

Bulletin of the Seismological Society of America

This copy is for distribution only by
the authors of the article and their institutions
in accordance with the Open Access Policy of the
Seismological Society of America.

For more information see the publications section
of the SSA website at www.seismosoc.org



THE SEISMOLOGICAL SOCIETY OF AMERICA
400 Evelyn Ave., Suite 201
Albany, CA 94706-1375
(510) 525-5474; FAX (510) 525-7204
www.seismosoc.org

Single-Station Sigma for Italian Strong-Motion Stations

by L. Luzi, D. Bindi, R. Puglia, F. Pacor, and A. Oth

Abstract A fundamental problem for site-specific ground-motion prediction, commonly required in seismic-hazard assessment, lies in the fact that ground-motion observations over long enough time periods are unavailable at the vast majority of sites. For this reason, most of the ground-motion prediction equations have been derived using observed data from multiple stations and seismic sources, and the standard deviation (sigma) is related to the statistics of the spatial variability of ground motion instead of temporal variability at a single site (ergodic assumption). In this paper, we explore the variability at single sites, decomposing sigma into different parts so that the various contributions to the variability can be identified and the standard deviation for empirical ground-motion prediction models quantified by removing the ergodic assumption. The analysis was conducted using three different data sets.

Sigma obtained for Italy using the ergodic assumption is about $0.35 \log 10$ units (Bindi, Pacor, *et al.*, 2011) and decreases to about 0.3 when single stations are considered (15% reduction). The values of single-station sigma obtained in this study for multiple-source data sets are rather stable, in the range 0.18 – $0.2 \log 10$ units, comparable to the findings of previous studies. The reduction of the epistemic uncertainty achieved through the restriction of the analysis to a particular seismic source leads to a sigma of about $0.25 \log 10$ units when the ergodic assumption is removed, suggesting that sigma at a particular site, due to a particular earthquake source, may reduce the sigma obtained for the Italian territory by Bindi, Pacor, *et al.* (2011) by about 30%.

Online Material: Tables of ground-motion prediction equation coefficients, site terms, and event-corrected single-station standard deviations.

Introduction

The standard deviation of the residuals about a median ground-motion prediction equation (GMPE), commonly termed sigma, is a fundamental parameter in probabilistic seismic-hazard analysis (PSHA). If the variability of the GMPE is larger, the expected ground motion of a PSHA at any given probability level is larger as well, and this effect is particularly pronounced at low exceedence probabilities. In GMPEs, sigma involves both the aleatory variability and the epistemic uncertainty, the latter referring to the lack of knowledge regarding the earthquake source process, the wave propagation in the region under study, and the ground-motion amplification due to the uppermost soil layers. In hazard calculations the epistemic uncertainty is usually taken into account through a logic tree approach (Bommer *et al.*, 2005), in which different models are associated with different weights, reflecting the relative confidence of the analyst in each of the adopted models.

In addition to epistemic uncertainty in the median ground-motion predictions, there is also epistemic uncertainty associated with the sigma for each equation; for example, there is still uncertainty about whether sigma is

dependent on earthquake magnitude (heteroscedastic) or independent of magnitude (homoscedastic).

In theory, sigma represents the randomness and should be irreducible; in practice, however, sigma is an apparent randomness in the observations with respect to a particular model that attempts to explain the latter. Therefore, if some model better explains the data, the variability should decrease. Refining the explanatory variables (i.e., classifying soil sites in terms of seismic response, or describing the source in terms of style of faulting) should thus lead to a decrease of the sigma of a GMPE. Nonetheless, the increase in the number of explanatory variables, coefficients, and the complexity of the functional forms of GMPEs developed in the past decades did not lead to a reduction of sigma (Strasser *et al.*, 2009), which remained on the order of 0.3 – $0.35 \log 10$ units.

In probabilistic seismic-hazard analysis, we should keep in mind that we are interested in the variation of ground-motion amplitudes at a particular site over time. Because, as a rule, we do not have observations over long periods of time at any given site, most of the GMPEs have been derived using

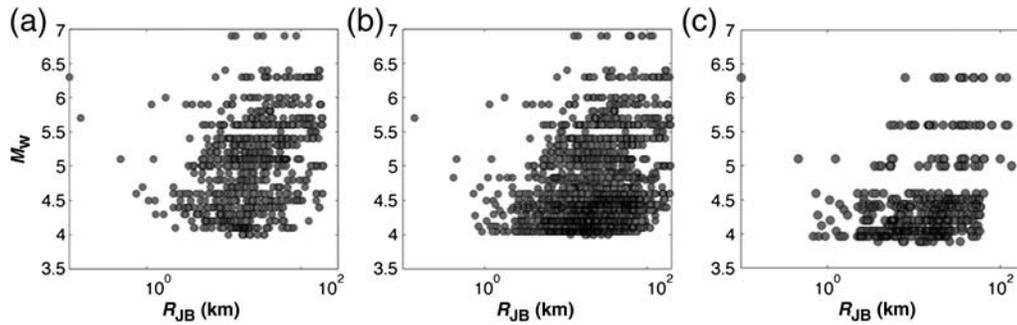


Figure 1. Magnitude–distance distribution of the three data sets: (a) *Blea*, (b) *Blea2*, and (c) *ABR*.

observed data from multiple stations and seismic sources, and the standard deviation is dominantly related to the statistics of the spatial variability of ground motion (Anderson and Brune, 1999). An ergodic assumption is made when PSHA treats that spatial uncertainty of ground motions as the uncertainty over time at a single point (Anderson and Brune, 1999); that is, the spatial variability of ground motion is assumed to be representative for the temporal one.

A solution for refining sigma is its decomposition into different parts (Al Atik *et al.*, 2010), so that if the various contributions to the variability can be identified and quantified, they can be subtracted from the total variability. When many recordings are available from a single station, the variability of the ground motion is smaller than the total sigma calculated for a standard GMPE (Atkinson, 2006; Morikawa *et al.*, 2008).

Atkinson (2006) found that the sigma value at individual sites, calculated from the ShakeMap data recorded at a group of stations in the Los Angeles region, were, on average, 10% smaller than the sigma calculated using all stations and first introduced the term single-station sigma. Morikawa *et al.* (2008) obtained a single-station sigma equal to 0.2 for six seismic sources using K-NET and KiK-net records. Rodriguez-Marek *et al.* (2011) used the KiK-net records for the breakdown of the sigma components and found the single-station standard deviation for multiple seismic sources was smaller by about 15% as compared with sigma evaluated using the ergodic assumption. Lin *et al.* (2011) quantified the reduction in the standard deviation for empirical ground-motion prediction models by removing the ergodic assumption with a data set of 64 shallow earthquakes in Taiwan. For peak ground acceleration and spectral accelerations at periods of 0.1, 0.3, 0.5, 1.0, and 3.0 s, they found the single-site standard deviations were 9%–14% smaller than the total standard deviation, whereas the single-path standard deviations were 39%–47% smaller.

These results led to single-station sigmas considerably lower than those calculated with the ergodic assumption, and therefore the measured ground-motion standard deviation at a single site can be used as a lower bound to the standard deviation for site-specific PSHA analysis. For practical applications, however, we should keep in mind that it is essential

that single-station standard deviation is used in PSHA only if estimates of the site term and of its epistemic uncertainty are also introduced in the analysis.

The goal of the paper is not the reduction of the overall variability of the residuals by reducing the epistemic uncertainty, but the evaluation of the influence of the ergodic assumption on the hazard assessment for a single site. To this aim, we explore the ground-motion variability using three different data sets. We verify to which extent the single-station sigma is lower than the global sigma evaluated for the Italian region considering the data set used by Bindi, Pacor, *et al.* (2011) to derive the most recent GMPEs for Italy. This data set has then been extended through addition of recent recordings (from 2009 to 2011) from events with magnitude lower than 5 and less accurate metadata, in order to explore the influence on sigma calculated at single stations. Finally, a third data set is composed of the recordings of the region recently struck by the 2009 L'Aquila sequence and is used to evaluate the single-station sigma for one seismic source zone (single-path sigma).

Data Set

The components of the variability of ground-motion models in the Italian region are evaluated using three acceleration data sets.

1. Data set *Blea*: The data set used to derive the most recent GMPEs for Italy (Bindi, Pacor, *et al.*, 2011) includes the recordings and the relevant metadata extracted from the Italian strong-motion database (Pacor *et al.* 2011) in the period 1972–2009. Bindi, Pacor, *et al.* (2011) selected the recordings such that they fulfilled the following criteria: (1) each earthquake should have an estimate of the moment magnitude, (2) each earthquake should be represented by at least two recordings, and (3) each station should be present in the data set with more than two recordings. The moment magnitude range is 4.0–6.9, and the Joyner–Boore distance range is 0–200 km (Fig. 1a). Only crustal events with focal depth shallower than 35 km were considered. Differently from Bindi, Pacor, *et al.* (2011) the events with only one record have been kept, with the aim of retrieving as many stations with

Table 1
Characteristics of the Three Data Sets Used in This Study

Data Set*	Number of Records	Number of Events	Number of Stations	M_w Range	Distance (km) [‡]
<i>Blea</i>	829	146	117	4.0–6.9	0–200
<i>Blea2</i>	2805	658	254	4.0–6.9 [†]	0–200
<i>ABR</i>	401	41	38	3.5–6.3 [†]	0–200

**Blea* is the data set used to derive the most recent GMPEs for Italy (Bindi, Pacor *et al.*, 2011); *Blea2* is an extension of *Blea* to include all records in the 4.0–6.9 magnitude range recorded from 1972 to 2011; *ABR* is a subset of the *Blea2*, composed of the records related to the 2009 L'Aquila sequence, in the 3.5–5.8 local magnitude range.

[†]Local magnitudes M_L have been converted into moment magnitudes M_w using the empirical relation of Castello *et al.*, 2007.

[‡]Distance is Joyner–Boore distance (R_{JB}) when the fault geometry is known ($M > 5.5$), otherwise it is the epicentral distance.

more than two records as possible. Therefore, the so-called *Blea* data set is composed of 829 records, from 146 earthquakes and 117 stations (whereas the data set by Bindi, Pacor, *et al.*, 2011, contains 769 records). This data selection has the consequence that only 25 stations recorded more than nine events.

2. Data set *Blea2*: This data set has been obtained by extending the *Blea* to include all records in the magnitude range 4.0–6.9 recorded from 1972 to 2011, in order to analyse a larger number of stations having recorded more than 10 events. Small events are characterized by local magnitudes M_L , which were converted into M_w using the empirical relation derived from the Italian earthquake catalog (Castello *et al.*, 2007). The magnitude and distance sampling is shown in Figure 1b. This data set consists of 2805 records from 658 events and 254 stations. The same Joyner–Boore distance and focal depth thresholds adopted for the *Blea* have been considered.
3. Data set *ABR*: This data set consists in a subset of the *Blea2*, composed of the records related to the 2009 L'Aquila sequence, in the local magnitude range 3.5–5.8. In order to isolate one seismic source, the events were selected in a spatial window with the latitude range 42.4–42.8 and the longitude range 13.2–13.6, approximately corresponding to the seismic source of the L'Aquila 2009 mainshock (M_w 6.3). The number of recordings in this data set is 401, relative to 41 events and 38 stations (Fig. 1c). All the events have normal style of faulting and shallow focal depths (<10 km).

The recording sites of all data sets are categorized into five classes, based on the shear-wave velocity intervals in the uppermost 30 m, V_{S30} , according to the Eurocode 8 (2004, EC8): class A, $V_{S30} > 800$ m/s; class B, $V_{S30} = 360 - 800$ m/s; class C, $V_{S30} = 180 - 360$ m/s; class D, $V_{S30} < 180$ m/s; class E, 5–20 m of C- or D-type alluvium underlain by stiffer material with $V_{S30} > 800$ m/s. A source of epistemic uncertainty common to all data sets is given by the fact that only about 130 recording stations are character-

ized by shear-wave velocity profiles. Therefore, the EC8 classification is mostly based on geologic considerations. The main characteristics of the three data sets are described in Table 1.

The waveforms have been processed uniformly, according to the procedure described in Paolucci *et al.* (2011): (1) baseline correction, (2) application of a cosine taper, based on the visual inspection of the record (typically between 2% and 5% of the total record length; records identified as late triggered are not tapered), (3) visual inspection of the Fourier spectrum to select the band-pass frequency range, (4) application of a second order acausal time-domain Butterworth filter to the acceleration time series padded with zeros, (5) double integration to obtain displacement time series, (6) linear detrending of displacement, and (7) double differentiation to get the corrected acceleration.

Theoretical Background

Following Bindi, Pacor, *et al.* (2011), the ground motion from data sets *Blea* through *ABR* are modeled as follows:

$$\log_{10} Y = e_1 + F_D(R, M) + F_M(M) + F_S + F_{sof}, \quad (1)$$

in which e_1 is a constant and the distance F_D and magnitude F_M functions can be expressed as

$$F_D(R, M) = [c_1 + c_2(M - M_{ref})] \log_{10}(\sqrt{R_{JB}^2 + h^2}/R_{ref}) - c_3(\sqrt{R_{JB}^2 + h^2} - R_{ref}), \quad (2)$$

$$F_M(M) = \begin{cases} b_1(M - M_h) + b_2(M - M_h)^2 & \text{for } M \leq M_h \\ b_3(M - M_h) & \text{otherwise} \end{cases}. \quad (3)$$

The functional form F_S in equation (1) represents the site amplification and is given by $F_S = s_j C_j$, for

$j = 1, \dots, 4$, in which s_j are the coefficients to be determined through the regression analysis, and C_j are dummy variables used to denote the four considered EC8 site classes (A through D, as class E has been disregarded due to the scarce number of samples). The functional form F_{sof} in equation (1) represents the style of faulting correction, and it is given by $F_{sof} = f_j E_j$, for $j = 1, \dots, 4$, in which f_j are the coefficients to be determined during the analysis and E_j are dummy variables used to denote the different fault classes. We considered four types of faulting style: normal, reverse, strike-slip, and unspecified. The variables M_{ref} , M_h , and R_{ref} (equations 2 and 3) have been fixed to 5, 6.75, and 1 km, respectively. As response variable Y , the peak ground acceleration (PGA in cm/s^2) is considered, along with 5% damped-spectral acceleration (SA in cm/s^2). Following [Brilinger and Preisler \(1985\)](#) and [Abrahamson and Youngs \(1992\)](#), a random effect model is introduced to describe the error terms as

$$Y = f(\mathbf{X}, \boldsymbol{\theta}) + \Delta, \quad (4)$$

in which Y is the base 10 logarithm of the observed ground-motion parameter, $f(\mathbf{X}, \boldsymbol{\theta})$ is the ground-motion model, \mathbf{X} is the vector of explanatory parameters (i.e., magnitude, distance, style of faulting, site conditions), $\boldsymbol{\theta}$ is the vector of model coefficients, and Δ is a random variable describing the total variability of the ground motion, with standard deviation σ . Δ is usually decomposed into between-events variability ΔB and within-event variability ΔW , which are zero-mean, independent, normally distributed random variables with standard deviations τ and ϕ , respectively.

The general form of the model is given by

$$y_{es} = \mu_{es} + \delta W_{es} + \delta B_e, \quad (5)$$

in which the subscripts e and s refer to event and station, respectively.

δB_e is the between-events residual, which corresponds to the average misfit of recordings from one particular earthquake with respect to the median ground-motion model; δW_{es} is the within-event residual, which corresponds to the difference between an individual observation and the event-corrected median estimate.

The single-station sigma has been calculated following the procedure described in [Rodriguez-Marek et al. \(2011\)](#). The within-event residuals computed from the GMPE are used to evaluate the site term for each station:

$$\delta S2S_s = \frac{1}{NE_s} \sum_{e=1}^{NE_s} \delta W_{es}, \quad (6)$$

in which $\delta S2S_s$ is a random variable that represents the average within-event residual at each station and is hereby referred to as the site term, and NE_s is the number of events recorded at station s . This is a zero-mean random variable, and its standard deviation is denoted by $\phi S2S$, which quan-

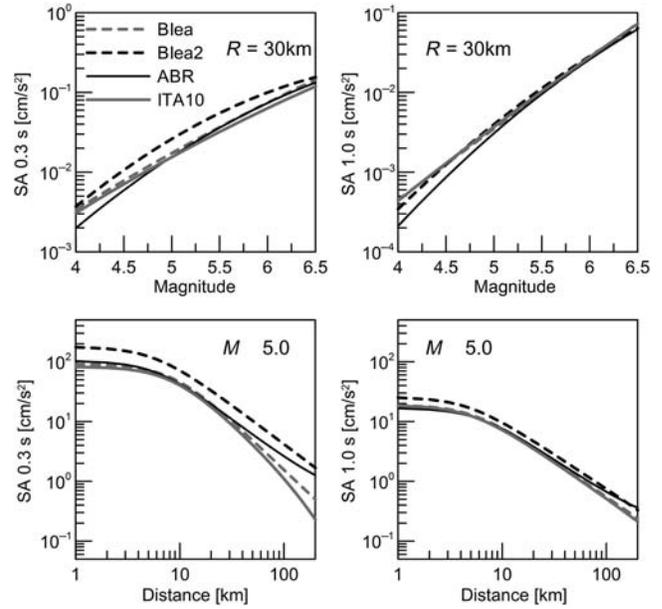


Figure 2. Spectral acceleration (SA) in function of magnitude (upper panel) and distance (lower panel). (Left) SA at 0.3 s; (right) SA at 1.0 s.

tifies the variability from site to site that cannot be explained by the model. The within-event residual is thus the sum of the site term and the event and site corrected residual as

$$\delta W_{es} = \delta S2S_s + \delta W_{o,es}. \quad (7)$$

The event-corrected single-station standard deviation can be computed for each site as

$$\phi_{ss,s} = \sqrt{\frac{\sum_{e=1}^{NE_s} (\delta W_{es} - \delta S2S_s)^2}{NE_s - 1}}, \quad (8)$$

and the event-corrected single-station standard deviation of all stations is

$$\phi_{ss} = \sqrt{\frac{\sum_{s=1}^{NS} \sum_{e=1}^{NE_s} (\delta W_{es} - \delta S2S_s)^2}{(\sum_{s=1}^{NS} NE_s - 1)}}, \quad (9)$$

in which NS is the number of stations in the data set.

Finally, the single-station standard deviation can then be computed as

$$\sigma_{ss} = \sqrt{\phi_{ss}^2 + \tau^2}. \quad (10)$$

Single-Station Sigma Evaluation

A separate regression has been carried out for each data set, in order to calculate the coefficients of the GMPEs and the components of the ground-motion variability.

Figure 2 shows the comparison among GMPEs. As expected the median predictions from the [Bindi, Pacor, et al.](#)

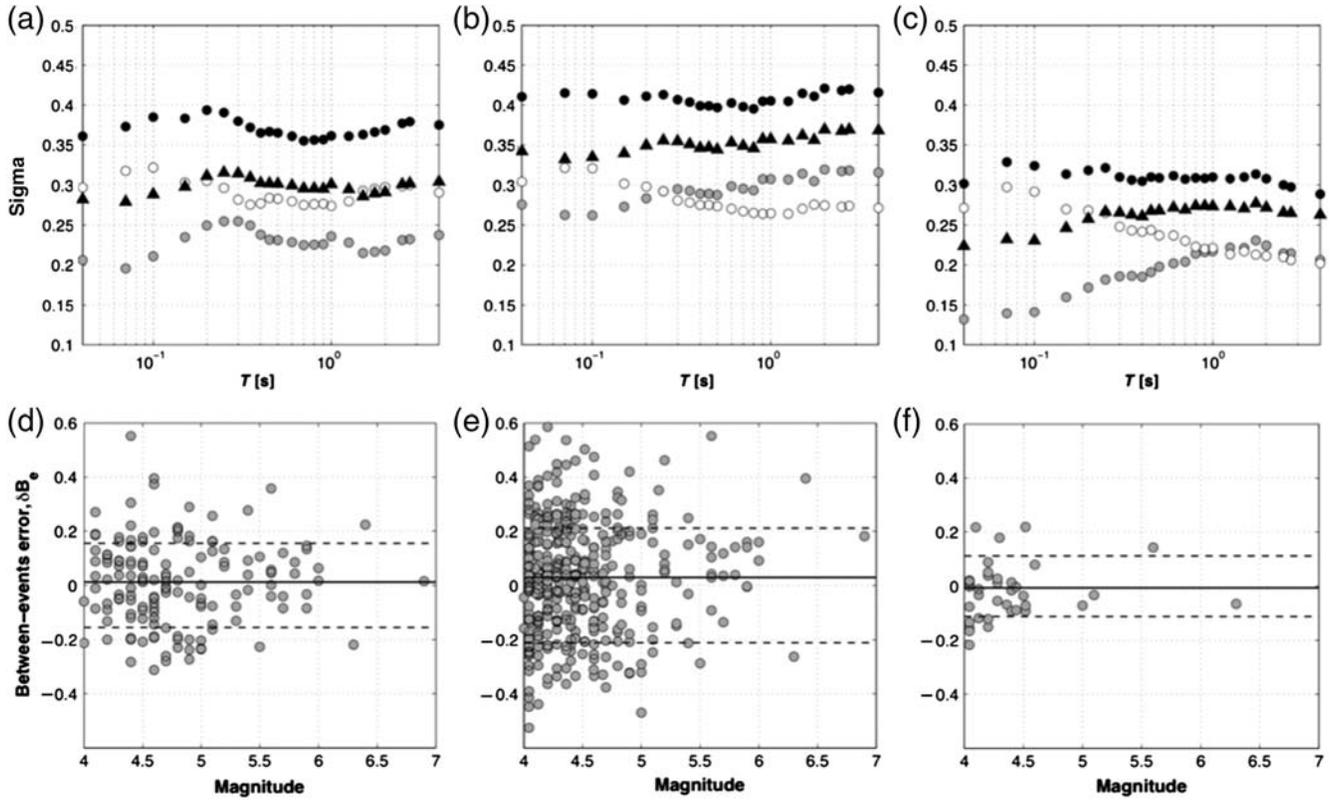


Figure 3. Sigma distribution as a function of period: (a) *Blea*, (b) *Blea2*, (c) *ABR* (black dots, total sigma with the ergodic assumption; white dots, within-event sigma; gray dots, between-events sigma; black triangles, total sigma after removal of the ergodic assumption; between-events error in function of magnitude at $T = 0.1$ s), (d) *Blea*, (e) *Blea2*, and (f) *ABR*.

(2011, termed ITA10) and the *Blea* data set coincide, as the data sets used to derive the GMPEs differ only for 60 records. Remarkable differences are observed for the GMPE derived from the *Blea2* data set. In particular, at short periods ($T = 0.3$ s) the median predictions are systematically larger than the rest of the GMPEs. This can be probably ascribed to the large proportion of small over moderate-to-large magnitude earthquakes. The GMPE derived from the L'Aquila fault data set (*ABR*) has median values close to the national GMPE, although lower values at small magnitudes and different attenuation trend at distances larger than 60 km can be observed. The \mathbb{E} coefficients of the three GMPEs derived in this study can be found in the electronic supplement to this paper.

Figure 3 shows, for the three data sets: (1) the total standard deviation evaluated with the ergodic assumption, σ ; (2) the different components of ground-motion variability (between-events sigma τ and within-event sigma ϕ), and (3) the total standard deviation without the ergodic assumption (σ_{ss}). As a complement, Table 2 lists all the values obtained for the single-station sigma and the total standard deviation without the ergodic assumption σ_{ss} as a function of periods. When the results, obtained using the *Blea*, are compared to those obtained with the extended data set, *Blea2*, it is evident how the introduction of low-magnitude earthquakes and the conversion from M_L to M_w increases the

epistemic uncertainty about the seismic events. As a matter of fact, it was shown for northwestern Turkey by [Bindi et al. \(2007\)](#) that the use of M_L instead of M_w for small earthquakes leads to lower estimates of between-event variability (τ). The between-events standard deviation is in fact lower than 0.25 in *Blea*, whereas it is about 0.3 (and even larger at some periods) in the case of *Blea2*. In the latter case, the between-events sigma becomes larger than the within-event sigma at $T > 0.25$ s, whereas it is always lower than the within-event one of the *Blea* data set. Conversely, in the case of the single-source data set, *ABR*, the epistemic uncertainty about the seismic events is reduced as the fault mechanisms are homogeneous and location and magnitudes are very well constrained. As a consequence, the between-events sigma is lower than 0.2 at periods up to $T = 2$ s. The within-event component of variability is the largest contributor to the overall standard deviation at short periods, whereas the between-events and within-event components are of the same order at periods longer than 1 s.

Figure 3d–f shows the between-events error in function of magnitude for the three data sets at $T = 0.1$ s, the period at which the largest discrepancy is observed among data sets. The variability of the between-events errors is larger for small magnitudes and when multiple seismic sources are sampled (*Blea*, *Blea2*). In particular, in the case of *Blea2*

Table 2
Single-Station Sigma (ϕ_{ss}) and Total Standard Deviation without the Ergodic Assumption (σ_{ss}) for the Three Data Sets

T (s)*	ϕ_{ss} <i>Blea</i> [†]	σ_{ss} <i>Blea</i> [‡]	ϕ_{ss} <i>Blea2</i>	σ_{ss} <i>Blea2</i>	ϕ_{ss} <i>ABR</i>	σ_{ss} <i>ABR</i>
PGA	0.1803	0.2824	0.1951	0.3351	0.1795	0.2257
0.04	0.1934	0.2825	0.2035	0.3423	0.1814	0.2240
0.07	0.1987	0.2789	0.2040	0.3324	0.1857	0.2322
0.10	0.1970	0.2886	0.2088	0.3351	0.1827	0.2309
0.15	0.1833	0.2978	0.2021	0.3394	0.1871	0.2461
0.20	0.1867	0.3114	0.2046	0.3495	0.1926	0.2579
0.25	0.1868	0.3159	0.2028	0.3558	0.1949	0.2665
0.30	0.1846	0.3145	0.1974	0.3549	0.1894	0.2654
0.35	0.1827	0.3091	0.1940	0.3510	0.1856	0.2630
0.40	0.1868	0.3027	0.1907	0.3464	0.1840	0.2611
0.45	0.1934	0.3019	0.1915	0.3469	0.1877	0.2677
0.50	0.1941	0.3017	0.1897	0.3445	0.1814	0.2683
0.60	0.1940	0.2997	0.1885	0.3532	0.1820	0.2718
0.70	0.1921	0.2956	0.1861	0.3489	0.1756	0.2690
0.80	0.1916	0.2958	0.1837	0.3459	0.1713	0.2741
0.90	0.1906	0.2954	0.1829	0.3574	0.1685	0.2738
1.00	0.1877	0.3012	0.1820	0.3573	0.1664	0.2734
1.25	0.1869	0.2950	0.1797	0.3558	0.1598	0.2735
1.50	0.1889	0.2861	0.1798	0.3621	0.1560	0.2709
1.75	0.1916	0.2893	0.1847	0.3564	0.1533	0.2769
2.00	0.1929	0.2912	0.1857	0.3692	0.1533	0.2716
2.50	0.1928	0.3900	0.1863	0.3679	0.1568	0.2655
2.75	0.1931	0.3020	0.1874	0.3695	0.1555	0.2649
4.00	0.1901	0.3040	0.1902	0.3683	0.1633	0.2629

* T is the period.

[†] ϕ_{ss} is the event-corrected single-station standard deviation (equation 9).

[‡] σ_{ss} is the single-station standard deviation (equation 10).

data set, the variability associated to small events increases, also because of the conversion of M_L into M_w .

The values of single station and the single-path sigma obtained by different authors (Atkinson 2006; Lin *et al.*, 2011; Rodriguez-Marek *et al.*, 2011) are shown in Tables 3 and 4 for several periods; the \oplus single-station sigma and the site terms for each station can be found in the electronic supplement. The single-station sigmas are of the same order, with the exception of those obtained by Lin *et al.* (2011) from a data set of 64 shallow earthquakes in Taiwan. Moreover, we report a recent study (Chen and Faccioli, 2013)

which exploits the data set of the 2010–2012 sequence of the Canterbury Region (New Zealand), finding single-path sigma in the range 0.15–0.2, consistent with the values shown in Table 4.

Figures 4–6 show the distribution of the within-event error (W_{es}), the site term ($S2S_s$), and the event and station corrected term ($W_{o,es}$) for the three data sets at different periods ($T = 0.3$ s, 1.0 s, and 2.0 s). The single-station sigma (the standard deviation of the event and station corrected term $W_{o,es}$) is lower for the *Blea* as compared with *Blea2* data set at short periods ($T = 0.3$ s). The sigma values invert

Table 3
Total Sigma (in log 10 Units) without the Ergodic Assumption for Different Studies

T (s)*	σ_{ss} <i>Blea</i> [†]	σ_{ss} Rodriguez-Marek <i>et al.</i> (2011) [†]	σ_{ss} Lin <i>et al.</i> (2011) [†]	σ_{ss} Atkinson (2006) [†]
PGA	0.282	0.291	0.253	0.268
0.1	0.289	—	0.267	—
0.3	0.314	0.286	0.269	0.295
0.5	0.301	—	0.272	—
1.0	0.301	0.265	0.278	0.268
3.0	—	—	0.290	0.260

* T is the period.

[†] σ_{ss} is the single-station standard deviation (equation 10).

Table 4
Single-Station Single-Path Sigma for Different Studies (in log 10 Units)

T (s)*	σ_{ie} <i>ABR</i> [†]	σ_{ie} Rodriguez-Marek <i>et al.</i> (2011) [†]	σ_{ie} Lin <i>et al.</i> (2011) [†]	σ_{ie} Atkinson (2006) [†]
PGA	0.179	0.188	0.146	0.18
0.1	0.183	0.194	0.158	—
0.3	0.189	0.195	0.157	0.2
0.5	0.181	0.178	0.172	—
1.0	0.166	0.171	0.187	0.18
3.0	—	—	0.200	—

* T is the period.

[†] σ_{ie} is the single-path standard deviation.

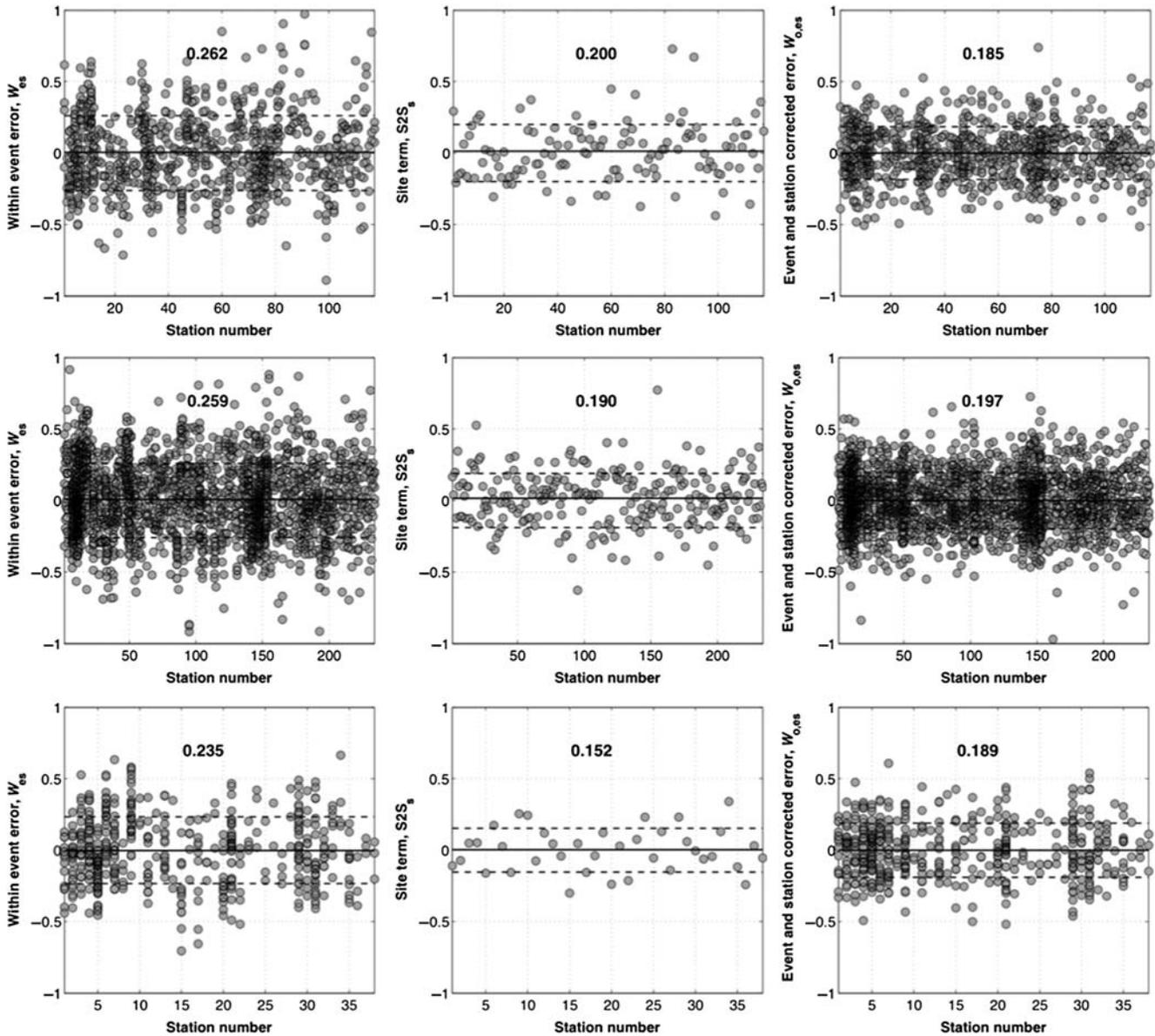


Figure 4. Average residuals at $T = 0.3$ s. (Left) Within-event error; (middle column) site term; (right) event and station corrected residuals. (Top) *Blea* data set; (middle) *Blea2*; (lower) *ABR* (numbers indicate the standard deviation of the parameter).

at long period ($T = 1$ and 2 s), as the between-events sigma is the largest component of ground-motion variability at long periods. The single-station sigma significantly reduces when considering the single-source data set *ABR* at all the considered periods.

Figure 7 shows the distribution of the single-station sigma at individual sites for the three data sets, at the same periods used in Figures 4–6. In Figure 7 the values of the average single-station sigma as well as the average of the sigma associated to the within-event residuals are shown.

Comparing the *Blea* and *Blea2* data sets (Fig. 7a–c and Fig. 7d–f, respectively) it is notable that, although the average single-station sigmas are similar, the one associated to the *Blea* at all periods has in general larger dispersion, which

means an increasing number of records could reduce the single-station sigma at individual sites. The distribution of single-station sigma at single sites for the *ABR* data set, shows the majority of values converge to the average sigma value, thus reducing the variability (Fig. 7g–i).

Figure 8 shows the single-station sigma associated to magnitude and distance bins at $T = 0.3$ s. Common to the *Blea* and *Blea2* data sets is the relatively large variability in the distance range 0–20 km at all magnitudes and the large variability associated with low-magnitude events as compared with moderate-to-large magnitudes, as observed in previous studies (Abrahamson *et al.*, 2008). At distances between 80 and 100 km a large variability of the $W_{o,es}$ can be found, probably due to the reflection of *S* waves at the Moho

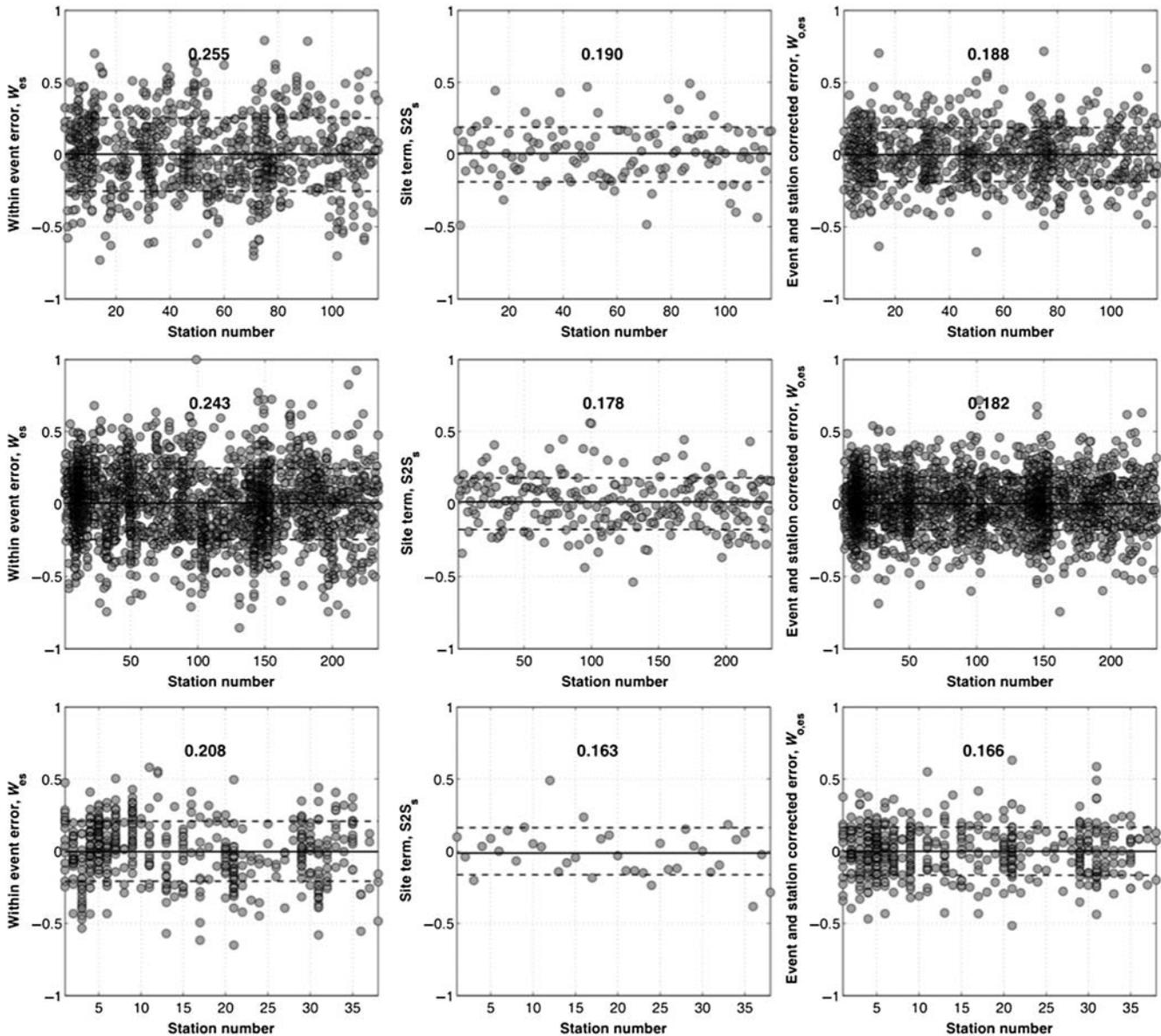


Figure 5. Average residuals at $T = 0.1$ s. (Left) Within-event error; (middle) site term; (right) event and station corrected residuals. (Top) *Blea* data set; (middle) *Blea2*; (lower) *ABR* (numbers indicate the standard deviation of the parameter).

(*SmS* phase) observed in literature (Bragato *et al.*, 2011, in northern Italy; Ponziani *et al.*, 1995, and Bindi *et al.*, 2004, in central Italy). This feature is not appreciable for the *ABR* data set, because of the scarcity of data samples in this distance range. The largest observed single-station standard deviation is for magnitudes in the range 5–5.5 and distances in the range 100–120 km.

In Figure 9 the single-station sigma associated to magnitude and distance bins is shown for $T = 1.0$ s. Similar patterns are observed as for $T = 0.3$ s, although the dispersion of the $W_{o,es}$ at distances between 80 and 120 km is smaller, as the phenomenon of the Moho reflection is mainly observed at high frequencies (Bragato *et al.*, 2011).

Figure 10 displays the single-station sigma for individual sites obtained with the *Blea* data set at different periods

(PGA, $T = 0.3$ s, 1.0 s, and 2.0 s). On average, the single-station sigma of the entire data set has values ranging from 0.18 to 0.2 log 10 units, although quite a large number of stations have sigmas larger than 0.25 and some even larger than 0.3 units.

Such large variability at single sites could be attributed to multiple factors (Bindi, Luzi, *et al.*, 2011). Stations Bronte (BNT) or Catania (CAT) are located in the Etna volcano area, where it is well known that crustal attenuation is different from the rest of Italy (De Natale *et al.*, 1988; Patanè *et al.*, 1994). Other stations are installed on peculiar sites, like Scafa (SCF), which is on an active landslide, Tolmezzo (TLM1) on a dam, and Cascia (CSC) on a slope. Tarcento (TRC) and Trenago (TRG) are the earliest Italian strong-motion sites and have mostly analog records. Gemona

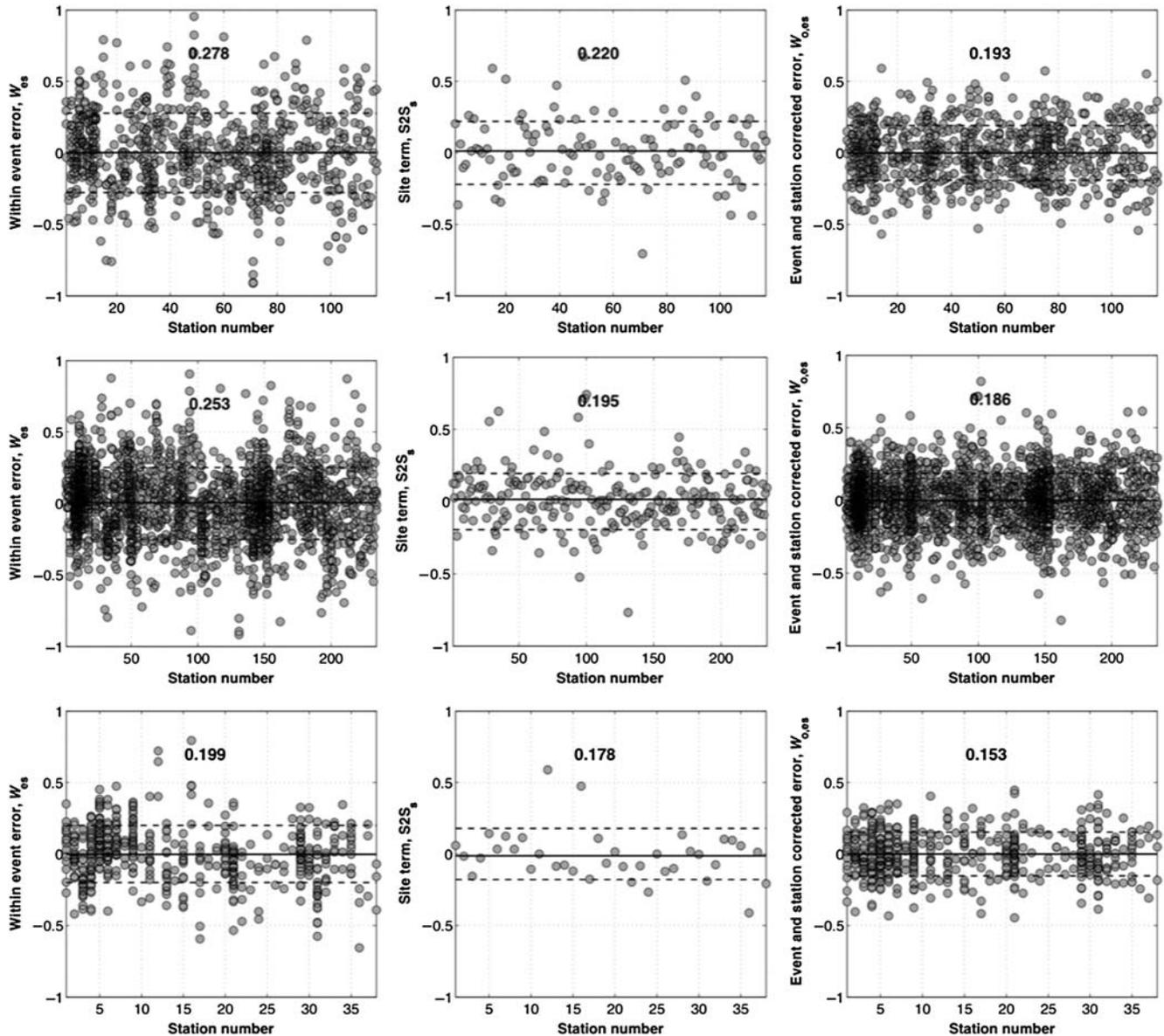


Figure 6. Average residuals at $T = 2.0$ s. (Left) Within-event error; (middle) site term; (right) event and station corrected residuals. (Top) *Blea* data set; (middle) *Blea2*; (lower) *ABR* (numbers indicate the standard deviation of the parameter).

(GMN) has recorded the strongest Friuli events at short epicentral distances (< 10 km), where the ground-motion variability is expected to be the largest. Part of the large variability might also be attributed to scarcity of recordings as the cited stations have less than five records. Figure 10d–f show the values of the single-station sigma evaluated with different thresholds of the number of recordings for each station. In general the sigma values converge to values in the range 0.18–0.2 log 10 units, although at long periods ($T = 2.0$ s) the stations with the largest number of records ($N > 15$) are located on deep sedimentary basins (e.g., stations AQK, AVZ) and are affected by largest variability (single-site sigma larger than 0.2), because of 2D or 3D amplification effects.

Figure 11 shows the distribution of $\phi_{ss,s}$ for the most populated EC8 site categories (A, B, and C), because a com-

parison with V_{S30} values would not be feasible, due to the scarcity of measured shear-wave velocity profiles. The largest $\phi_{ss,s}$ values are observed for site category C, at short ($T = 0.3$ s) as well as at long ($T = 2.0$ s) periods. Category C of EC8, as stated above, can be affected by large variability, due to complex site effects. In the following we will compare the single-station sigmas evaluated for the individual sites, either using the *Blea* data set or the single-source data set *ABR* in order to verify whether several seismic sources and source-to-site paths might affect the ground-motion variability. The two data sets have in common only the mainshock of L'Aquila (M_w 6.3, 6 April 2009, 01:32:39 GMT) and its strongest aftershock (M_w 5.6, 7 April 2009, 17:47:37 GMT). Therefore variations in ground motion between the data sets can be ascribed to the different source-to-site paths

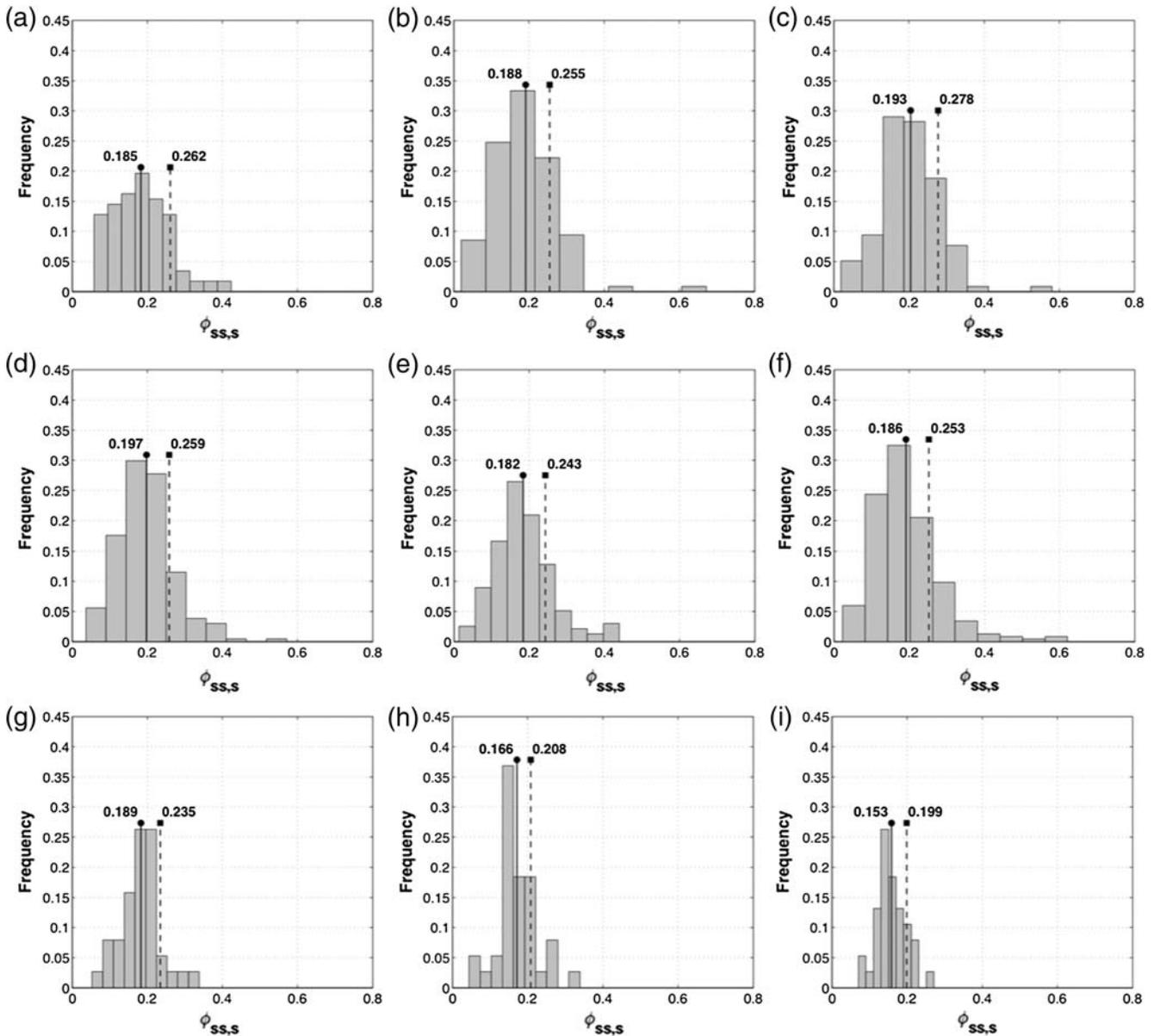


Figure 7. Frequency of the single-station standard deviation at individual sites. (Left) $T = 0.3$ s; (middle) $T = 1.0$ s; (right) $T = 2.0$ s. (Top) *Blea* data set; (middle) *Blea2*; (lower) *ABR* (continuous line with dot indicates the single-station standard deviation; dashed line with square indicates the standard deviation of the within-event error).

and to the characteristics of the seismic sources. Eight stations have been selected that have recorded L'Aquila and other seismic sequences (Lazio–Abruzzo 1984, Umbria–Marche 1997, Molise 2002) or sparse events (e.g., Gargano, Adriatic Sea). Figure 12 displays the distribution of the seismic sources for which the events have been recorded by the eight stations as well as the location of the recording sites.

Figure 13 shows the site terms in the 0.04–4 s period range (i.e., the average within-event residual at each station $S2S$, and the single-station sigma for each site $\phi_{ss,s}$). Site terms close to zero mean that the station, on average, has a response that closely follows its class. Positive site terms mean that, on average, the residuals of the station, corrected

for the between-events error, indicate amplification with respect to its class, whereas negative terms mean deamplification.

Three cases can be identified.

1. Stations with amplification with respect to the average of the class and low single-station sigma. Stations AQK and AQV are both classified as EC8 class B ($V_{S30} = 717$ m/s and 474 m/s, respectively). Station AQK is characterized by a velocity inversion, as high-velocity conglomerates overlay slower lacustrine deposits, and by a low fundamental frequency (Ameri *et al.*, 2009), whereas station AQV has a stratigraphic profile that indicates 50 m of alluvial deposits overlaying carbonate rock (see [Data](#)

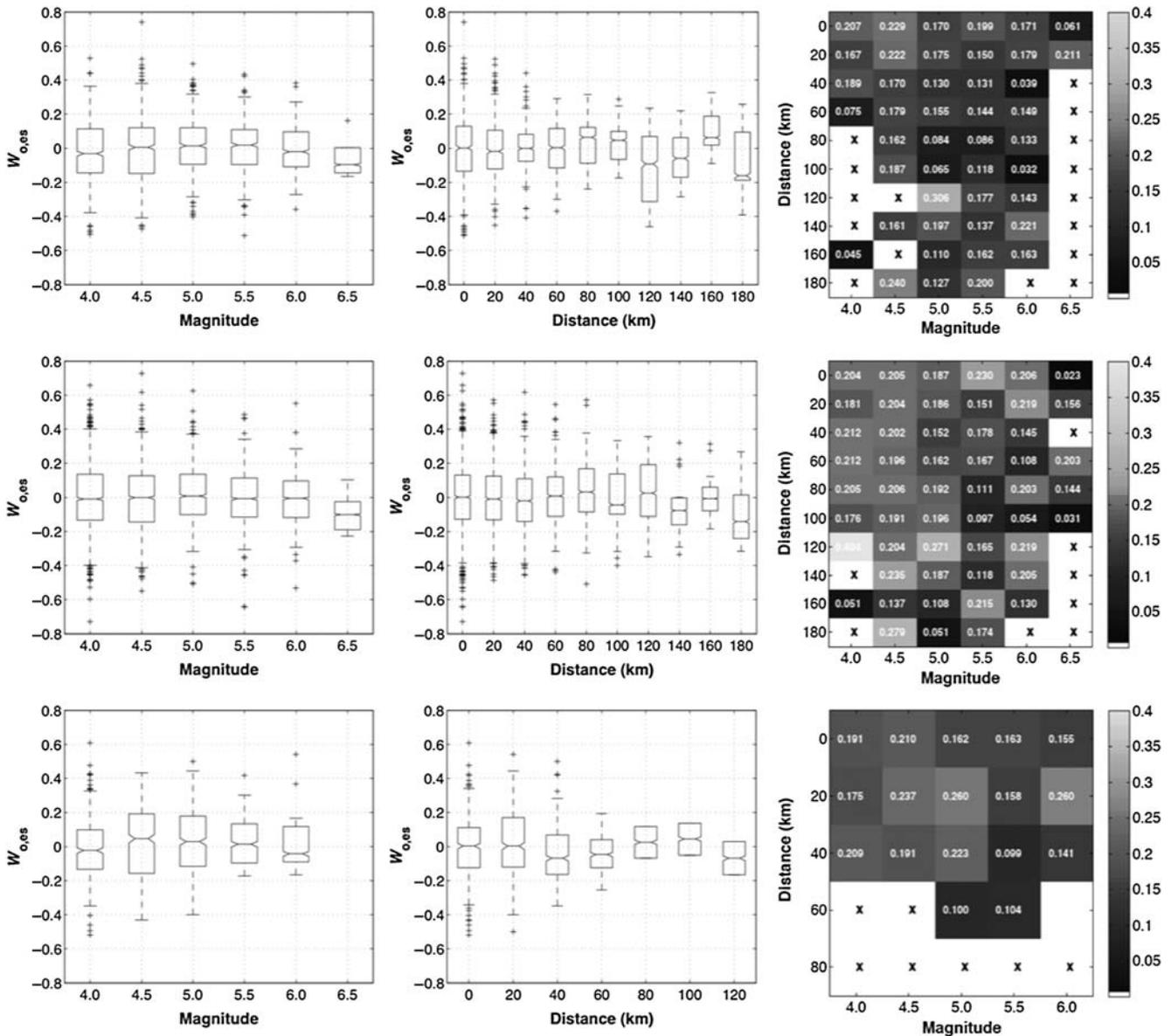


Figure 8. $T = 0.3$ s. (Left) Distribution of the event- and site-corrected errors for a range of magnitude bins; (middle) distribution of the event and site corrected error for a range of distance bins; (right) single-station sigma for distance–magnitude bins (in matrix view). The top panel is relative to the *Blea*, the middle panel to the *Blea2*, and the bottom panel to the *ABR* data set.

and Resources). Both stations show amplification with respect to their class, although at different periods, but very low variability at periods longer than 0.2 s (sigma equals to about 0.15 units). This means the random variability associated to the station is very low, but the misclassification can increase the overall sigma evaluated with the ergodic assumption.

2. Stations with no amplification with respect to the average of the class and low sigma. Other sites, such as AQA (EC8 class B, $V_{S30} = 552$ m/s) or NOR (EC8 class C, on the basis of geologic considerations), have negligible site terms (close to zero, indicating a similar response to the average response of the class) and very low sigma (indicating a stable behavior, irrespective of the different

seismic sources and source to site paths). Stations with these characteristics are extremely stable and therefore can be considered as benchmarks for ground-motion variability studies.

3. Stations with variable amplification with respect to the average of the class and variable sigma. Sites like AVZ (EC8 class C, $V_{S30} = 199$ m/s) or CHT (EC8 class B, on the basis of geologic considerations) exhibit different site terms depending on the recorded events. Site CHT shows large amplification at long periods when events originate on the L'Aquila fault, in combination with low-to-moderate variability (about 0.2 log 10 units). The site amplification is close to the average of the class in case of events from multiple sources and source-to-site

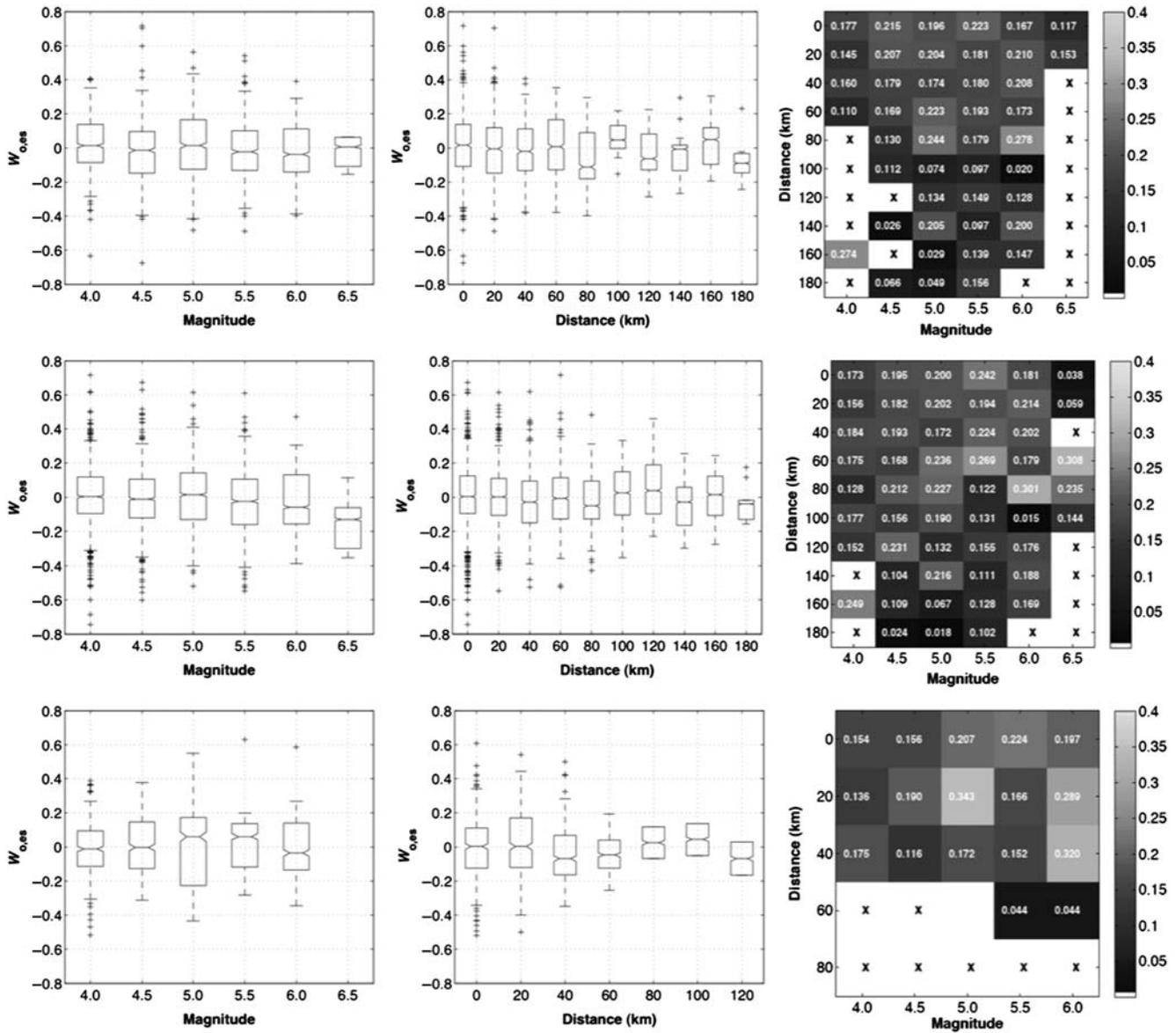


Figure 9. $T = 1.0$ s. (Left) Distribution of the event- and site-corrected error for a range of magnitude bins; (middle) distribution of the event- and site-corrected error for a range of distance bins; (right) single-station sigma for distance–magnitude bins (in matrix view). The top panel is relative to the *Blea*, the middle panel to the *Blea2*, and the bottom panel to the *ABR* data set.

paths, in combination with very high variability at long periods.

Different site terms associated to large sigma are also evident for GSA (EC8 class B, $V_{S30} = 488$ m/s) and SCF (EC8 class B, on the basis of geologic considerations). For both stations the variability of the response is larger when events are from multiple sources. When several seismic sources contribute to the hazard at the site, the single-station-sigma should be determined with data sets representative of multiple sources, in order to avoid biases of ground-motion variability.

Chen and Faccioli (2013) conclude that between-events standard deviation reflects the goodness of fit of data to the

GMPE adopted, whereas single-station sigma is mostly governed by site properties. On the contrary, this study demonstrates that different source-to-site paths can influence site amplification as well as single-station sigma.

Discussion and Conclusions

In this article, we explored the ground-motion variability at single sites decomposing sigma into different components, so that the various contributions to the variability could be identified and quantified. We could observe the variation of the different components of the standard deviation by controlling the introduction of epistemic uncertainty, through the usage of three different data sets.

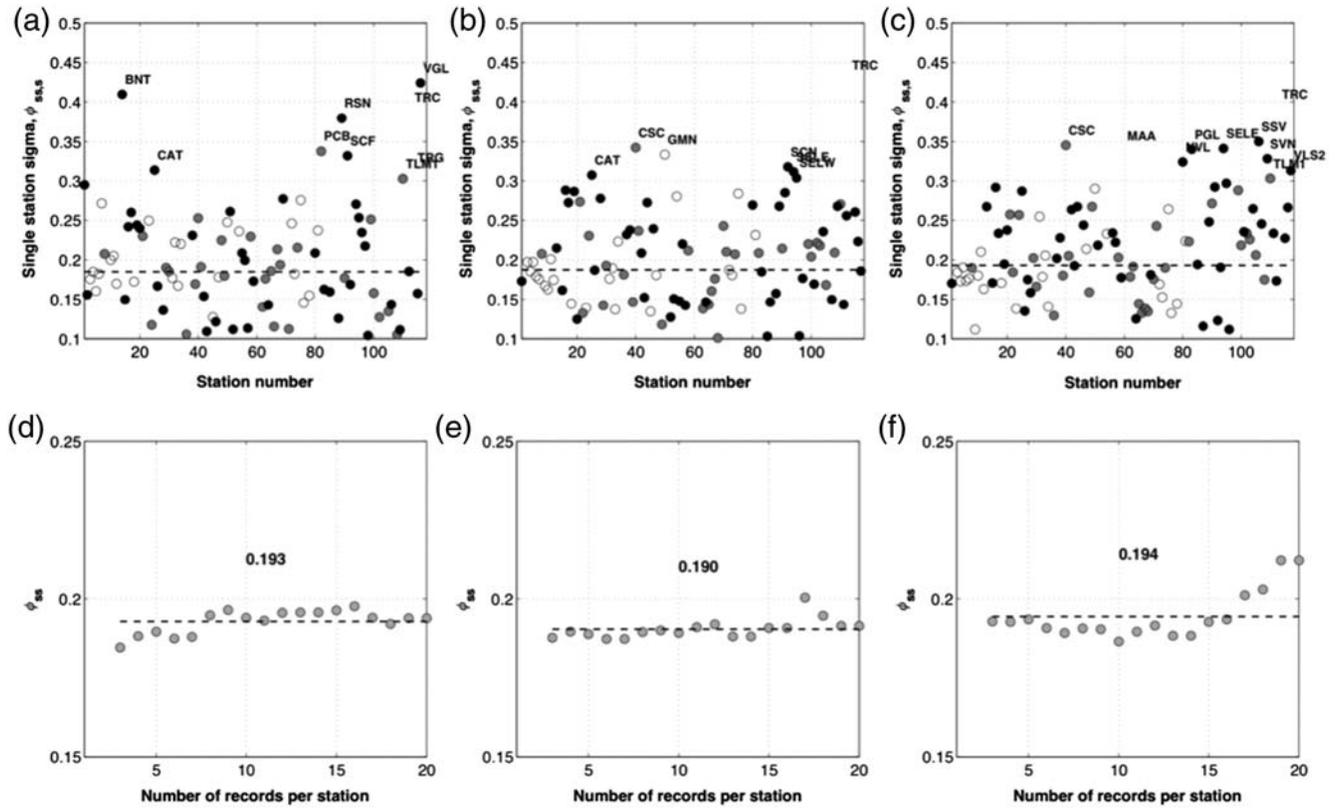


Figure 10. Single-station sigma at individual sites for *Blea*: (a) $T = 0.3$ s; (b) $T = 1.0$ s; and (c) 2.0 s (black dots, < 5 records; gray dots, 5–10 records; white dots, > 5 records). Single-station sigma evaluated as a function of the number of recordings per station: (d) $T = 0.3$ s; (e) $T = 1.0$ s; and (f) 2.0 s.

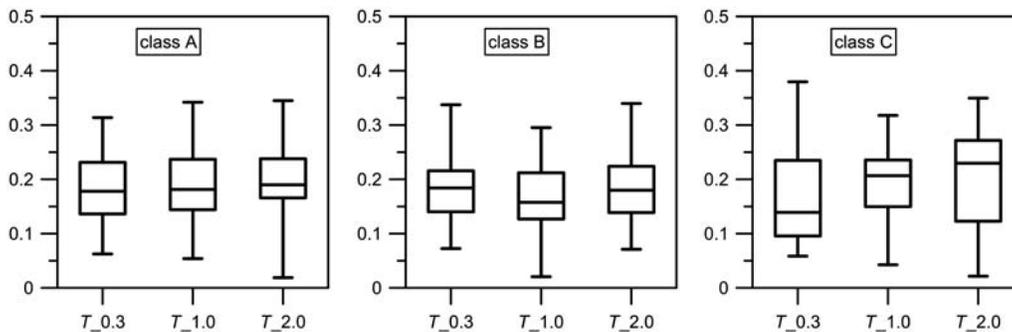


Figure 11. Distribution of $\phi_{ss,s}$ for EC8 site categories and spectral acceleration at three periods: $T = 0.3$ s ($T_{0.3}$), $T = 1.0$ s ($T_{1.0}$), and $T = 2.0$ s ($T_{2.0}$).

The main conclusions are the following:

1. Sigma obtained for Italy using the ergodic assumption is about 0.35 log 10 units (Bindi, Pacor, *et al.*, 2011) and decreases to about 0.3 when the single stations are considered (15% reduction).
2. Comparing the result with other studies (Table 3, Table 4), we observe that sigma without the ergodic assumption, $\sigma_{ss,s}$, evaluated with the Italian data set is slightly larger than the corresponding sigmas evaluated with other data sets (Atkinson, 2006; Lin *et al.*, 2011; Rodriguez Marek *et al.*, 2011); this suggests that sources of uncertainty could still be present, as several Italian stations have an insufficient number of records, and the event and station metadata could still be improved. Although there are minor differences, Italian, Californian, Japanese, and Taiwanese data sets converge to similar results.
3. The values of single-station sigma obtained in this study are rather stable, in the range 0.18–0.2 for the multiple-source data sets, and are comparable to the findings of Rodriguez Marek *et al.* (2011), between 0.17 and 0.21

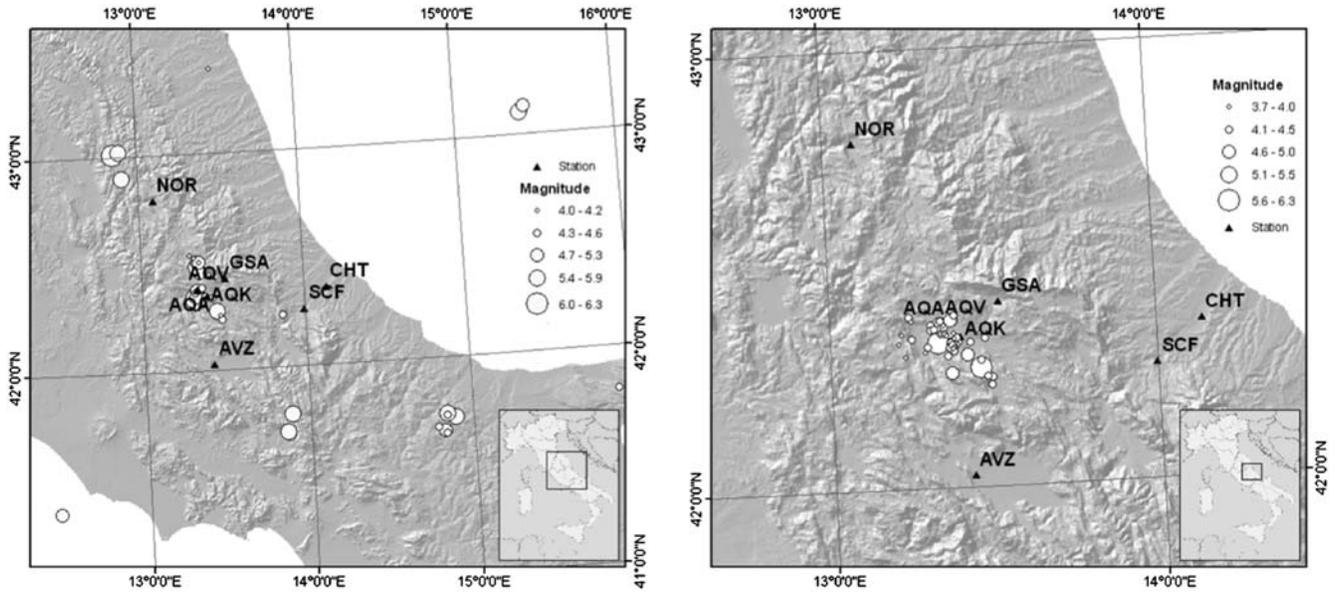


Figure 12. (Left) Epicenters of the event for the multiple-source data set (*Blea*); (right) epicenters of the events for the single-source data set (*ABR*). The triangles represent the recording stations in Figure 13.

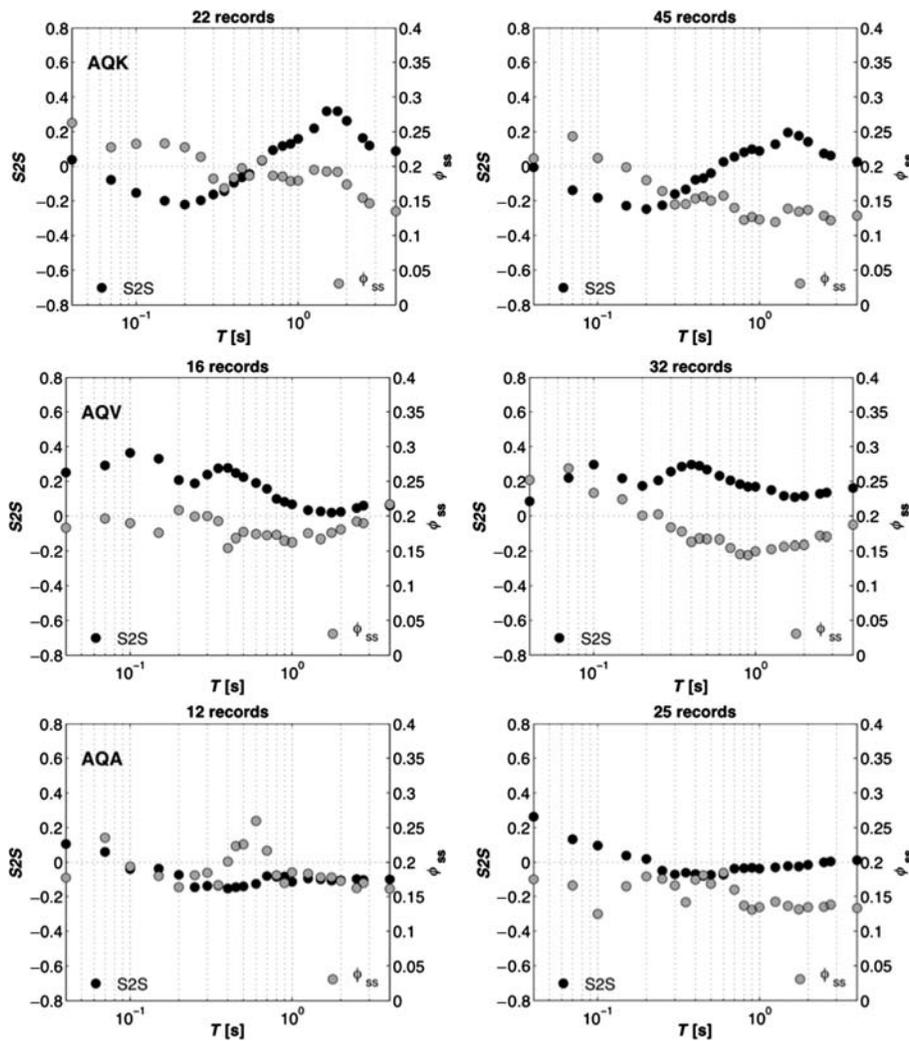


Figure 13. Site term and single-station sigma for representative stations: (left) *Blea* data set; (right) *ABR* data set. (Continued)

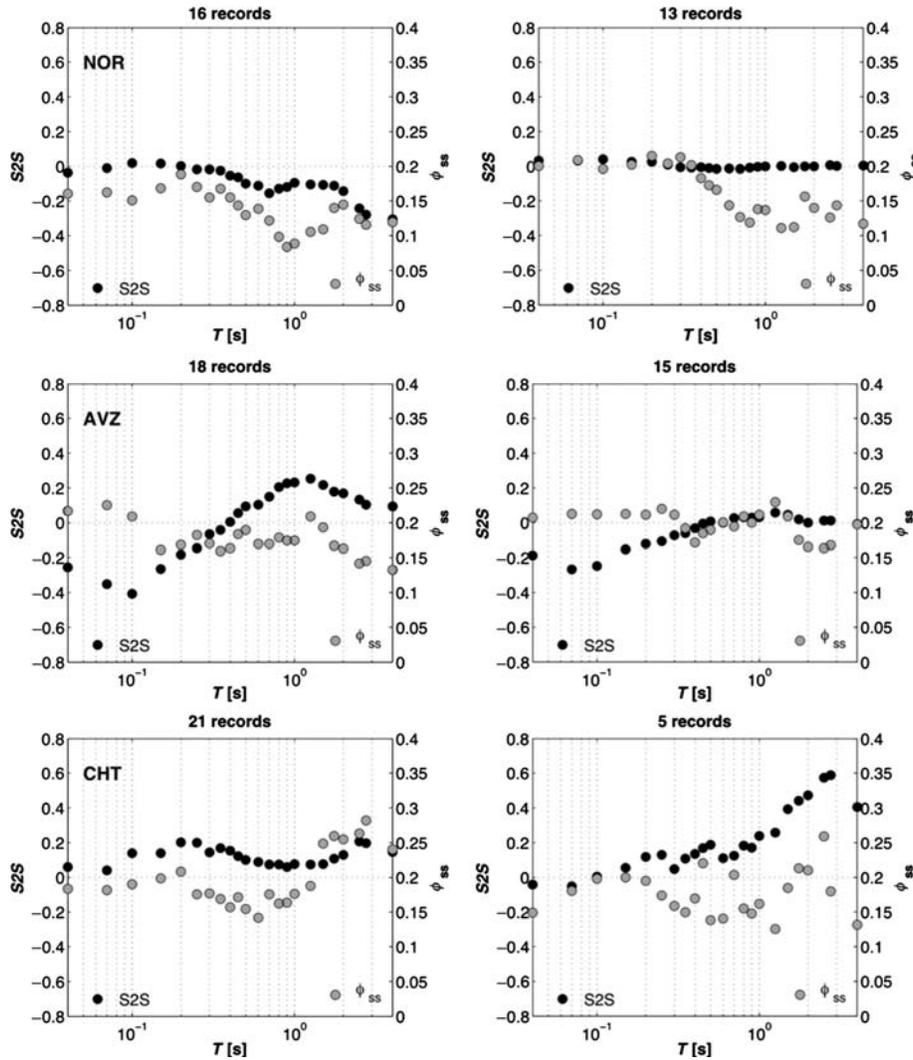


Figure 13. Continued.

for surface recordings of the KiK-net data set and multiple sources.

4. The introduction of epistemic uncertainty related to the seismic events (events with magnitude lower than 5 and M_L empirically converted to M_w) increases the between-events sigma (Fig. 3, data set *Blea* versus *Blea2*), which in turn increases the ergodic sigma up to 0.4. When single stations are considered, sigma decreases to 0.35, with a reduction of about 15% as observed for the *Blea* data set. On the other hand, the increase in the number of recordings slightly reduces the within-event sigma (Fig. 3, and Figs. 4–6 data set *Blea* versus *Blea2*).
5. The reduction of the epistemic uncertainty, restricting the analysis to a particular seismic-source zone, leads to a sigma of about 0.25 when the ergodic assumption is removed. This result is quite significant in that it suggests sigma at a particular site due to a particular earthquake

source may be reduced by about 30%, compared to the overall sigma for the Italian territory of about 0.35.

The single-station sigma is generally high at short distances from the seismic source (0–40 km), as shown by the matrix entries larger than 0.18 in Figures 8 and 9, particularly at short periods ($T = 0.3$ s). Large values of single-station sigma are mainly observed in case of low-magnitude events, suggesting once again the higher ground-motion variability for low magnitudes with respect to larger ones. Moreover, in the Italian region, an increase of ground-motion variability can be observed at distances of 80–120 km from the source, due to the reflection from the Moho especially at short periods ($T < 0.5$ s), as shown in Figures 8 and 9, in which a relative increase of the average event and station corrected term ($W_{o,es}$) is observed for the three data sets. This effect could be managed by adopting functional forms that more closely reflect the physical phenomena.

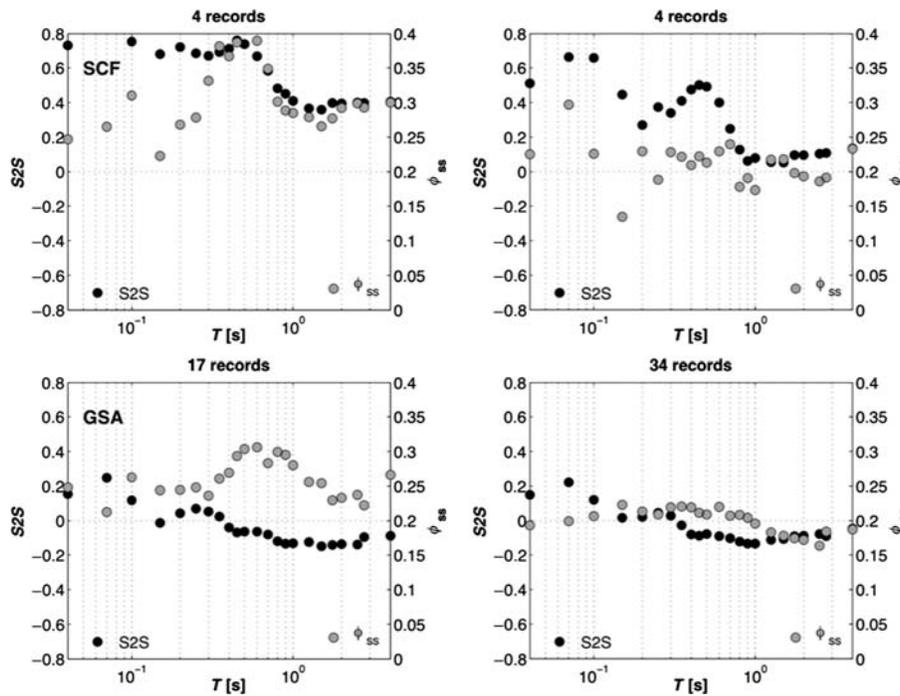


Figure 13. Continued.

The encouraging results of this study suggest to us that in Italy new efforts in data collection should be spent, in order to increase the strong-motion data set including events that occurred after 2011 (e.g., the M_w 6.1 2012 Emilia sequence); in addition a careful revision of event and station metadata will be of primary importance for the reduction of epistemic uncertainty. Finally a more careful site selection should be carried out when deriving new GMPEs, in order to remove those stations that might be affected by peculiar effects (landslides, stations affected by the response of a nearby infrastructure, stations located on volcanic areas, etc.), and, for this purpose, the preliminary analysis of the single-station sigma at individual sites might be of primary importance.

Data and Resources

The Italian strong-motion data and the geologic and geotechnical information regarding the Italian strong-motion stations have been searched using <http://itaca.mi.ingv.it> (last accessed April 2013).

Acknowledgments

The authors acknowledge Fabrice Cotton and one anonymous reviewer for the valuable contribution in improving the quality of this paper. The research activity of D. Bindi have been founded through the REAKT (Strategies and Tools for Real Time Earthquake Risk Reduction) EU-FP7 project.

References

Abrahamson, N. A., and R. R. Youngs (1992). A stable algorithm for regression analyses using The Random Effects Model, *Bull. Seismol. Soc. Am.* **82**, no. 1, 505–510.

- Abrahamson, N. A., G. Atkinson, D. Boore, Y. Bozorgnia, K. Campbell, B. Chiou, I. M. Idriss, W. Silva, and R. Youngs (2008). Comparisons of the NGA ground-motion relations, *Earthq. Spectra* **24**, no. 1, 45–66, doi: [10.1193/1.2924363](https://doi.org/10.1193/1.2924363).
- Al Atik, L., N. A. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81**, no. 5, 794–801, doi: [10.1785/gssrl.81.5.794](https://doi.org/10.1785/gssrl.81.5.794).
- Ameri, G., M. Massa, D. Bindi, E. D'alema, A. Gorini, L. Luzi, S. Marzorati, F. Pacor, R. Paolucci, R. Puglia, and C. Smerzini (2009). The 6 April 2009 M_w 6.3 L'Aquila (central Italy) earthquake: Strong-motion observations, *Seismol. Res. Lett.* **80**, no. 6, 951–966, doi: [10.1785/gssrl.80.6.951](https://doi.org/10.1785/gssrl.80.6.951).
- Anderson, J. G., and J. N. Brune (1999). Probabilistic seismic hazard analysis without the ergodic assumption, *Seismol. Res. Lett.* **70**, no. 1, 19–28.
- Atkinson, G. M. (2006). Single-station sigma, *Bull. Seismol. Soc. Am.* **96**, no. 2, 446–455, doi: [10.1785/0120050137](https://doi.org/10.1785/0120050137).
- Bindi, D., R. R. Castro, G. Franceschina, L. Luzi, and F. Pacor (2004). The 1997–1998 Umbria-Marche sequence (central Italy): Source, path, and site effects estimated from strong motion data recorded in the epicentral area, *J. Geophys. Res.* **109**, no. B4, B04312, doi: [10.1029/2003JB002857](https://doi.org/10.1029/2003JB002857).
- Bindi, D., L. Luzi, F. Pacor, and R. Paolucci (2011). Identification of accelerometric stations in ITACA with distinctive features in their seismic response, *Bull. Earthq. Eng.* **9**, no. 6, 1921–1939, doi: [10.1007/s10518-011-9271-5](https://doi.org/10.1007/s10518-011-9271-5).
- Bindi, D., F. Pacor, L. Luzi, R. Puglia, M. Massa, G. Ameri, and R. Paolucci (2011). Ground motion prediction equations derived from the Italian strong motion database, *Bull. Earthq. Eng.* **9**, no. 6, 1899–1920, doi: [10.1007/s10518-011-9313-z](https://doi.org/10.1007/s10518-011-9313-z).
- Bindi, D., S. Parolai, H. Grosser, C. Milkereit, and E. Durukal (2007). Empirical ground-motion prediction equations for northwestern Turkey using the aftershocks of the 1999 Kocaeli earthquake, *Geophys. Res. Lett.* **34**, L08305, doi: [10.1029/2007GL029222](https://doi.org/10.1029/2007GL029222).
- Bommer, J. J., F. Scherbaum, H. Bungum, F. Cotton, F. Sabetta, and N. A. Abrahamson (2005). On the use of logic trees for ground-motion prediction equations in seismic-hazard analysis, *Bull. Seismol. Soc. Am.* **95**, no. 2, 377–389, doi: [10.1785/0120040073](https://doi.org/10.1785/0120040073).

- Bragato, P. L., M. Sukan, P. Augliera, M. Massa, A. Vuan, and A. Sarao (2011). Moho reflection effects in the Po Plain (Northern Italy) observed from instrumental and intensity data, *Bull. Seismol. Soc. Am.* **101**, no. 5, 2142–2152, doi: [10.1785/0120100257](https://doi.org/10.1785/0120100257).
- Brillinger, D. R., and H. K. Preisler (1985). Further analysis of the Joyner–Boore attenuation data, *Bull. Seismol. Soc. Am.* **75**, no. 2, 611–614.
- Castello, B., M. Olivieri, and G. Selvaggi (2007). Local and duration magnitude determination for the Italian earthquake catalog, 1981–2002, *Bull. Seismol. Soc. Am.* **97**, no. 1B, 128–139, doi: [10.1785/0120050258](https://doi.org/10.1785/0120050258).
- Chen, L., and E. Faccioli (2013). Single-station standard deviation analysis of 2010–2012 strong-motion data from the Canterbury region, New Zealand, *Bull. Earthq. Eng.* **11**, 1617–1632, doi: [10.1007/s10518-013-9454-3](https://doi.org/10.1007/s10518-013-9454-3).
- De Natale, G., E. Faccioli, and A. Zollo (1988). Scaling of peak ground-motions from digital recordings of small earthquakes at Campi Flegrei, Southern Italy, *Pure Appl. Geophys.* **126**, 37–53.
- Eurocode 8 (2004). Design of structures for earthquake resistance, part 1: General rules, seismic actions and rules for buildings, EN 1998-1, European Committee for Standardization (CEN), Brussels, <http://www.cen.eu/cenorm/homepage.htm> (last accessed December 2013).
- Lin, P. S., B. Chiou, N. Abrahamson, M. Walling, C. T. Lee, and C. T. Cheng (2011). Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models, *Bull. Seismol. Soc. Am.* **101**, no. 5, 2281–2295, doi: [10.1785/0120090312](https://doi.org/10.1785/0120090312).
- Morikawa, N., T. Kanno, A. Narita, H. Fujiwara, T. Okumura, Y. Fukushima, and A. Guerpinar (2008). Strong motion uncertainty determined from observed records by dense network in Japan, *J. Seismol.* **12**, no. 4, 529–546, doi: [10.1007/s10950-008-9106-2](https://doi.org/10.1007/s10950-008-9106-2).
- Pacor, F., R. Paolucci, L. Luzi, F. Sabetta, A. Spinelli, A. Gorini, M. Nicoletti, S. Marucci, L. Filippi, and M. Dolce (2011). Overview of the Italian strong motion database ITACA 1.0, *Bull. Earthq. Eng.* **9**, no. 6, 1723–1739, doi: [10.1007/s10518-011-9327-6](https://doi.org/10.1007/s10518-011-9327-6).
- Paolucci, R., F. Pacor, R. Puglia, G. Ameri, C. Cauzzi, and M. Massa (2011). Record processing in ITACA, the new Italian strong-motion database, in *Earthquake Data in Engineering Seismology. Predictive Models, Data Management and Networks*, S. Akkar, P. Gulkan, and T. Van Eck (Editors), Geotechnical, Geological and Earthquake Engineering series, Vol. 14, Chapter 8, Springer, New York, 99–113.
- Patanè, D., F. Ferrucci, and S. Gresta (1994). Spectral feature of microearthquakes in volcanic areas: Attenuation in the crust and amplitude response of the site at Mt. Etna, Italy, *Bull. Seismol. Soc. Am.* **84**, 1842–1860.
- Ponziani, F., R. De Franco, G. Minelli, G. Biella, C. Federico, and G. Piali (1995). Crustal shortening and duplication of the Moho in the northern Apennines: A view from seismic refraction data, *Tectonophysics* **252**, 391–418.
- Rodriguez-Marek, A., G. A. Montalva, F. Cotton, and F. Bonilla (2011). Analysis of single-station standard deviation using the KiK-net data, *Bull. Seismol. Soc. Am.* **101**, no. 3, 1242–1258, doi: [10.1785/0120100252](https://doi.org/10.1785/0120100252).
- Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: Issues, insights, and challenges, *Seismol. Res. Lett.* **80**, no. 1, 40–56, doi: [10.1785/gssrl.80.1.40](https://doi.org/10.1785/gssrl.80.1.40).

Istituto Nazionale di Geofisica e Vulcanologia
Via Bassini 15
20133 Milano, Italy
lucia.luzi@mi.ingv.it
rodolfo.puglia@mi.ingv.it
francesca.pacor@mi.ingv.it
(L.L., R.P., F.P.)

GFZ German Research Centre for Geosciences
Centre for Disaster Management and Risk Reduction Technology (CEDIM)
Helmholtz Centre Potsdam
Telegrafenberg, 14473 Potsdam, Germany
bindi@gfz-potsdam.de
(D.B.)

European Center for Geodynamics and Seismology
19, rue Josy Welter
L-7256 Walferdange, Luxembourg
adrien.oth@ecgs.lu
(A.O.)

Manuscript received 15 April 2013;
Published Online 21 January 2014