Strong ground-motion prediction equations from induced earthquakes in St. Gallen geothermal field, Switzerland

Nitin Sharma [iD][1],[*], **Vincenzo Convertito**[2], **Raffaella De Matteis**[3] **and Paolo Capuano**[4]

[1] CSIR-National Geophysical Research Institute, Hyderabad - 500068, India

[2] Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Via Diocleziano 328, 80124 Napoli, Italy

[3] Dipartimento di Scienze e Tecnologie, Università degli Studi del Sannio, Via de Sanctis, 82100, Benevento, Italy

[4] Dipartimento di Fisica "E. R. Caianiello", Università di Salerno, 84084, Fisciano (SA), Italy

*Corresponding author: Nitin Sharma. E-mail: withnitin@gmail.com

Abstract

Ground shaking, whether it is due to natural or induced earthquakes, has always been a matter of concern since it correlates with structural/non-structural damage and can culminate in human anxiety. Industrial activities such as water injection, gas sequestration and waste fluid disposals, promote induced seismicity and consequent ground shaking that can hinder ongoing activities. Therefore, keeping in mind the importance of timely evaluation of a seismic hazard and its mitigation for societal benefits, the present study proposes specifically designed ground-motion prediction equations (GMPEs) from induced earthquakes in the St. Gallen geothermal area, Switzerland. The data analysed in this study consist of 343 earthquakes with magnitude $-1.17 \leq M_{L,\,corr} \leq 3.5$ and hypocentral distance between 4 and 15 km. The proposed study is one of the first to incorporate ground motions from negative magnitude earthquakes for the development of GMPEs. The GMPEs are inferred with a two-phase approach. In the first phase, a reference model is obtained by considering the effect of source and medium properties on the ground motion. In the second phase the final model is obtained by including a site/station effect. The comparison between the GMPEs obtained in the present study with GMPEs developed for the other induced seismicity environments highlights a mismatch that is ascribed to differences in regional seismic environment and local site conditions of the respective regions. This suggests that, when dealing with induced earthquakes, GMPEs specific for the study should be inferred and used for both monitoring purposes and seismic hazard analyses.

Keywords: induced seismicity, strong ground motion, seismic hazard, geothermal, ground-motion prediction equations

1. Introduction

Understanding strong ground motion from earthquakes is a key aspect for seismic hazard analyses. The idea becomes more interesting when ground motion from induced or triggered earthquakes changes the existing hazard scenarios, which are predominantly estimated from tectonic activity only. Anthropogenic activities, such as fluid injection for geothermal energy production, gas sequestrations, wastewater disposal, enhanced oil recoveries and hydraulic
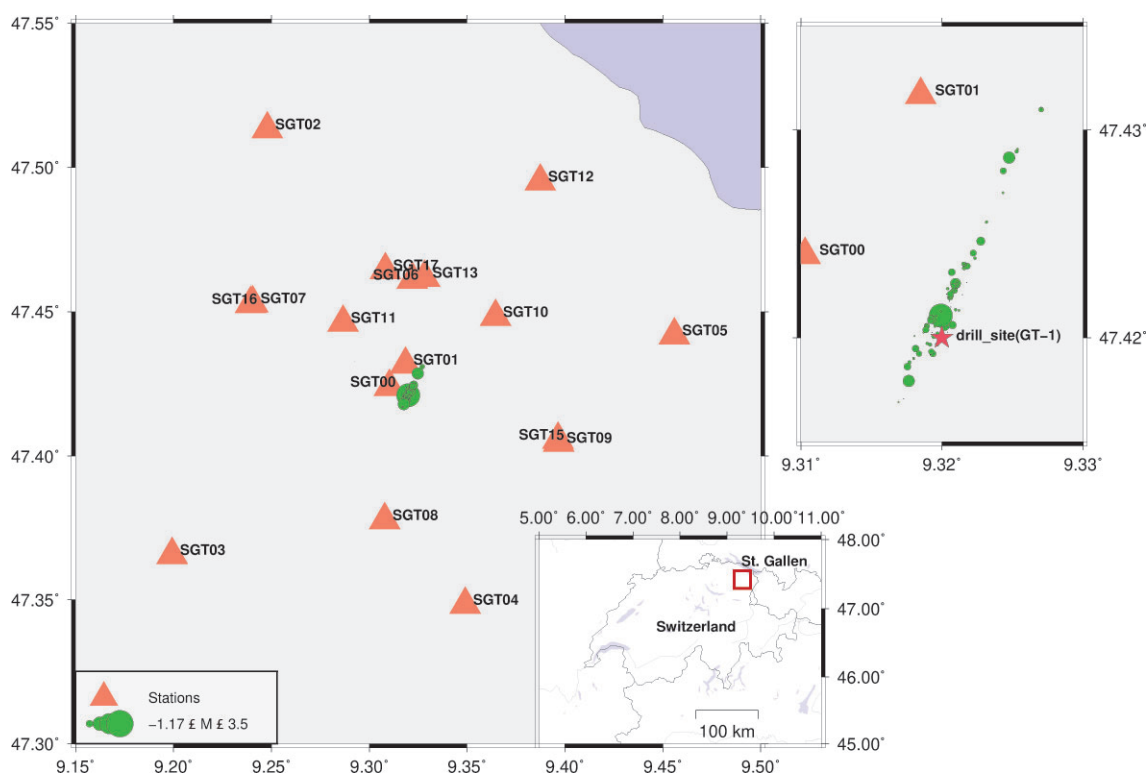
Figure 1. The distribution of seismic stations of the Swiss Seismological Service (SED) and the seismic sequence occurred in 2013 used for the analysis in the present work. The location of study region is shown in the red square in the inset. The right panel shows the zoom-in section of the sequence.

fracking, have opened a new avenue of earthquake hazards from man-made earthquakes. Many studies such as Bao & Eaton (2016) and Ruiz-Barajas, *et al.* (2017) have revealed that the stress changes during fluid injection activities may activate fault slips, which results in periodic seismic activities that may last for months. There are many examples of injection-induced seismicity reported worldwide, such as the Geysers, Basel, Hengill, Roswinkel, Soultz and Voerendaa Paralana, KiGam at Pohang, Eastern coast of Spain, Western Canada, Oklahoma in USA (see Majer & McEvilly 1979; Jousset & François 2006; Häring *et al.* 2008; Bao & Eaton 2016; Barajas, *et al.* 2017; Lu 2017; Schultz, *et al.* 2020), to demonstrate this. The race to obtain clean energy is forcing the exploitation of geothermal regions around the globe. It is proving a powerful source when compared with traditional renewable sources of energy (water, solar, wind, etc.). Thus, geothermal systems emerge as a preferred solution to meet demand for clean energy supplies. Inspired by the success of commercial deep hydrothermal projects in Munich, a geothermal project at the St. Gallen hydrothermal reservoir located in Switzerland was initiated (Diehl *et al.* (2017) and references therein). The region lies over a Mesozoic sediments deposit in the Molasse basin of northeast Switzerland and was characterized by very low seismicity in recent years. Only one earthquake with $M_L \geq 3.0$ occurred in the proxim-

ity of the area of interest and no other significant earthquake was reported from 1984 until 2012. The detailed knowledge of the St. Gallen geothermal system and the origin of the sequence of induced earthquakes and their societal impact have been described by Moeck *et al.* (2015), Wolfgramm *et al.* (2015) and Obermann *et al.* (2015).

Therefore, motivated by these discussed facts, in the present study, we decided to develop the ground-motion prediction equations (GMPEs) because there are no specific ones available for the St. Gallen area. Studies such as time-dependent analysis of GMPEs and seismic hazards might prove helpful in monitoring and controlling the effects of the seismicity rate levels on inhabitants living in surrounding areas and on structures as well (Convertito *et al.* 2012; Convertito *et al.* 2021).

The GMPEs are developed by analysing the 2013 seismic sequence comprising 343 earthquakes with magnitude $-1.17 \leq M_{L,corr} \leq 3.5$ and hypocentral distance range between 4 to 15 km. The sequence was recorded by 16 surface seismic stations along with one station in a borehole (see figure 1). The distribution of hypocentral distances versus magnitude and depth versus magnitude of the events is shown in figure 2.

Nearly 16,524 waveforms have been processed to measure the peak-ground accelerations (PGA) and peak-ground velocities (PGV) from all the events at all stations. We use
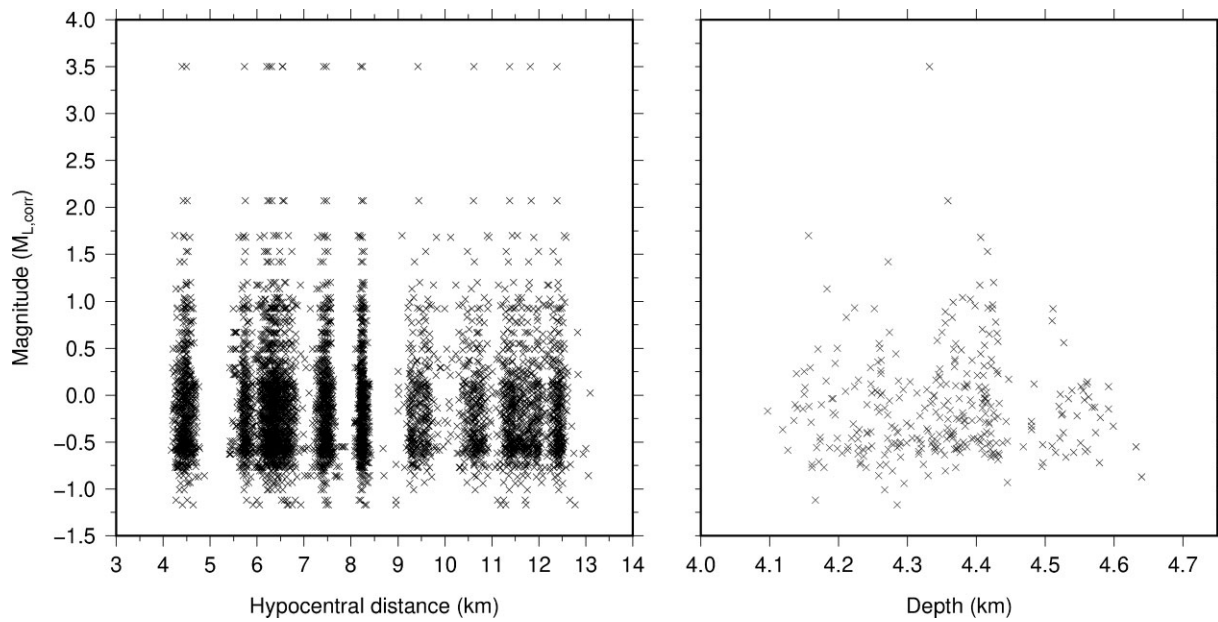
Figure 2. The scatter plot showing distribution of (a) magnitude of earthquakes versus hypocentral distance and (b) magnitude versus depth.

the nonlinear mixed-effects regression technique to infer the GMPEs (Lindstrom & Bates 1990; Abrahamson & Youngs 1992). We consider the GMPE as a mixed model in which the total uncertainty is a combination of both aleatory (random) and epistemic (informative) uncertainties, which are further reflected into inter- and intra-event uncertainties. The latter are considered to be independent, normal distributed variables with their own variances. The inter- and intra-event components of the total uncertainty represent the earthquake-to-earthquake variability and the variability among observations within a single event, respectively (see Abrahamson & Silva 1997; Atik *et al.* 2010; Baltay *et al.* 2017). The mixed model allows us to understand the respective contribution of aleatory and epistemic uncertainties, which can be accounted individually to improve the ground-motion models. In this work, apart from developing the GMPEs, we also demonstrated how the addition of medium properties and site information reduces epistemic uncertainties and results in an overall reduction of the total uncertainty in the final models.

We estimated the GMPEs from induced seismicity records by implementing the two-phase approach as proposed by Emolo *et al.* (2011, 2015) and Sharma *et al.* (2013), Sharma and Convertito (2018). In the first phase, a reference model is inferred by considering only the effect of the source (through the magnitude) and propagation medium (through the geometric attenuation). The model developed during phase I is called a reference model. Then, the corresponding ground motions are predicted for each magnitude and distance combination available in the catalogue using a reference model. The residuals at each station are obtained

through the difference between the observed and predicted PGA and PGV values, respectively. The residuals carry the information that are not captured by the reference model. Thus, the modal residual value is considered to be corrective factor and in phase II the final corrected GMPE is obtained through a new regression by including additional model parameters accounting for anelastic attenuations and site effects. The modal values of the residual distribution at each station are again obtained and compared with those of the reference model to evaluate the improvement in the final model and robustness of the adopted technique. The results show that the maxima of residuals for most of the stations are centred at zero and the total standard deviations are significantly reduced when compared with the reference model. This two-phase method is found to be very effective when there is little or no information available about the local recording sites, e.g. Vs_{30}. The obtained GMPEs are also compared with the GMPEs obtained from other induced seismic regions.

2. Data processing

Nearly 343 induced earthquakes occurred in the St. Gallen geothermal region during the fluid injection process in July 2013. The seismicity was monitored by the Swiss National Seismic Network operated by the SED and a local seismic network whose configuration has been modified during the period of interest for the present study. Overall, data from 17 surface three-component stations, which include broadband (SGT01-SGT05) equipped with Nanometrics Trillium Compact 120s, short period (SGT06–SGT16) and one borehole station SGT00 installed at the depth of 205 m
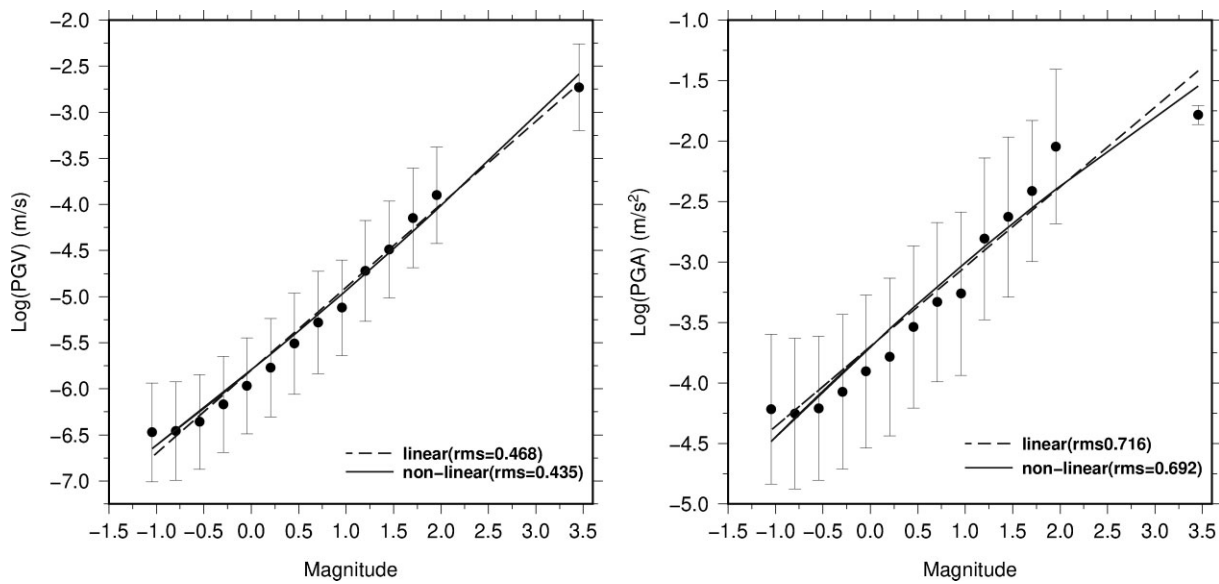
Figure 3. The nonlinear trend observed between magnitude and corresponding ground-motion parameters (a) PGV and (b) PGA. The circle represents the binned average value of ground-motion parameters (within magnitude bin size = 0.25) and deviation from the mean value is shown by vertical bars. The linear and nonlinear fits are also shown with relative RMS values.

Table 1. Coefficients and associated uncertainty of the reference model (equation 3) obtained in phase I

GM	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_b$	$d \pm \sigma_d$	τ (inter-event)	Φ(intra-event)	σ total	R^2
PGV	-1.5893 ± 0.0600	0.7824 ± 0.0200	0.0718 ± 0.0141	-4.3708 ± 0.0591	0.1383	0.3230	0.3514	0.7552
PGA	1.0326 ± 0.0793	0.5743 ± 0.0214	0.1034 ± 0.0151	-4.9358 ± 0.0786	0.1223	0.4305	0.4476	0.6377

and equipped with an OYO Geospace HS-1LT sensor (2.0–28 Hz) were used (see figure 1). Details about the configuration of the network and its temporal evolution can be found in Diehl *et al.* (2017) and Edwards *et al.* (2015). In total, 16,524 waveforms from induced earthquakes are processed by first removing mean and trend. Using the response files, the instrumental response is removed from the waveforms that are finally filtered with Butterworth band pass filter in the frequency range 2–90 Hz. We chose this range to consider the corner frequencies of even small magnitude earthquakes. PGA and PGV are obtained in the same frequency range and correspond to the maximum between the two horizontal components. To measure the correct values of ground motion we cut the waveforms in a selected time window starting at a time corresponding to 0.05% and ending at 95% of the total energy contained in the waveform. The windowed waveforms are also tapered at a 0.1 taper width with a cosine window. The magnitude $M_{L,corr}$ of the earthquakes, range between -1.17 and 3.5 with the hypocentral distances between 4 and 15 km. The catalogue information is obtained from Swiss National Seismic Network operated by the SED and contains all the metadata information such as station locations and hypocentre parameters of the events. The magnitude of the earthquake was provided on a $M_{L, corr}$ magnitude scale (see Edwards *et al.* 2015).

3. Methodology

In general, the functional form of empirical GMPE is expressed as:

$$\log Y = f(M) + f(R) + f(\text{site}) + \xi(\eta, \varepsilon) \tag{1}$$

where Y is the ground-motion value, $f(M)$ is a function of the magnitude M, $f(R)$ is a function of the source-to-site distance R and $f(\text{site})$ accounts for possible site effect. The variability in ground-motion model, which is a function of distance, magnitude and site, is expressed through the function ξ. This latter can include two independent components, namely the inter-event variability (η) and the intra-event variability (ε). The equation can be expressed both as natural or common logarithms (here, common log is used to perform the comparison with similar existing models). Once the formulation for each of the functions listed in equation (1) is selected, the $\log Y$ will be related to the predictive parameters M, R and site through a set of coefficients that are inferred by using the available data. In particular, in the present study model parameters are obtained through a nonlinear mixed-effects regression technique (Lindstrom & Bates 1990; Abrahamson & Youngs 1992). The adopted regression technique allows the total variability of the ground motion to be partitioned into two terms (see equation 2): the inter-event
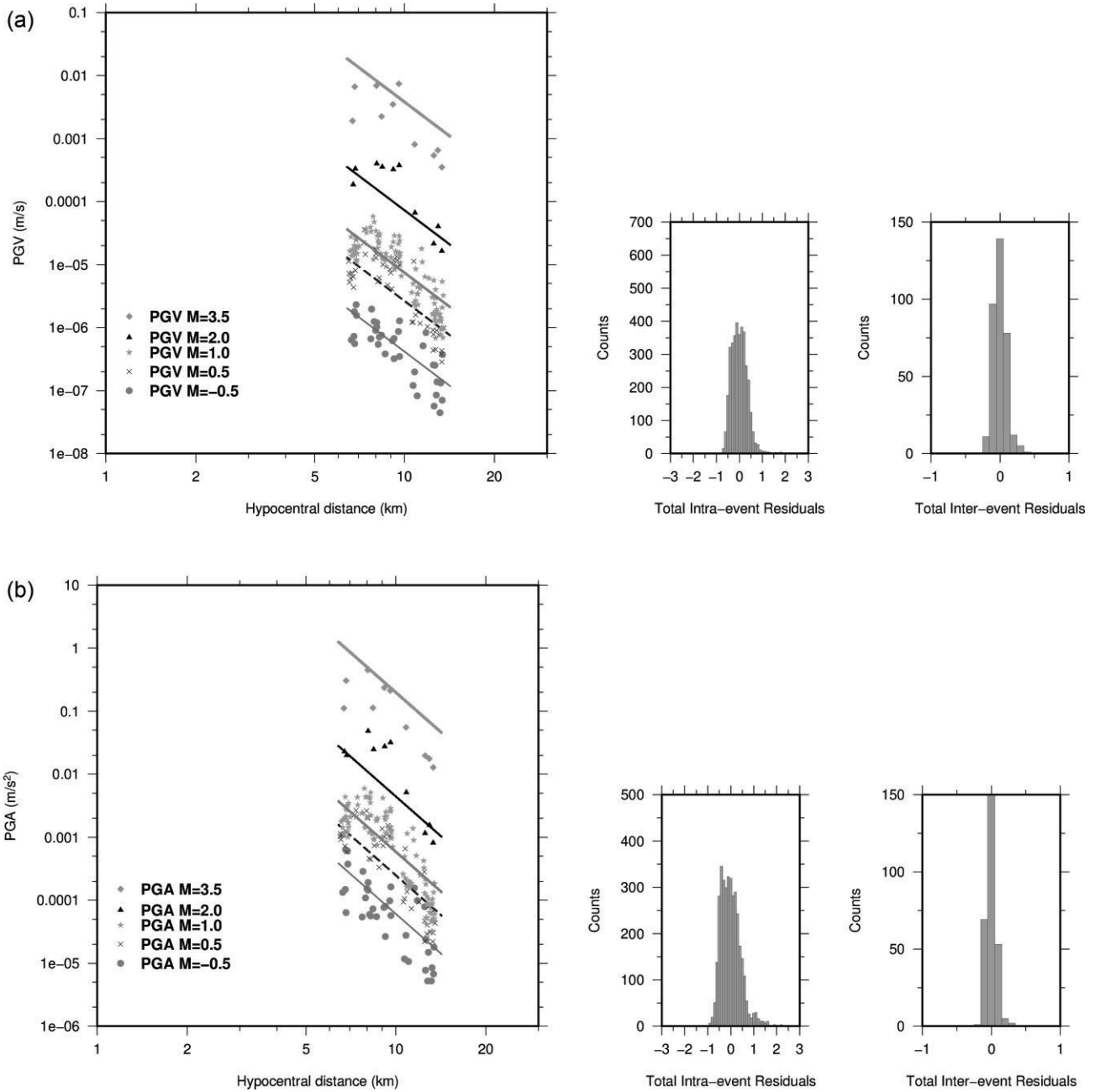
Figure 4. Reference model obtained for magnitudes $M_{L, corr}$ = −0.5, 0.5, 1.0, 2.0 and 3.5 with respect to hypocentral distance. The inter- and intra-event residual distributions for each ground-motion parameter (PGV, PGA) are also depicted in the right side of panel.

variability (η) and the intra-event variability (ε). The η and ε are assumed to be independent, normally distributed variables with standard errors τ and Φ, respectively. Therefore, the total standard deviation of the ground-motion model σ_{Total} can be expressed as:

$$\sigma_{Total} = \sqrt{\tau^2 + \Phi^2} \qquad (2)$$

The nonlinear mixed-effects regression technique is an optimization iterative procedure that requires an initial model. The two-phase algorithm is used to estimate the random effects, the variances and the model parameters (Lindstrom & Bates 1990). The algorithm to calculate inter-

and intra-event variability is described in work done by Abrahamson & Youngs (1992). The distribution of variability allows to understand the contributions from source and medium along with their respective aleatory and epistemic components.

4. Ground-motion prediction equation (GMPE)

4.1. Development of the reference GMPEs phase I

To develop GMPEs from induced seismicity records from St. Gallen sequence, in phase I we infer a reference model (see Sharma *et al.* 2013; Emolo *et al.* 2015; Sharma & Convertito
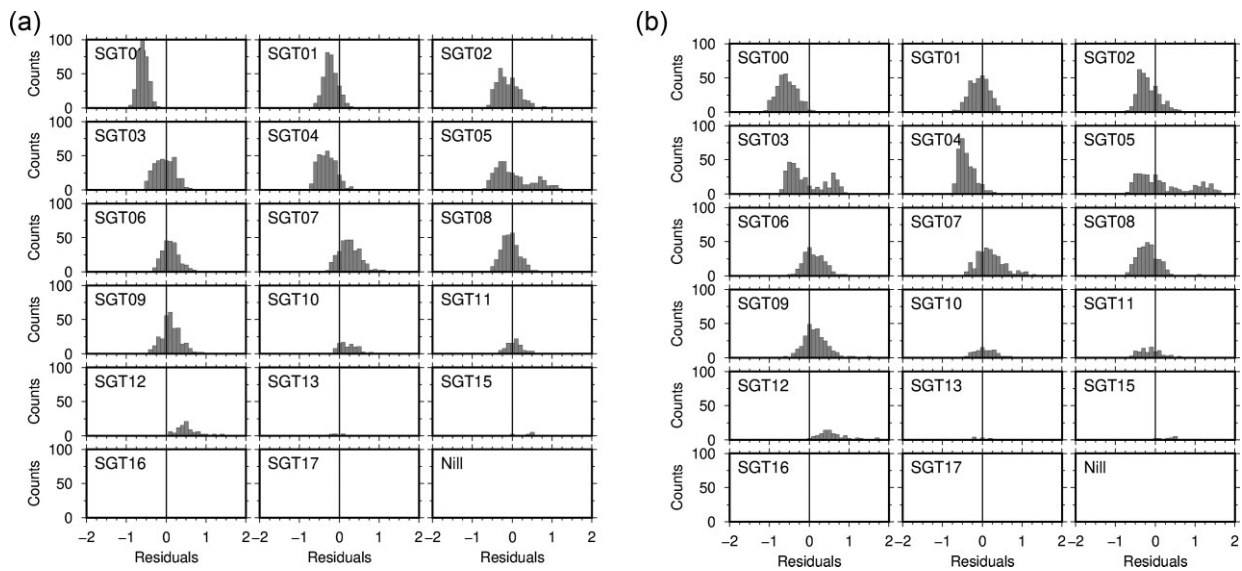
Figure 5. Residual histograms at each station for reference model for PGV (a) and PGA (b). It can be observed from the distribution of residuals that most of the stations need correction because the maxima of residual at most stations is not centred at zero. This shows that station/site correction is not accounted in reference model.

2018). This model accounts for the source and geometrical attenuation effects only and is formulated as:

$$\log Y = a + b \cdot M + c \cdot M^2 + d \cdot \log\sqrt{(R^2 + depth^2)} \tag{3}$$

where Y is the ground motion, that is, PGV (ms^{-1}) and PGA (ms^{-2}), whereas the magnitude is measured on a $M_{L, corr}$ magnitude scale and epicentral distance R and depth of the events are measured in kilometres.

The coefficients b and c are associated with the magnitude account for the effect of the size of earthquake on ground motions, whereas d accounts for the decay in terms of geometrical spreading.

The quadratic dependency on the magnitude is chosen according to the minimum of the Akaike information criterion (AIC) (Akaike 1974). In particular, we compare the RMS of the data fitting for the two formulations: $f(M) = a + b*M$ and $g(M) = a + b*M + c*M^2$ (figure 3). The AIC value for the linear fit (i.e. for $f(M)$) for PGV and PGA is 2936.7 and 4537.8 and for the quadratic fit (i.e. $g(M)$) is 2917.4 and 4502.4, respectively. The lower AIC values indicate a better-fit model for $g(M)$, which is then used for the magnitude dependence. Since earthquake depths vary in a very narrow range between 4.0 and 4.7 km, they are co-located with respect to the source-to-site distances. Douglas & Jousset (2011) suggested that PGA does not depend on stress drop for very small earthquakes since it only affects spectral amplitudes at frequencies higher than the corner frequency that for small events may be highly attenuated by site attenuation. If this is correct, the observed quadratic dependency can be explained only if the site attenuation shows a nonlinear behaviour (the larger

the magnitude the larger the peak-ground motion), which, however, seems to be not very plausible for the peak-ground motion values measured in the present study. Thus, for the St. Gallen earthquakes, we favour a difference in the stress drop between the events as a more plausible mechanism.

The regression coefficients (a, b, c and d), together with their associated uncertainty and the inter- and intra-event residuals, are then obtained through a nonlinear mixed-effects regression technique. The results are listed in Table 1. Figure 4 shows the comparison between the data and the reference model for five different magnitude values ($M_{L, corr}$ = −0.5, 0.5, 1.0, 2.0 and 3.5) in terms of PGV and PGA. In the same figure, we display the histograms for the inter- and intra-event residuals. Note that GMPEs slightly over predict the ground motions for a magnitude of 2 and above. The reason is that there are only two events with a magnitude $2.0 < M_{L, corr} \leq 3.5$ in the data set. The main contribution to the residuals comes from the intra-event component (Φ) that is larger than the inter-event component (τ) for both PGV and PGA. Moreover, the intra- and inter-event residual distributions peak around zero, and the overall dispersion among residuals appears to be small, even in the case when there is noticeable dispersion in the ground-motion data.

4.2. Station residual analysis phase I

The reference model developed here considers the effect of source and elastic attenuation (geometrical spreading) on the ground motion. Residuals analysis allows to understand the robustness of model and the appropriateness of the parameters, which are considered while inferring the

Table 2. The model parameters along with associated errors obtained after regression during phase II for the final models (equations (4) and (5))

GM	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$d \pm \sigma_d$	$e \pm \sigma_e$	$s \pm \sigma_s$	τ (inter-event)	Φ (intra-event)	σ total	R^2
PGV	-3.3442 ± 0.3066	0.8017 ± 0.0170	0.0644 ± 0.0120	-0.5529 ± 0.5445	-0.1522 ± 0.0241	0.8898 ± 0.0239	0.1224	0.2676	0.2943	0.8203
PGA	1.7654 ± 0.0743	0.6116 ± 0.0182	0.0907 ± 0.0128	-4.9284 ± 0.0645	N/A	1.1140 ± 0.0313	0.1060	0.3794	0.3940	0.7083

Table 3. Coefficients used for site correction to obtain final model

Station code	Corrective parameters for PGA (mu)	Corrective parameters for PGV (mu)
SGT00	−1.1880	−1.1390
SGT01	−0.7315	−0.6090
SGT02	−0.6648	−0.6294
SGT03	−0.7293	−0.6210
SGT04	−0.8148	−0.7302
SGT05	−0.6834	−0.6899
SGT06	−0.4992	−0.3624
SGT07	−0.5466	−0.4261
SGT08	−0.8032	−0.6417
SGT09	−0.6001	−0.4495
SGT10	−0.4918	−0.3971
SGT11	−0.6592	−0.3543
SGT12	−0.0147	−0.0177
SGT13	−0.3669	−0.2059
SGT15	−0.0448	−0.0573
SGT16	−0.1979	0.0718
SGT17	−0.5989	−0.3248

reference model. The residuals ($\log Y^{obs} - \log Y^{pred}$) distribution at each station is shown in figure 5a and b. It can be observed that most stations need a correction because the maximum of residuals at most of them is not centred at zero. This shift means that the reference model is either under or over predicting the ground motions. Thus, the parameters considered to develop the reference model are not sufficient.

Therefore, other components, i.e. local geology and site effects, beneath the recording sites are subsequently included into the model. Local site conditions play a significant role in the amplification and/or attenuation of ground motion immediately before striking the free surface. Due to a lack of local geological information, such as, for instance, the average V_{S30} (i.e. the average value of shear-wave velocity in the upper 30 m of the crust) at recording sites, we implemented the technique proposed by Emolo *et al.* (2011) and then successfully implemented by Sharma *et al.* (2013), Sharma and Convertito (2018) to account for a first-order station specific site effect.

To apply site correction, the modal residual value is obtained at each station and is used as a corrective factor (see table 3). Thus, in addition to the anelastic attenuation, the proposed corrective coefficient allowed us to include effects of local site conditions in the final model, which were not included in the reference model developed in phase I. It should be noted that the station effect considered here is not solely based on average V_{S30} values but includes, more generally, all the effects concurring to a systematic site amplification and attenuation effect.

Further, note that we did not consider the waveforms from borehole stations, that is, SGT00, during the phase I development of the GMPE (reference model). The reason is that
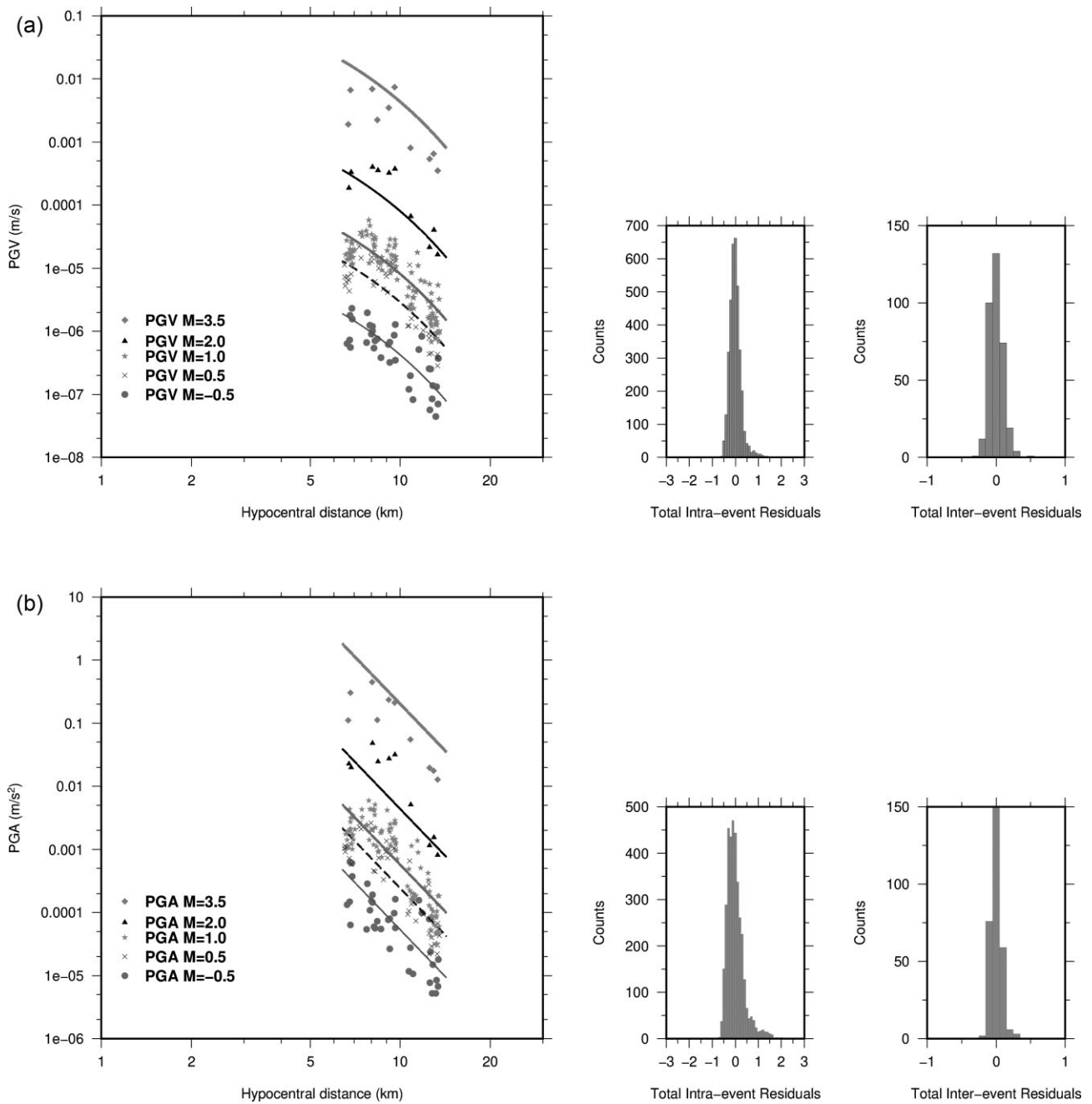
Figure 6. The final model obtained after station/site correction for magnitudes −0.5, 0.5, 1.0, 2.0 and 3.5 with respect to hypocentral distance. The inter- and intra-event residual distributions for each ground-motion parameter (PGV, PGA) are also depicted on the right side of panel.

the recording station SGT00 is underground and it might not capture the local site effects.

4.3. Development of final GMPEs phase II

During phase II, the dependency of ground motion on anelastic property of the medium and contribution of site effects on peak-ground motions are considered. The terms accounting for anelastic attenuation and the local site effect are added explicitly in the reference model to obtain the

following final model:

$$\log PGV = a + b \cdot M + c \cdot M^2 + d$$

$$\cdot \log \sqrt{(R^2 + \text{depth}^2)} + eR + mu \cdot s \quad (4)$$

In equation (4), *e* accounts for the anelastic attenuation, which represents loss of energy per unit cycle during seismic wave propagation and *s* accounts for the effect of local site conditions on ground-motion recordings.
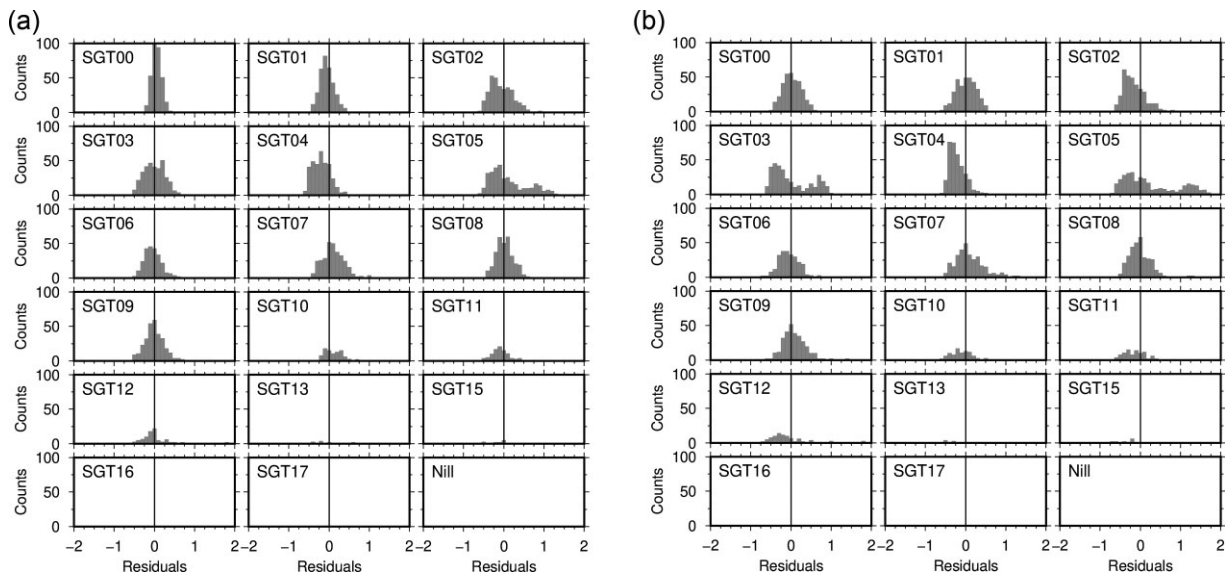
Figure 7. Residual histograms at each station for final model for PGV (a) and PGA (b). It can be observed from the distribution of residuals that the site/station correction works for most stations because the maxima of residuals at most of the station are now centred at zero. This shows that after station/site correction the model is improved.

The obtained coefficients together with their respective uncertainties are shown in Table 2.

It should be noted that data recorded at the borehole station SGT00 are also used to obtain the final GMPE. It is worth mentioning that the total standard deviation is reduced for the final GMPEs (see Table 2). There is significant reduction in intra-event component of standard deviation for both PGVs and PGAs. The improvement in the fit is demonstrated by the increase in the R-squared values for the final GMPEs as compared with the reference model. However, it is observed that PGA data do not allow to separately resolve the anelastic attenuation contribution and the geometric attenuation. Indeed, we obtain a positive *e*-coefficient that cannot be physically interpreted. This can be probably due to the relatively higher frequency content of PGA with respect to PGV. Thus, the final model for PGA is as follows:

$$\log(\text{PGA}) = a + b \cdot M + c \cdot M^2 + d$$
$$\cdot \log\sqrt{(R^2 + \text{depth}^2)} + mu \cdot s \quad (5)$$

The plot of the final model for both PGV and PGA along with the histograms of intra-event and inter-event residuals is shown in figure 6.

4.4. Station residuals analysis phase II

The residuals at each station are again calculated by following the procedure mentioned previously for the phase I and reported in figure 7a and b. The final model obtained after necessary medium and site corrections shows a better fit to the observed data. It is clearly evident that the maximum of

residuals shifts towards zero for most stations. This signifies that the final model now better explains the observed ground motions showing the effectiveness of the adopted two-phase method to obtain the GMPE in the absence of average V$_{S30}$ to account for local site effects.

4.5. Total residual distribution

The residual trend as function of the hypocentral distance and magnitude is also analysed for both the reference and final model respectively (see figure 8a and b). In general, no clear trend is observed for both PGA and PGV. We binned the residuals for both PGA and PGV at a bin size of 0.25 (Sturges 1926), both in terms of magnitude and hypocentral distance to demonstrate the improvement in the fit of the model. The binned residuals values show a clear reduction in the scatter particularly when considered as a function of the hypocentral distance.

5. Comparison with other GMPEs

The final model developed for the St. Gallen geothermal region is compared with the GMPEs obtained specifically for induced earthquakes occurred in other geothermal regions. The latest GMPEs proposed by Sharma & Convertito (2018) for the Geysers and GMPEs developed by Douglas *et al.* (2013) using mixed data from six different geothermal areas including the Geysers in USA, Basel, Hengill, Roswinkel, Soultz and Voerendaal in Europe, respectively, are compared with the GMPEs developed in the present study. The comparison for different magnitude ranges is shown in
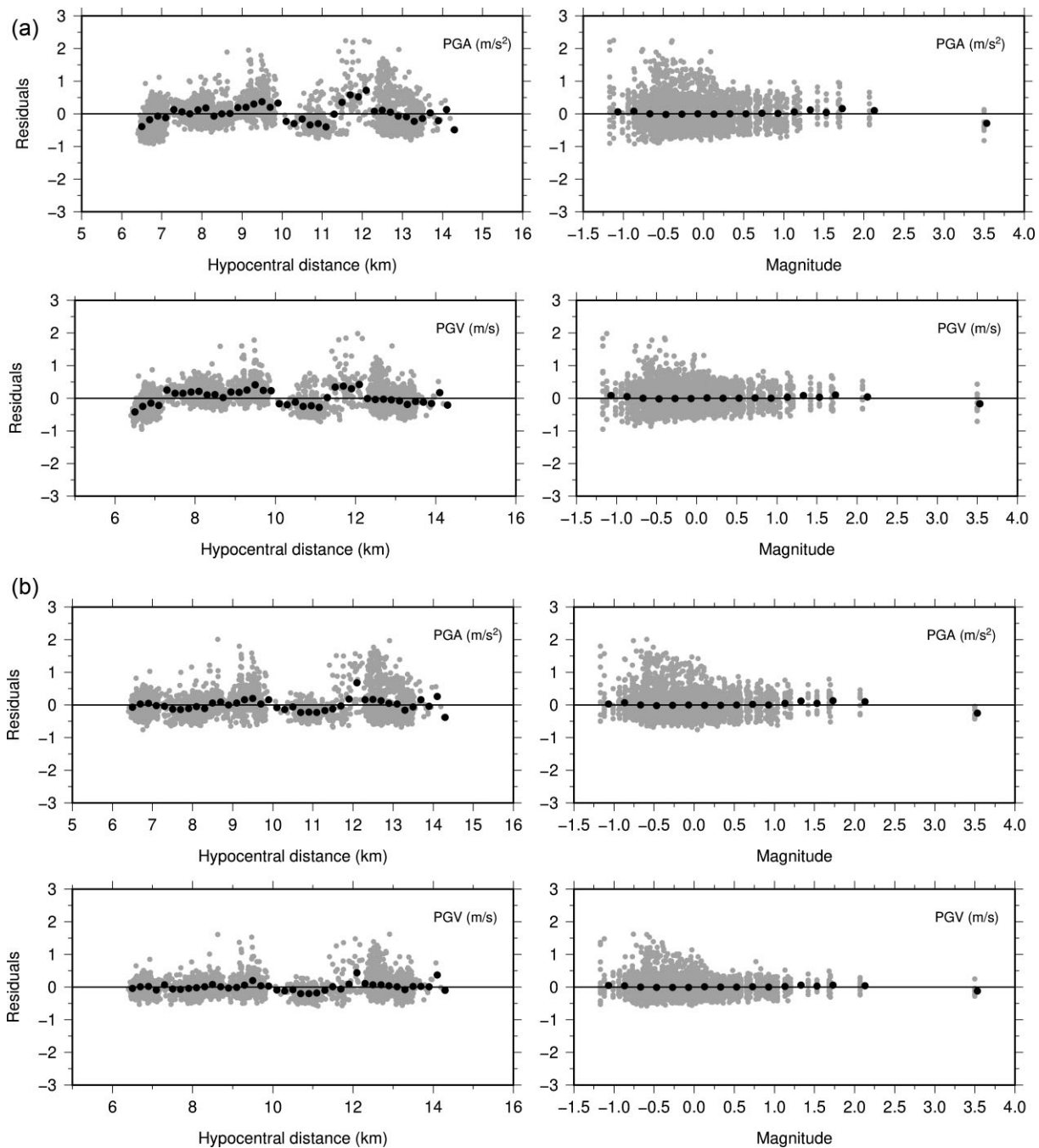
Figure 8. Distribution of residuals (logY^{obs} − logY^{pred}) obtained from reference model (upper panel) and final model (lower panel) after station/site correction as function of the hypocentral distance and magnitude of events for PGV (a) and PGA (b) respectively. The binned average value of residuals is represented by black circles. It can be clearly observed that after the station/site correction the residuals are sifted towards zero.

figure 9a and b. The results indicate that the GMPEs developed for other geothermal areas under predict the data from the St. Gallen geothermal area. There might be several reasons to explain the mismatch among the GMPEs. One may be related to the different magnitude range and, in particular, to the number of larger magnitude events contained in the different datasets, which can affect the final regression. A second reason can be related to either the seismogenesis or the

local site conditions of the respective regions. Moreover, if we consider the case of the model developed for the Geysers (Sharma & Convertito 2018) and Douglas *et al.* (2013), most of the earthquakes used for analysis are at shallow depths. The events in case of St. Gallen region are confined to particular source layer (between 4 and 4.7 km). This might be the reason why the PGA and PGV values faintly vary with depth and have higher values at the St. Gallen region. This
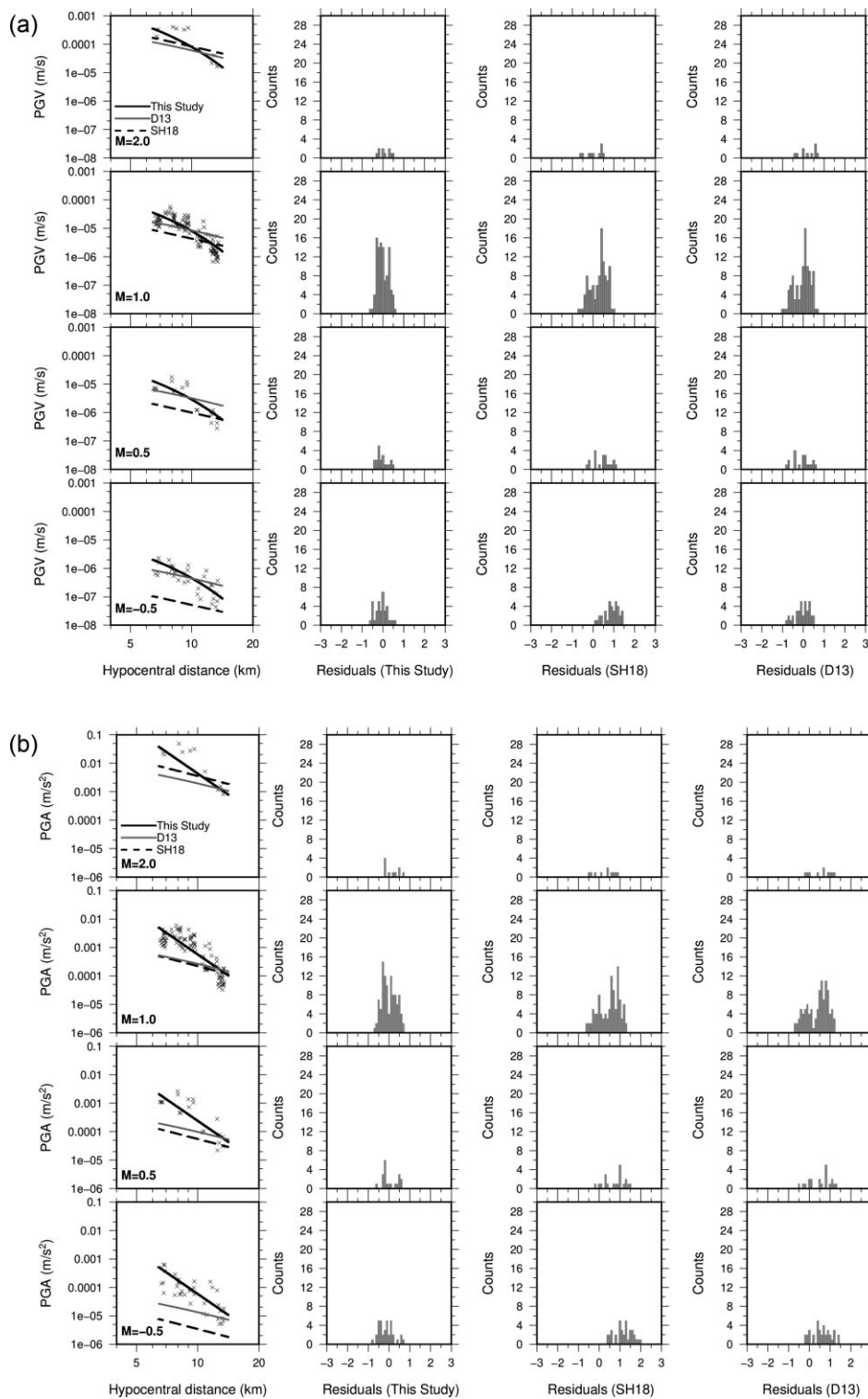
Figure 9. Comparison of GMPEs (solid-line, black) developed for the St. Gallen region with the GMPEs proposed by Sharma & Convertito (2018), SH18 (dashed-line), Douglas *et al.* (2013) (D13) (solid-line, grey), for PGV (a) and PGA (b). The comparison is shown fitting for magnitudes $M_{L, corr}$ = −0.5, 0.5, 1.0 and 2.0.

fact is further supported by Douglas *et al.* (2013), where they clearly demonstrated that the ground-motion differs at different depths even if the magnitude and the hypocentral distances are kept the same for the earthquakes.

6. Conclusions

In this work, we present the first GMPEs for the St. Gallen geothermal region. The GMPEs are developed by using the

2013 seismic sequence comprising 343 earthquakes with magnitudes $-1.17 \leq M_{L,corr} \leq 3.5$, hypocentral distances between 4 and 15 km and depths confined between 4 and 4.7 km. We used the nonlinear mixed-effects regression technique to update the GMPEs (Lindstrom & Bates 1990; Abrahamson & Youngs 1992). We consider the GMPE to be a mixed model in which the total uncertainty is accounted as combination of both aleatory (random) and epistemic (informative) uncertainties, which are further reflected into inter- and the intra-event uncertainties. The GMPEs are developed for PGA and PGV. We estimated the GMPEs coefficients by implementing a two-phase approach. During the first phase, a reference model is developed by considering the effect of the source (magnitude) and propagating medium (in terms of elastic attenuation) only. The inferred model represents the reference model. At each station, we analyse the residual distribution, which is calculated as difference between the observed and predicted PGA and PGV values, respectively. The modal values of the residual distributions are considered to be corrective factors and in the second phase the final corrected GMPEs are obtained through regression with additional model parameters accounting for anelastic attenuations and site effects. We demonstrate that the addition of medium properties and site information reduces the epistemic component on intra-event uncertainties with a significant reduction in total uncertainties in the final GMPEs. This two-phase method is found to be very effective when there is a lack of or no information available about the local recording sites such as the Vs_{30} values. The obtained GMPEs are finally compared with the GMPEs obtained from the analysis of induced earthquakes recorded in other regions. In particular, we consider the models proposed by Sharma & Convertito (2018) and by Douglas *et al.* (2013). We verified that the GMPEs developed for other geothermal areas under predict the data from the St. Gallen geothermal area. This mismatch could be associated with the regional seismic environment and the local site conditions of the respective regions, which suggests that, when dealing with induced earthquakes, GMPEs specific for the study should be inferred and used for both monitoring purposes and seismic hazard analyses.

Availability of data and materials

The catalogue and waveform data are available from the Swiss National Seismic Network operated by the SED (Edwards *et al.* 2015; Diehl *et al.* 2017), last accessed during April 2021. The link to website is: http://www.seismo.ethz.ch/en/earthquakes/monitoring/.

Seismic Analysis code was used to perform all the basic waveform processing. The Generic Mapping tool (Wessel & Smith 1991) was used to generate figures.

References

Abrahamson, N.A. & Silva, W.J., 1997. Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismological Research Letters,* **68**, 94–127.

Abrahamson, N.A. & Youngs, R.R., 1992. A stable algorithm for regression analysis using the random effect model, *Bulletin of the Seismological Society of America,* **82**, 505–510.

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **6**, 716–723.

Atik, A.L., Abrahamson, N.A., Bommer, J.J., Scherbaum, F., Cotton, F. & Kuehn, N., 2010. The variability of ground-motion prediction models and its components, *Seismological Research Letters,* **81**, 794–801.

Baltay, A.S., Hanks, T.C. & Abrahamson, N.A., 2017. Uncertainty, variability and earthquake physics in ground-motion prediction equations, *Bulletin of the Seismological Society of America,* **107**, doi:10.1785/0120160164.

Bao, X.W. & Eaton, D.W., 2016. Fault activation by hydraulic fracturing in western Canada, *Science*, **354**, 1406–1409.

Convertito, V., Ebrahimian, H., Amoroso, O., Jalayer, F., De Matteis, R. & Capuano, P., 2021. Time-dependent seismic hazard analysis for induced seismicity: the case of St Gallen (Switzerland), geothermal field, *Energies*, **14**, 2747.

Convertito, V., Maercklin, N., Sharma, N. & Zollo, A., 2012. From induced seismicity to direct time-dependent seismic hazard, *Bulletin of the Seismological Society of America*, **102**, 2563–2573.

Diehl, T., Kraft, T., Kissling, E. & Wiemer, S., 2017. The induced earthquake sequence related to the St. Gallen deep geothermal project (Switzerland): fault reactivation and fluid interactions imaged by microseismicity, *Journal of Geophysical Research: Solid Earth*, **122**, 7272–7290.

Douglas, J., Edwards, B., Convertito, V., Sharma, N., Tramelli, A., Kraaijpoel, D., Cabrera, B.M., Maercklin, N. & Troise, C., 2013. Predicting ground motion from induced earthquakes in geothermal areas, *Bulletin of the Seismological Society of America,* **103**, 1875–1897.

Douglas, J. & Jousset, P., 2011. Modeling the difference in ground motion magnitude-scaling in small and large earthquakes, *Seismological Research Letters,* **82**, 504–508.

Edwards, B., Kraft, T., Cauzzi, C., Kästli, P. & Wiemer, S., 2015. Seismic monitoring and analysis of deep geothermal projects in St Gallen and Basel, Switzerland, *Geophysical Journal International,* **201**, 1022–1039.

Emolo, A., Convertito, V. & Cantore, L., 2011. Ground-motion predictive equations for low-magnitude earthquakes in the Campania-Lucania area, southern Italy, *Journal of Geophysics and Engineering,* **8**, 46–60.

Emolo, A., Sharma, N., Festa, G., Zollo, A., Convertito, V., Park, J.H., Chi, H.C. & Lim, I.S., 2015. Ground-motion prediction equations for South Korea Peninsula, *Bulletin of the Seismological Society of America,* **105**, 2625–2640.

Häring, M.O., Schanz, U., Ladner, F. & Dyer, B.C., 2008. Characterisation of the Basel 1 enhanced geothermal system, *Geothermics,* **37**, 469–495.

Jousset, P. & François, B., 2006. Set-up of a broadband seismological network at Hengill geothermal field. BRGM report RP-54971-FR, 51.

Lindstrom, M.J. & Bates, D.M., 1990. Nonlinear mixed-effects models for repeated measures data, *Biometrics* **46**, 673–687.

Lu, S.-M., 2017. A global review of enhanced geothermal system (EGS), *Renewable and Sustainable Energy Reviews,* **81**, 2902–2921.

Majer, M.L. & McEvilly, T.V., 1979. Seismological investigations at The Geysers geothermal field, *Geophysics,* **44**, 246–269.

Moeck, I., Bloch, T., Graf, R., Heuberger, S., Kuhn, P., Naef, H., Sonderegger, M., Uhlig, S. & Wolfgramm, M., 2015. The St. Gallen project: development of fault controlled geothermal systems in urban areas, In *Proceedings, World Geothermal Congress 2015, Melbourne, Australia,* 19–25 April.

Obermann, A., Kraft, T., Larose, E. & Wiemer, S., 2015. Potential of ambient seismic noise techniques to monitor the St. Gallen geothermal site (Switzerland), *Journal of Geophysical Research: Solid Earth,* **120**, 4301–4316.

Ruiz-Barajas, S., Sharma, N., Convertito, V., Zollo, A. & Benito, B., 2017. Temporal evolution of a seismic sequence induced by a gas injection in the Eastern coast of Spain. *Scientific Reports,* **7**, 1–15.

Schultz, R., Skoumal, R.J., Brudzinski, M.R., Eaton, D., Baptie, B. & Ellsworth, W., 2020. Hydraulic fracturing-induced seismicity. *Reviews of Geophysics,* **58**, https://doi.org/10.1029/2019RG000695.

Sharma, N. & Convertito, V., 2018. Update, comparison, and interpretation of the ground-motion prediction equation for "The Geysers" geothermal area in the light of new data. *Bulletin of the Seismological Society of America,* **108**, 3645–3655.

Sharma, N., Convertito, V., Maercklin, N. & Zollo, A., 2013. Ground motion prediction equation for The Geysers geothermal area based on induced seismicity records, *Bulletin of the Seismological Society of America,* **103**, 117–130.

Sturges, H.A., 1926. The choice of a class interval. *Journal of the American Statistical Association,* **21**, 65–66.

Wessel, P. & Smith, W.H.F., 1991. Free software helps map and display data, *EOS Transactions American Geophysical Union,* **72**, 445–446.

Wolfgramm, M., Bloch, T.M., Bartels, J., Heuberger, S., Kuhn, P., Naef, H., Seibt, P.J., Sonderegger, M., Steiger, T.H., Uhlig, S. & Stadtwerke, S.G., 2015. Reservoir-Geological Characterization of a Fractured Limestone: Results Obtained from the Geothermal Well St. Gallen GT-1 (Switzerland), Proceedings of World Geothermal Congress.