used in this study is as follows: the waveforms are first windowed to provide data that are categorized as signal and noise: one of each being necessary for each record. As $S$-wave arrivals are not automatically picked, a $P:S$ ratio of 1.73 is used to estimate the arrival time from the automatic $P$-wave picks. The signal window initially starts 5 s before the $S$-pick and is composed of unclipped data of at least 10 s, and at most 50 s in duration, depending on the availability of data. This initial window is then refined, following the method proposed by Raof et al. (1999) to encapsulate 5–75 per cent of the cumulative squared velocity of the record. After this procedure, the signal windows usually start at the $S$-wave arrival with lengths that are significantly shorter than the original maximum of 50 s, dependent on the source–station distance. This approach also means that the results are relatively insensitive to the pick-times of the $P$ and $S$ waves, and in addition, provides a quantitative measure of shaking duration based on the cumulative squared velocity for potential further analysis. An example of data from one of the stations used in this paper is shown in Fig. 1.

Following the selection of the signal window, a pre-$P$-wave noise window is selected. Its duration is chosen to be as long as possible to enable the correct recovery of any long-period noise. Specifically, the noise window starts at the beginning of the trace, and continues over a duration equal to 75 per cent of the time until the $P$-wave arrival. This ensures that potentially inaccurate $P$-wave arrival time picks (or estimates) do not lead to the noise window being contaminated by the $P$-arrival. It should be noted that for events recorded

© 2010 The Authors, GJI, 183, 407–420
Geophysical Journal International © 2010 RAS
near to the source, the \( P \)-coda may contaminate the signal window. As we estimate the noise level using data prior to the event, this may lead to an underestimation of the noise level. However, as described later in the paper, we conservatively increase the estimated noise level to account for any underestimation. Zero padding is then used to create equal length windows (signal and noise), and provide the 2\(^{n}\) samples required for the Fast Fourier Transform (FFT). Both signal and noise windows (excluding the zero padding) are then demeaned and tapered using multitaper algorithms (Park et al. 1987; Lees & Park 1995) with 5–3\(\pi\) prolate tapers, before applying an FFT and multiplying by the sampling period to obtain the two-sided continuous Fourier velocity spectrum of each record. To account for different window lengths used for the signal and noise spectra and the subsequent zero-padding, the noise spectra are also normalized by \( N_{\text{signal}}/N_{\text{noise}} \), where \( N \) is the number of points in the time-series (excluding the zero-padding). For accelerometric records, the spectral data are integrated by dividing by \( i\omega \). All data are then corrected to records of ground–velocity through deconvolution with the complex instrument response function.

To extract only the most reliable data, the noise estimate is automatically increased to insure intersection of the signal and noise at both the lowest and highest frequencies of each spectrum. Although this may reduce the available bandwidth above the signal-to-noise ratio, it prevents unwanted effects from being included in the data. For instance, this includes effects due to the lack of periodicity due to trends in the time-series, the high variance of low-frequency spectral levels for relatively short window lengths, the time-dependence of the noise, or the presence of \( P \)-coda in the signal window. The useful frequency range is then defined for each recording as where the signal-to-noise ratio is greater than 3 over a continuous bandwidth of at least 1 log\(_{10}\) units of frequency (e.g. 1–10 Hz or 0.5–5 Hz). For inclusion in the spectral inversion, only frequencies up to 30 Hz are considered.

2.2 Inversion model formulation

A simplification of the method detailed in Edwards et al. (2008) and Edwards & Rietbrock (2009) is applied to the data in order to robustly determine \( M_W \) using the following model of the Fourier velocity spectrum:

\[
\Omega_i(f, r) = 2\pi f E_i(f, f_c)B_i(f, t'_0)S_i(r, r_0, a, \lambda, A_j, f),
\]

where \( E_i(f, f_c) \) is the (Brune 1970, 1971) source model for event \( i \) with a defining corner-frequency \( f_c \), \( B_i(f, t'_0) \) is the attenuation along the ray path

\[
B_i(f) = e^{-\pi f t'_0},
\]

where \( t'_0 \) is the frequency independent whole path attenuation operator (both crustal and site attenuation), which is sometimes...
separated into a crustal attenuation parameter $r^*$ and site attenuation parameter $\kappa$ (e.g. Anderson & Hough 1984) and is defined by

$$t_{ij}^* = \frac{t_{ij}}{Q_{ij}}, \quad (6)$$

where $t_{ij}$ is the traveltime and $Q_{ij}$ is the path-dependent dimensionless quality factor (Rietbrock 2001). It should be noted that due to the relatively long window lengths used (encapsulating the duration of significant shaking), the degree of attenuation may vary at different points in the window due to different paths and travel times of the individual phases. In this case, therefore the attenuation parameter, $t_{ij}^*$, obtained from inversion will describe the average attenuation over all the travel paths (e.g. Ottemöller 2002). $S_i(r, r_0, \ldots, r_{n-1}, \lambda_1, \ldots, \lambda_n)$ is the amplitude decay with distance or 'apparent geometrical spreading', given by

$$S_i(r, r_0, \ldots, r_{n-1}, \lambda_1, \ldots, \lambda_n) = \left\{ \begin{array}{ll}
\frac{(r)}{r}^{\lambda_1} & r \leq r_1, \\
S(r_1) \left( \frac{r_1}{r} \right)^{\lambda_2} & r_1 \leq r \leq r_2, \\
\vdots & \vdots \\
S(r_n) \left( \frac{r_n}{r} \right)^{\lambda_n} & r \geq r_n,
\end{array} \right. \quad (7)$$

where $r$ is the hypocentral distance, $r_{0, \ldots, n}$ are distances at which the rate of decay changes from the $\lambda_0$ to the $\lambda_{n+1}$ decay value, and $r_0$ is the rupture radius, which can be normalized to $r_0 = 1 \text{ km}$ without affecting $M_w(\text{SP})$ as long as $r_0 = 1 \text{ km}$ is maintained in eq. (13). $T_i(f, A_i)$ is the site transfer function (including the generic 'rock-amplification' due to increasing velocity with depth (e.g. Boore & Joyner 1997)) at the recording station which is given by

$$T_i(f, A_i) = A_i a_i(f), \quad (8)$$

where $A_i$ is the average site amplification and $a_i(f)$ the normalized frequency-dependent site amplification, relative to a reference site (that may not necessarily be known). Finally, $I_i(f)$ is the instrument response function which is set to $I_i(f) = 1$ as the response of the instrument is removed from the data prior to the inversion.

In this paper, we refer to the signal moment, which is the far-field long period spectral amplitude, or the frequency-independent parameters in eq. (4). It is defined as

$$\hat{Q}_{ij} = \Psi_i A_i S_i (r, r_0, \ldots, r_{n-1}, \lambda_1, \ldots, \lambda_n), \quad (9)$$

where $S_i(r, r_0, \ldots, r_{n-1}, \lambda_1, \ldots, \lambda_n)$ is given by eq. (7), $A_i$ is the site amplification parameter independent of frequency, and $\Psi_i$ is the effective long period plateau value at the source location.

### 2.2.1 Inversion procedure

The inversion approach is summarized as follows. First, an initial inversion of the spectra is performed for three defining parameters—the source corner-frequency ($f_c^*$) of the Brune (1970) far-field source spectrum, the signal moment (the frequency-independent far-field amplitude) and path-dependent $Q_{ij}$. The inversion routine implements a combined grid-search for $f_c^*$ and Powell’s minimization for $t_{ij}^*$ and the signal moment (Press et al. 1997). The grid search of $f_{c0}$ is performed at 10 per cent intervals, starting at the equivalent $f_c$ value for a 0.001 MPa approximate stress-drop and increasing to at the equivalent $f_c$ for a 100 MPa approximate stress-drop, such that the $n$th grid-search value of $f_{c0}$ is given by

$$f_{c,n} = 1.1^n \left( 0.4906 \beta \left( \Delta \sigma_{\text{max}}/10^{1.5 \text{ MPa}+9.1} \right)^{1/3} \right) \quad (10)$$

while

$$f_{c,n} < \left( 0.4906 \beta \left( \Delta \sigma_{\text{max}}/10^{1.5 \text{ MPa}+9.1} \right)^{1/3} \right) \quad (11)$$

with $\Delta \sigma_{\text{max}} = 0.001 \text{ MPa}$, $\Delta \sigma_{\text{max}} = 100 \text{ MPa}$ and $\beta = 3500 \text{ ms}^{-1}$. Note that an assumption of $M_L \approx M_w$ is made in eqs (10) and (11) to approximate the starting point of the grid search over $f_c$ (as the $M_w$ value is as yet unknown). Note, however, that the wide search range (0.001 < $\Delta \sigma < 100$ MPa) prevents any bias in results. Each earthquake is assumed to have a common corner-frequency across all recordings. This may not be necessarily the case due to directivity and radiation effects, but it actually provides a stable average consistent with the Brune (1970, 1971) instantaneous slip model. In fact, Mayeda & Walter (1996) found that in the case of shear-wave coda, scattering in the Earth’s crust averages out such azimuthal variations. The use of long windows is therefore advantageous in reducing the influence of directivity effects in addition to increasing the reliability of long-period spectral estimation. For the choice of the minimization function, it was found in Edwards et al. (2008) and Edwards & Rietbrock (2009) that the log-space least squares fit significantly reduces the covariance between $f_{c0}, f_c$ (the source corner-frequency), and the signal moment.

In Edwards et al. (2008) and Edwards & Rietbrock (2009), we proceeded to tomographically reconstruct the $Q$ structure using the obtained path-dependent $f_{c0}$ values (which represent path-dependent $Q_{ij}$) to stabilize the inversion and obtain robust determinations of the source corner-frequency and stress-drop. However, it is shown in this study that the initial signal moments (allowing for path variable attenuation and therefore without the need for a priori knowledge of attenuation) can be used to successfully compute $M_w$ from even small earthquakes without significant trade-off or bias. The decision to allow the attenuation to vary for each path also prevents any bias associated with assuming a constant $Q_0$, from influencing the $M_w$ value. This bias was tested for the 2009 January 4 $M_L = 4.1$ event (Table 1) using fixed $Q_0 = 325, 650$ and 1300. This led to a $\Delta M_w$ of $-0.07$ (for $Q_0 = 325$) and $+0.09$ (for $Q_0 = 1300$) from a reference $M_w$ computed with $Q_0 = 650$. However, it should be noted that several authors have previously demonstrated that due to the lack of a priori knowledge of the attenuation, and the strong trade-off of this with the source corner-frequency (or stress-drop), values of $f_{c0}$ and $f_c$ obtained with this unconstrained inversion can be unreliable (e.g. Abercrombie & Leary 1993; Abercrombie 1997; Bay et al. 2003; Sonley & Abercrombie 2006; Allmann & Shearer 2007; Edwards et al. 2008). In this paper, we focus specifically on the determination of the low-frequency spectral level, used to define $M_w$, which was shown to be relatively insensitive to the parameter trade-off using the $M_L = 4.1$ event. Nevertheless, to fully assess its effect, the uncertainty of all the parameters is taken into account. We compute parameter uncertainty following the method of Viegas et al. (2010): from the grid-search around $f_{c0}$, we find those values that are within 5 per cent of the minimum misfit. The corresponding range of signal moment and $f_{c0}$ values are then obtained. The uncertainty in the individual signal moment values are then propagated through the inversion procedure, leading to an error estimate on $M_w$ which takes into account the uncertainty in the spectral deconvolution.

A principal difference in the spectral fitting in this study relative to Edwards & Rietbrock (2009) is the application of a weighting parameter to the data, equal to $1/f$. This gives preference to fitting the long-period spectral amplitudes, similar in effect to a log-space resampling in frequency (e.g. Ide et al. 2003). The misfit function for a particular spectrum (assuming equal variance in the spectral
values) is given by
\[ s_{ij} = \sum_{f_{i}, \text{start}}^{f_{i}, \text{end}} \frac{\log(\Omega_{ij}(f, r)) - \log(\Omega_{ij}^{\text{m}}(f, r))}{\sigma_{ij}}. \]
(12)
where superscripts \( r \) and \( m \) denote spectral values for the real data and the model (eq. 4), respectively. \( f_{i}, \text{start} \) and \( f_{i}, \text{end} \) are the minimum and maximum frequencies that define the selected bandwidth for a particular spectrum, respectively. A Powell’s minimization (Press et al. 1997) is used to solve eq. (12) for each iteration of the grid-search over \( f \). The inverted signal moments that correspond to the \( f \) value leading to the minimum sum of residual misfits, \( s_{ij} \), over all stations, \( j \), of an event, \( i \), are then extracted.

Having successfully extracted the signal moment (the far-field, long-period spectral displacement amplitude), we can proceed with the deconvolution to find the long-period plateau value at the source by accounting for mean site amplifications and the distance decay function (eq. 9). This is solved using a linear L2 SVD decomposition in the log-space. For fast computation of individual \( M_{W} \) in real-time (i.e. with only the recordings of one event available), the \( a \) priori knowledge of \( S_{0}(r, r_{0}, \ldots, r_{n}, \lambda_{1}, \ldots, \lambda_{n}) \) and \( A_{j} \) must be available. In this case it is assumed that \( A_{j} = 1 \) and \( S_{0}(r, r_{0}, \ldots, r_{n}, \lambda_{1}, \ldots, \lambda_{n}) = r^{-1.0} \) for \( r < 150 \) km and \( S_{0}(r, r_{0}, \ldots, r_{n}, \lambda_{1}, \ldots, \lambda_{n}) = 150^{0.5}(r/150)^{-0.5} \) for \( r \geq 150 \) km. The site transfer-term (\( T_{j}(f, A_{j}) \)) was initially taken as that given in Boore & Joyner (1997) for a generic rock-site with \( V_{S30} = 620 \) ms\(^{-1} \). In addition, a homogeneous half-space model was tested, such that \( T_{j}(f, A_{j}) = 1 \). This resulted in negligible differences in the \( M_{W} \) values (an average of +0.02) compared to those computed when accounting for the generic rock-site site transfer-term of Boore & Joyner (1997). For Switzerland, the average \( V_{S30} \) is higher than 620 ms\(^{-1} \), so the true rock amplification lies somewhere in between the Boore & Joyner (1997) generic rock-site and the homogeneous half-space (Poggi et al. 2010). We therefore expect the rock amplification effect to be negligible for computing \( M_{W} \). For simplicity and speed, the half-space model is therefore used. For application in other regions, the average site condition should be assessed and the effect of the site amplification terms considered. However, for NEHRP (BSSC 2003) site classes A and B it is unlikely that any correction is necessary to obtain \( M_{W} \). For a more physically realistic determination of \( M_{W} \), for example including the amplification due to Moho reflections, region specific \( S_{0}(r, r_{0}, \ldots, r_{n}, \lambda_{1}, \ldots, \lambda_{n}) \) and \( T_{j}(A_{j}, f) \) can be derived. The reader should refer to Edwards et al.
B. Edwards et al. (2008) for a full description of the inversion methodology and the robust determination of these parameters.

It follows that seismic moments can be calculated using the following equation (Brune 1970):

\[
M_0 = \frac{\Psi_i \psi r_0 v 4\pi}{\xi F \Theta_1 \Phi_{1i}},
\]

where \(M_0\) is the seismic moment (in SI units), \(\Psi_i\) is the effective long-period plateau displacement at source \(i\) (from eq. 9), \(\Theta_{1s}\) is the average radiation pattern (\(\Theta_{1s} = 0.55\) for \(S\) waves; Aki & Richards 1980; Boore & Boatwright 1984), \(v\) is the shear wave velocity at the source (it is assumed that \(v = 3.5\) km s\(^{-1}\)), \(F\) is the free surface amplification (\(F = 2.0\) for normally incident \(SH\) waves and a good approximation for \(SV\)) and \(\rho\) is the average crustal density (\(\rho = 2800\) kg m\(^{-3}\); \(r_0\) is the fault radius, or 1 km, depending on the definition used in eq. (9). \(\xi\) is an empirically derived factor to account for the difference in the horizontal and vertical ground-motion amplitudes if using vertical recordings (in this case \(\xi = 1\) as we use only horizontal records) (Newmark & Hall 1982). The use of the average radiation effect coefficient, \(\Theta_{1s}\), relies upon a good range of azimuthal coverage when analysing individual phases, at least over one continuous half of the radiation pattern (due to its symmetry). However, the use of multiply scattered waves reduces the effect of the radiation pattern (e.g. Mayeda & Malagnini 2010) such that the azimuthal coverage is less important. The effect of fault orientation was discussed in Boore & Boatwright (1984) where it was shown that for close observations of a vertical strike slip, 30\(^{\circ}\) dip-slip and 45\(^{\circ}\) oblique slip fault the average radiation coefficients for \(S\) waves were 0.50, 0.64 and 0.57, respectively. This would lead to \(\Delta M_W\) of 0.03 for the vertical strike slip event, −0.04 for the dip-slip event and 0.01 for the oblique dip-slip event. A possible extension of the method could be to follow work such as Kissling (1980) and Kissling et al. (1981) to attempt to obtain an estimate of the focal solution for small events and remove the requirement of an average radiation effect, however, it is likely that this would not be compatible with the real-time application presented here.

3 QUALITY EVALUATION OF \(M_W\) FROM SPECTRAL FITTING

3.1 Comparison of \(M_W\)

To test the \(M_W\) computed using the method described in this paper [referred to as \(M_W(SP)\)], a comparison is made with \(M_W\) values from automatic moment tensor solutions computed at the SED \([M_W(MT)]\) (Bernardi et al. 2004; Clinton et al. 2006). 40 earthquakes were selected for comparison, all having the highest category ‘A’ quality for the \(M_W(MT)\) (Table 1). The magnitudes range from \(M_W(MT) = 2.8\) to 5.0. When using broadband full-waveform moment tensor methods, the retrieval of \(M_W(MT)\) for small events (e.g. \(M_W < 3\)) is difficult and only possible when there are nearby recordings and a low level of noise. As a result, moment tensor solutions for small events are not abundant despite their frequent occurrence. The earthquakes used in this comparison are shown in Fig. 2 and Table 1.

The distribution of hypocentral distance and depth of sources with respect to \(M_L\) is shown in Fig. 3. In addition, the range of

![Figure 2. Distribution of earthquakes (circles), instrumentation (squares) and ray-geometry used in the comparison between \(M_W(SP)\) and \(M_W(MT)\).](image-url)
minimum and maximum frequencies used in the spectral inversion for $M_W$ are shown.

The comparison of $M_W$ (MT) with the $M_W$ (SP) derived in this study is shown in Table 1 and Fig. 4. The $M_W$ (SP) were computed using a simple model with $T(f/A) = 1$ and $S(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_{n}) = r^{-1/6}$ for $r < 150$ km and $S(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_{n}) = 150^{1/0.5}(r/150)^{-0.5}$ for $r \geq 150$ km (i.e. a homogeneous half-space with spherical leading to cylindrical geometrical decay). The error in the $M_W$ (SP) value presented in Table 1 has been propagated through the matrix inversion using the estimates of signal moment variance obtained from the spectral inversion. An example of the spectral fit for the horizontal recordings shown in Fig. 1 is given in Fig. 5. Note that the noise estimate has been increased to intersect the signal at low and high frequencies. This ensures a conservative selection of the useful bandwidth, such that noise does not influence the results. The mean difference between the computed $M_W$ is

$$M_W = M_W^{(MT)} + 0.001 \pm 0.063$$

as shown in Fig. 4(a). We found that, as most of the large magnitude events occurred at the borders of Switzerland (Fig. 2), a requirement of $\sigma_m > 30$ (the circular standard deviation of source–station azimuths) removes most of the $M_L > 4$ events. Furthermore, the requirement did not lead to any decrease of the misfit between $M_W$(SP) and $M_W$(MT).

Following Poggi et al. (2010), site transfer functions ($T(f/A)$) were obtained and removed from the data to check for any improvement in variance or introduction of bias in the $M_W$(SP) with respect to $M_W$(MT). In this case however, with generally well recorded events (a minimum of eight stations were available for the smallest events and 29 stations recorded the largest event) the inversion was robust enough that this did not affect the $M_W$(SP). Where few recordings are available, or for strong systematic site effects (e.g. strong-motion sensors located in a sedimentary basin) the prior-removal of $T(f)$ will be necessary for non-biased, robust solutions.

A study of attenuation in Switzerland was performed by Bay et al. (2003, 2005) in which they determined $Q_0 = 270^{0.3}$, and $S_i(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_n) = 50^{-1.1}(r/50)^{-0.3}$ for $50 < r < 70$ km; $S_i(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_n) = 50^{-1.1}(r/70)^{-0.6}$ for $70 < r < 100$ km and $S_i(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_n) = 50^{-1.1}(r/100)^{-0.8}(r/70)^{-0.2}$ for $r \geq 100$ km. This attenuation model includes the amplification of ground motion in the range 70–100 km due to $S_{5r}$ reflections using: $S_i(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_n) = 50^{-1.1}(r/70)^{-0.6}(r/100)^{-0.2}$ for $r \geq 100$ km. The model of Bay et al. (2003, 2005) was used to compute the same $M_W$ values determined using the simple attenuation model described earlier. However, using the Bay et al. (2003, 2005) model, the $M_W$ values are overestimated by an average of 0.1 relative to the $M_W$(MT) (Fig. 4b). Furthermore, the scatter in the difference between $M_W$(MT) and $M_W$(SP) is slightly increased. This could be due to the over-attenuation of low frequencies, or alternatively the incompatibility of the amplitude decay function $S_i(r; r_0, \ldots, r_{n-1}, \lambda_{1}, \ldots, \lambda_n)$ with the moment tensor velocity model. To have consistency between the $M_W$(MT) and $M_W$(SP) it is therefore better to use the simple model of amplitude decay and to allow attenuation to be path variable. In terms of the individual errors assigned to the $M_W$ values, using the Bay et al. (2003, 2005) model significantly reduces these errors. This highlights a problem of using an a priori attenuation model, although we reduce the variance of the inverted parameters, we may introduce bias that is difficult to quantify.

### 4 Application to $M_L$–$M_W$ Magnitude Scaling

#### 4.1 Scaling of $M_L$–$M_W$ in Switzerland

$M_W$ is the preferred homogeneous measure of magnitude for the Swiss earthquake catalogue (Fäh et al. 2003). Although the method shown here is able to remove the requirement for scaling $M_L$–$M_W$ for all but the smallest events, it is useful to have scaling relationship for cases where only $M_L$ is available. Previously a simple scaling relation of

$$M_W = M_L(SED) - 0.2$$

was used at the SED (Braunmiller et al. 2005). This was the result of a regression of 34 events for which both $M_W$(MT) and $M_L$(SED) were available. Due to the necessity of having signal above the noise at long periods (e.g. less than 5 s) for computation of broadband $M_W$(MT), this was limited to the larger events ($M > 3.5$) in the catalogue. This also led to a potential bias in the statistics, as the scaling equation was actually then used outside its valid parameter range: i.e. for $M < 3.5$.

A comparison of $M_W$ and $M_L$ of European earthquakes has been made by of Grünthal et al. (2009). They showed that

$$M_W = 0.53 + 0.646 M_L + 0.0376 M_L^2$$

However, this is based on $M_L$ values from a number of catalogues of different seismological agencies, which, as previously mentioned, scale differently. It should, nevertheless represent an average European scaling relation and provide a worthwhile comparison to a unique $M_L$(SED) to $M_W$ scaling relation.
The new relation computed in this study is based on $M_w(SP)$ computed using the spectral fitting technique described previously. The quantity of data (679 events) for this regression was substantially larger than the original regression of Braunmiller et al. (2005). The minimum magnitude was $M_L = 1.3$. The data were binned into groups of $M_L$ 1–2, 2–3, 3–4 and 4–5.5. A bootstrap technique was used with 1000 subsets comprising of 80 per cent of the full data set, and 20 per cent random repetitions, although enforcing an equal number of data in each $M_L$ bin. This was to evaluate the robustness of the solution and avoid bias due to data selection and magnitude. The regression was an orthogonal L1 solution assuming equal variance in both magnitudes. From the 1000 possible scaling relations given by the bootstrap analysis, a final regression was performed to find the orthogonal L1 solution of the bootstrap possibilities. The new relation is given by

$$M_w = 1.02 + 0.472 M_L(SED) + 0.0491 M_L(SED)^2 \pm 0.15.$$  

(17)

The range of fits found by the bootstrap approach is shown in Fig. 6. The scaling law of Grünthal et al. (2009) (eq. 16) is included for comparison. The scaling relation in this study matches that of Grünthal et al. (2009) well for magnitudes greater than $M_L = 3.5$. For lower magnitudes, the scaling relation of this study gives higher $M_w$ for respective $M_L$ values.

It is important to note that the validity of the derived relation does not extend beyond the limits for which data were available: specifically for $M_L > 1.3$ and $M_L < 5.3$ as there is no theoretical basis for a purely quadratic relation between $M_L$ and $M_w$. This apparent in Fig. 6, where it is shown that the possibilities for the scaling relation grow significantly for $M_L < 1.3$ and $M_L > 5.3$.

4.2 Testing the scaling relation

To test the results of the regression, the $M_w(MT)$ computed by the automatic moment tensor solutions of Clinton et al. (2006) are compared with the values predicted by both the old and new scaling relations. These $M_w(MT)$ are completely independent from the $M_w(SP)$ derived for computing the new scaling relation [only $M_w(SP)$ were used] and therefore can provide an unbiased assessment of any residual misfit dependency. The results are shown in Fig. 7. Both the old and new relations tend to underpredict $M_w(MT)$ of events smaller than $M_L = 3.5–3.75$. However, this is due to the fact that, for such earthquakes, it is only possible to obtain high-quality broadband moment tensor solutions for events with particularly high signal-to-noise ratios at nearby recording stations. As a result, $M_w$ values for events with $M_L < 3.75$ tend only to be available for when their $M_w$ is particularly high. Effectively, the

© 2010 The Authors, GJI, 183, 407–420
Geophysical Journal International © 2010 RAS
Computation and scaling of moment magnitudes

Figure 5. Spectral fit plotted over the normalized input velocity spectra (spectra divided by the peak spectral velocity) for the horizontal components of recording for the events shown in Fig. 1. The top panel shows the north–south and bottom panel the east–west orientations. The adjusted noise estimates are included with the signal data (see text for a description of the noise adjustment). The vertical dotted lines indicate the limits between which the automatic processing selected data for inclusion in the spectral inversion. The vertical dash-dot lines indicate the range of possible $f_c$ based on misfit within 5 per cent of the minimum.

Figure 6. Top panel: plot of all events used to define the scaling relation. The $M_L$ error bars indicate the standard deviation in $M_L$ ($\sigma_{M_L}$) as given by the earthquake catalogue of Switzerland (0.15 for $1 \leq M_L < 2$ and 0.1 for $M_L \geq 2$). The $M_W$–$M_L$ error bars are given by $\sqrt{\sigma_{M_L}^2 + \sigma_{M_W}^2}$. Bottom left: plot of possible polynomial fits found during the bootstrap process to identify the optimal $M_L$–$M_W$ scaling relationship. Bottom right: density plot of the events used in the bootstrap procedure. The solid lines indicate the best model given by eq. (17), the dash-dot line indicates the old linear scaling relation (eq. 15) while the dashed lines indicate the scaling relation of Grünthal et al. (2009).

5 $M_L$–$M_W$ SCALING SIMULATIONS

In the remaining sections, we attempt to understand the reason for the shape of the observed $M_L$–$M_W$ scaling relation. Building upon the theoretical framework (e.g. Hanks & Boore 1984; Ben-Zion & Zhu 2002), Deichmann (2006) presented an argument for the scaling of $M_L$–$M_W$, where it was shown that, provided the self-similarity of earthquake sources, the scaling of $M_L$–$M_W$ should be 1:1. The observation that the theoretical 1:1 relationship breaks down in empirical relationships was explored by Hanks & Boore (1984). They postulated that a complex interaction of the Wood–Anderson high-pass filter, the far-field source corner-frequency (of which stress-drop is a proxy; Brune 1970) and the $f_{\text{max}}$ (Hanks 1982), requires a polynomial relation to describe the scaling between $M_L$ and $M_W$. Ben-Zion & Zhu (2002) proposed that an alternative cause of the magnitude of statistical completeness for the $M_W$ catalogue of Clinton (personal communication) is around $M_L = 4$ (as a conservative estimate) and statistical interpretation should only be made above this magnitude. For larger earthquakes, there does not seem to be any dependency of the residual misfit on magnitude for either the linear or quadratic relations. Overall there is negligible difference in the $M_W$ residual between the old and new relations for events of $M_L > 3.5$, although the quadratic scaling relation tends to predict higher $M_W$ for $M_L > 4$. The most significant difference between the two scaling relations is at low magnitudes, where it is not possible to independently verify the predictions with $M_W(M_T)$.
deviation from 1:1 scaling could result from non-self-similarity of earthquake sources. They suggested that a change in the effective stress regime could lead to the observed polynomial scaling between $M_L$ and $M_W$: for small events the heterogeneity of a tectonic stress-field may lead to slip that does not grow substantially with rupture dimension, whereas for larger events the effect of small scale stress heterogeneities is smoothed out. They suggest that the evolution between the two extremes of small and large events could lead to the observed polynomial form of $M_L - M_W$ scaling (e.g. Millet et al. 1999).

To improve our understanding of the influence of attenuation and stress-drop on the scaling of $M_L - M_W$, scaling relationships were simulated through the generation of stochastic waveforms. Following the method of Boore (2003), accelerograms can be synthesized by convolving the Fourier acceleration spectrum of an earthquake with Gaussian noise shaped with a tapering function. Furthermore, by applying the resulting acceleration time-series to a damped simple harmonic oscillator with properties of the Wood–Anderson Seismometer (damping factor equal to 0.69, natural period of 0.8 s, gain of 2800), the peak Wood–Anderson amplitude can be obtained. Note that to maintain consistency, the properties used to simulate the Wood–Anderson Seismometer response are chosen to match those used for the computation of $M_L$ (SED) rather than the empirically derived response (e.g. Uhrhammer & Collins 1990). The Fourier acceleration spectrum required for the stochastic simulation can be reconstructed using eqs (3), (4) and (13), requiring only the $M_W$, stress-drop, attenuation, site amplification and duration of shaking, $T_d$. The source corner-frequency is obtained from the stress-drop using (Boore 2003)

$$f_c = 0.4906 \beta \left( \frac{\Delta \sigma}{M_0} \right)^{1/3}$$

which is based on Brune (1970, 1971) and Eschelby (1957). Attenuation is parameterized in terms of the crustal attenuation, $Q$, the site attenuation $\kappa$ (synonymous to the $f_{max}$; Hanks 1982) and geometrical spreading.

### 5.1 Bay et al. (2003, 2005) attenuation and source scaling model

In a first step, a synthesized $M_W$ versus $M_L$ scaling relation is produced based on the geometrical spreading, $Q$, stress-drop and $\kappa$ given by Bay et al. (2003). These are $S_0(r) = 10^{-0.6} r^{-0.6}$ for $10 \leq r < 70$ km; $S_0(r) = 10^{-0.6} r^{-0.6}$ for $70 \leq r < 100$ km and $S_0(r) = 10^{-0.6} r^{-0.6}$ for $r \geq 100$ km (in the form of eq. 7); constant $\Delta \sigma = 0.5 - 1.0$ MPa; $Q(f) = 270 f^{0.5}$ and $\kappa = 0.015$ s. Site amplification is set to $T(f, A) = 1$ as the model is referenced to an undefined average site condition and the duration of shaking is detailed in Bay et al. (2003). $M_L$ values are computed over a range of distances (10, 20, 30, 50, 100 and 200 km) to prevent bias in the computed $M_L$. The synthesized $M_L$ versus $M_W$ are shown in Fig. 8. It can be seen that $M_L$ synthesized using the model of Bay et al. (2003) do not follow the empirical curve. This is consistent with Bay et al.'s (2005) findings where it was shown that ground motion tended to be underpredicted using the model with constant stress-drop. In Bay et al. (2005), it was proposed that an increasing stress-drop with seismic moment model, such that

$$\Delta \sigma = 0.79 \times 10^{0.3(M_W - 4.2)} \text{ MPa},$$

along with $\kappa = 0.0125$ s may lead to better estimates of ground-motion at higher magnitudes. Despite the change, the resulting synthesized scaling relation (Fig. 8) still does not adequately match the observed relation. The inadequacy of the Bay et al. (2003) and Bay et al. (2005) models in synthesizing the observed scaling relation leads us to explore why this is the case. Using a grid-search it was found that the observed scaling relation could be adequately replicated by a synthetic relation (Fig. 8). However, the uncertainty of the model was unacceptably high. In the following section the

---

**Figure 7.** Comparison of $M_W$ (MT) computed by Clinton et al. (2006) and those predicted using (squares) the new quadratic and old linear scaling relation (diamonds).

**Figure 8.** Synthetic $M_L - M_W$ scaling relations based on Bay et al. (2003, 2005) studies along with the adjusted Bay et al. (2005) gridsearch results. $M_L$ are computed for a range of distances, resulting in the observed scatter. Third degree polynomials are fit to the data to highlight the trend. The observed $M_L - M_W$ scaling relation is included for comparison.
sensitivity of the $M_L-M_W$ scaling relation to the source and attenuation parameters is therefore explored instead.

### 5.2 Sensitivity of $M_L-M_W$ scaling to $\Delta\sigma$ and $Q$

To reconstruct the curvilinear form of the observed $M_L-M_W$ scaling relationship (eq. 17), a non-constant change of $\Delta \sigma$ (the change in the amplitude of the synthesized Wood–Anderson Seismogram) with respect to $M_W$ is required. A simple sensitivity test showed that the shape of the scaling relationship was relatively insensitive to the duration of shaking, $T_d$; doubling the duration of shaking resulted in a decrease of $M_L$ by around 0.1 for all $M_W$. As $T_d$ has been defined empirically (Bay et al. 2003) this was therefore fixed in the simulations. Geometrical spreading, by definition, is independent of $M_W$ and frequency in the case of a point-source simulation. Spherical geometrical spreading in the first 150 km, followed by cylindrical spreading is assumed, consistent with the method for obtaining $M_W$(SP) applied earlier. Site effects (excluding non-linear effects) are independent of magnitude, but are frequency dependent therefore may affect the shape of the $M_L-M_W$ scaling relationship. The generic rock site amplification function of Boore & Joyner (1997) was tested in the simulations. The site amplification term resulted in a small change in shape of the $M_L-M_W$ scaling relationship, with an increase of 0.3 in $M_L$ for $M_W \approx 1$, dropping linearly down to an increase of 0.2 in $M_L$ for $M_W \approx 6$. Site effects or reference site conditions (Steidl et al. 1996) are however, usually available (e.g. Fäh et al. 2003b), hence the parameter is fixed in the simulations, as described in the text.

The effect of the site attenuation term $\kappa$ is analogous to that of stress-drop $\Delta \sigma$ (or source corner-frequency $f_c$), in that it acts as a low-pass filter, independent of the propagation distance. Due to their similar effects, it is non-trivial to separate $\Delta \sigma$ and $\kappa$ (e.g. Boore et al. 1992), particularly for small magnitude events. We therefore assume a value of $\kappa$ for the simulations based on other studies. However, it should be noted that increasing the stress-drop is analogous to decreasing the $\kappa$ value. Furthermore, high $\kappa$ values can give the impression of an increasing stress-drop with magnitude for small events.

### 5.3 Effect of $\Delta \sigma$ on $M_L-M_W$ scaling

The effect of the stress-drop on the $M_W-M_L$ scaling relation is shown in Fig. 9. Intrinsic attenuation is not applied in the simulation (specifically $Q = 10^4$ and $\kappa = 0.0$) to analyse only the effect of $\Delta \sigma$. As described by Hanks & Boore (1984), it can be seen that the interaction of the stress-drop and the Wood–Anderson filter plays a strong role in the deviation of the scaling from a linear form. In fact, the stress-drop defines at which point the $M_L$ scale begins to saturate. For $\Delta \sigma$ of 1 MPa, the upper bound of the Bay et al. (2003) model, the saturation begins to take effect (i.e. the turning point of the $M_W-M_L$ versus $M_L$ scaling plot in Fig. 9) as early as $M_L = 4$. On the other hand, a $\Delta \sigma$ of 10 MPa delays the start of saturation of $M_L$ to around $M_L = 5$. Extrapolating to larger magnitudes, the 5 MPa stress-drop model is roughly consistent with the maximum $M_L$ value of $\sim 7$ suggested by Hanks & Kanamori (1979).

### 5.4 Effect of $Q$ on $M_L-M_W$ scaling

Published values of $Q$ vary significantly from region-to-region (e.g. Bay et al. 2003; Drouet et al. 2005, 2008; Edwards et al. 2008; Atkinson & Boore 1995; Malagnini et al. 2000). One reason for this is the strong covariance of the parameters in eq. (4). To analyse the effect of $Q$ on $M_L$, it is first assumed for simplicity that the high-frequency $\omega^2$ decay of earthquake source spectra (Brune 1970) is not present within the frequency band of recording. The effect of $Q$ alone on the synthesized $M_L$ values can therefore be analysed. To do this, a sufficiently high stress-drop is chosen for the simulations ($\Delta \sigma = 1000$ MPa). The resulting $M_L$ and $M_W$ are shown in Fig. 10(a). For increasing attenuation (decreasing $Q$), the synthesized $M_L$ values decrease. The effect of changing $Q$ is the same over all magnitudes. This is expected, as $Q$ and, in this case, the source spectral shape (flat in displacement due to the high stress-drop), are both independent of magnitude. In other words, the shape of the displacement spectra at a particular distance from the source is identical for all $M_W$: only the amplitude differs. In the absence of source spectral decay, intrinsic attenuation is therefore indistinguishable from geometrical decay in terms of its effect on $M_L$. Finally, it is interesting to note that in the absence of source spectral decay within the recording bandwidth, $M_L$ scales exactly proportional to $1.5M_W$, or alternatively to $\log(M_W)$ rather than 1:1 with $M_W$ (Fig. 10a).

The introduction of a realistic stress-drop (e.g. $\Delta \sigma = 5$ MPa; Allmann & Shearer 2009) appears to diminish the effect of $Q$ at high magnitudes (Fig. 10b). The use of the realistic stress-drop, and hence a source spectrum with high-frequency $\omega^2$ decay within the frequency band of recording, introduces a magnitude

---

**Figure 9**. Synthetic $M_L-M_W$ scaling relations for different $\Delta \sigma$ values without attenuation. $M_L$ are computed for a range of distances, resulting in the observed scatter for each $M_L$. Third degree polynomials are fit to the data to highlight the trend. The observed $M_L-M_W$ scaling relation is included for comparison.
dependency into the effect of $Q$ on the peak Wood–Anderson amplitude. At low magnitudes, the effect of $Q$ is dominant: the $o^2$ decay of the source spectrum begins at a sufficiently high frequency that it does not affect the synthesized peak Wood–Anderson amplitude (as Fig. 10a). However, for increasingly larger magnitudes, the effect of $Q$ becomes less significant as the $o^2$ source decay begins at lower frequencies: for a given stress-drop, the corner-frequency of the $o^2$ source decay is inversely proportional to the cube of the seismic moment, Brune (1970). As a result of this magnitude dependency, the $Q$ value can change the slope of the $M_{L}-M_w$ relationship.

5.5 Impact of the $M_{L}-M_w$ relationship on source scaling

The observed scaling of $M_{L}-M_w$ can be shown to be due to a combination of the earthquake source spectra (defined by $\Delta\sigma$), the attenuation (defined by the $Q$ value and $\kappa$) and the effect of the Wood–Anderson high pass filter. The $Q$ value controls the shape of the scaling relation between $M_{L}$ and $M_w$ at low magnitudes, while the stress-drop controls the saturation of the $M_{L}$ scale at high magnitudes. The comparison of the synthetically produced $M_{L}-M_w$ relation with the observed relation produced for Switzerland can provide an insight into the regional average stress-drop and crustal attenuation. For instance, the observed scaling in this case clearly does not allow for a stress-drop as low as 0.5–1 MPa, as found by Bay et al. (2003), throughout the magnitude range (Fig. 8). Furthermore, it suggests that the modification made in Bay et al. (2005) is insufficient and the increase of stress-drop with magnitude needs to be stronger still. The variability of stress-drop with magnitude would be consistent with the postulation of Ben-Zion & Zhu (2002) that increasing fault dimensions smooth out the heterogeneity of the tectonic stress-field. The comparison of the observed and synthesized $M_{L}-M_w$ scaling relation does give us some insight into source and path effects. However, the resolution of source scaling properties is notoriously poor and a detailed analysis in the context of the $M_{L}-M_w$ scaling relation goes beyond the scope of this paper.

6 SUMMARY AND CONCLUSIONS

It has been shown that the $M_w$ computed using a moment tensor solution based on broadband waveform fitting of local Swiss earthquakes of $M_w = 2.8–5.0$. The standard deviation of the difference between the two solutions was only 0.06. This error is consistent with the expected error from different moment tensor solutions, for example due to the use of different velocity models, locations, source depths or stations used. Hence we are confident of the magnitude similarity between the spectral and broadband waveform fitting methods. The method was integrated with the automatic picking and location procedure at the SED, and hence provides real-time moment magnitude determination with a delay of about five minutes after automatic waveform pick and earthquake location information are available. Although this is not critical for small events, fast magnitude determination is still important in cases of moderate earthquakes (e.g. $M = 4.5–6$) such that the extent of strong ground shaking can be rapidly assessed. Results are stored in an XML database and are then available on the web in real-time by using an XSL stylesheet. Manual updates to the earthquake location and phase picks are also automatically reprocessed and included when available.

Using the newly computed $M_w(SP)$ a scaling relation between $M_{L}(SED)$ and $M_w$ was determined. A second degree polynomial was fit to the data and a bootstrap-L1 regression used to robustly determine the scaling law. This law was then applied to $M_{L}$ for which $M_w(MT)$ were available from independent broadband moment tensor solutions. The old linear scaling law, defined at larger magnitudes ($M_{L} > 3.5$), was shown to underpredict the smaller broadband moment tensor $M_w(MT)$. The new relation was shown to reduce the magnitude dependence of the residual between predicted and moment tensor $M_w(MT)$. The difference between the two scaling relations is most significant for small magnitude events. For instance, for $M_{L} = 2.0$, the linear scaling relation predicts $M_w = 1.8$, whereas the quadratic scaling relation defined in this study predicts $M_w = 2.16$. This difference increases for decreasing magnitude. The form of the new relation is similar to a regression of European earthquakes with $M_{L}$ and $M_w$ by Grunthal & Wahlström (2004) and Grunthal et al. (2009), although for $M_{L} < 3.5$ the relation presented in this study predicts slightly higher $M_w$.

Theoretically the method detailed in this article should allow determination of $M_w$ on low-cost strong motion networks, as entire spectra are modelled (so high-frequency information can be utilized), and the respective site transfer function (which can be strong

Figure 10. Synthetic $M_{L}-M_w$ scaling relations for different $Q$ values (a) without influence from the stress-drop (b) with a stress-drop of 5 MPa. $M_{L}$ are computed for a range of distances, resulting in the observed scatter. Third degree polynomials are fit to the data to highlight the trend—note that in (a) the best fit is linear. The observed $M_{L}-M_w$ scaling relation is included for comparison.
for strong-motion installations) could be removed before processing. This would be useful where broadband data does not exist. However, further tests will be undertaken to assess the robustness of the method in such situations.

The use of a quadratic term in scaling $M_L-M_W$ can cause problems when the magnitude of interest is outside the range of data used in the derivation of the scaling relation. Further work is to be undertaken to incorporate even smaller magnitudes in addition to larger magnitudes into the regression: hence avoiding potential artefacts if the quadratic scaling does not continue to lower or higher magnitudes. This revision of the scaling relation will then be used to obtain a homogeneous moment magnitude scale for the forthcoming earthquake catalogue of Switzerland (ECOS09).

Synthesized $M_L$ values were computed for a range of $M_W$ using a stochastic modelling technique. It was shown how the stress-drop and the $Q$ value interact to affect the $M_L-M_W$ scaling relation. The shape of the observed scaling relation could be very closely approximated by varying these parameters. The $Q$ value mainly affected low magnitude scaling, while the stress-drop affected high magnitude scaling, specifically the saturation of the $M_L$ scale due to the high-pass Wood–Anderson filter. It was noted that the trade-off between $\kappa$ and stress-drop meant that the interpretation of increasing stress-drop with magnitude could be substituted with the use of a high $\kappa$ value. However, $\kappa$ can be determined from large events which do not exhibit such trade-off (e.g. Anderson & Hough 1984), so $\kappa$ was fixed based on similar studies for Switzerland. In the absence of $\omega^2$ source spectral decay within the frequency band of interest (i.e. for very small events), $M_L$ scales proportionately to log($M_0$) or $1.5M_W$. At some point, defined by the average stress-drop and therefore the lower frequency limit of the $\omega^2$ spectral decay, this scaling reduces to 1:1 with $M_W$, followed by the saturation of the $M_W$ scale at higher magnitudes.

ACKNOWLEDGMENTS

Some figures in this paper were produced using GMT 4 (generic mapping tools) software written by Wessel & Smith (1991, 1998). Parts of the code used in creating the inversion tools were provided by Pozo (2004) and Press et al. (1997). Thanks to Nicholas Deichmann, Falko Bethmann, Yehuda Ben-Zion, Rachel Abercrombie and the anonymous reviewers for their helpful comments. This work was partly funded by swissnuclear within the Pegasos Refinement Project (Renault et al. 2010).

Moment magnitudes of Swiss earthquakes were taken from the online database at: http://www.seismo.ethz.ch/mt/RTMT/RTMT_SWISS.php. Only ‘class A’ events were used. All data used in this study are available via AutoDRM (see http://www.seismo.ethz.ch/autodrm/). Station maps can be obtained from http://www.seismo.ethz.ch/div佥ks/plots/plots.html.

REFERENCES


