Bulletin of the Seismological Society of America

NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground motion models in near-source

--Manuscript Draft--

Manuscript Number:	BSSA-D-21-00080R2
Article Type:	Article
Section/Category:	Special Section - Fault Displacement and Near-Source Ground Motion Models
Full Title:	NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground motion models in near-source
Corresponding Author:	Sara Anna Sgobba, PhD Istituto Nazionale di Geofisica e Vulcanologia (INGV) Milano, Italy ITALY
Corresponding Author's Institution:	Istituto Nazionale di Geofisica e Vulcanologia (INGV)
Corresponding Author E-Mail:	sara.sgobba@ingv.it
Order of Authors:	Sara Anna Sgobba, PhD
	Chiara Felicetta, MSc
	Giovanni Lanzano, PhD
	Fadel Ramadan, MSc
	Maria Clara D'Amico, PhD
	Francesca Pacor, PhD
Abstract:	We present an extended and updated version of the worldwide NEar-Source Strong motion (NESS) flat-file, which includes an increased number of moderate-to-strong earthquakes recorded in epicentral area, new source metadata and intensity measures, comprising spectral displacements and fling-step amplitudes retrieved from the extended BASeline COrrection (eBASCO) processing of velocity time series. The resulting dataset consists of 81 events with moment magnitude larger or equal than 5.5 and hypocentral depth shallower than 40 km, corresponding to 1189 three-component waveforms, which are selected to have a maximum source-to-site distance within one fault length. Details on the flat-files, metadata and ground-motion parameters, processing scheme and statistical findings are presented and discussed. The analysis of these data allows to recognize the presence of distinctive features, such as pulse-like waveforms, large vertical components and hanging-wall effects, that can be exploited to assess their impact on near-source seismic motion. As an example, we use the NESS2.0 dataset for calibrating an empirical correction factor of a regional ground-motion model mainly based on far-field records. In this way, we can adjust the median predictions to account for near-source effects not fully captured by the reference model. The final aim of this work is to promote the use of the NESS2 flat-file as a tool to disseminate qualified and referenced near-source data and metadata in the light of improving the constraints of ground-motion models (both empirical and physics-based) close to the source.
Author Comments:	
Suggested Reviewers:	Roberto Paolucci, PhD Full Professor, Politecnico di Milano roberto.paolucci@polimi.it Main expert in the field of ground-motion definition for design actions
	Karim Tarbali, PhD Research Associate, University College London, London, England, United Kingdom karim.tarbali.p@gmail.com Expert on near-source ground motion models
	Georgios Baltzopoulos, PhD

	Università degli studi di Napoli Federico II, Naples, Italy georgios.baltzopoulos@unina.it Expert in the field on near-source ground motion features
Opposed Reviewers:	
Response to Reviewers:	
Additional Information:	
Question	**Response**
Key Point #1: <i>Three key points will be printed at the front of your manuscript so readers can get a quick overview. Please provide three COMPLETE sentences addressing the following: 1) state the problem you are addressing in a FULL sentence; 2) state your main conclusion(s) in a FULL sentence; and 3) state the broader implications of your findings in a FULL sentence. Each point must be 110 characters or less (including spaces).	We propose an extended version of the worldwide NEar-Source Strong motion flat-file (NESS2)
Key Point #2:	The new dataset exhibits the typical near-source features
Key Point #3:	NESS2 can be used for calibrating new ground motion models or for constraining existing models in near-source

NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground-motion models in near-source

Authors: Sgobba S., Felicetta C., Lanzano G., Ramadan F., D'Amico M., Pacor F.

Corresponding Author

Sara Sgobba

Istituto Nazionale di Geofisica e Vulcanologia, Milano Department

Via Alfonso Corti, 12 – 20133 Milano, Italy

E-mail: sara.sgobba@ingv.it

Declaration of Competing Interests

The authors acknowledge there are no conflicts of interest recorded

Abstract

We present an extended and updated version of the worldwide NEar-Source Strong motion (NESS) flat-file, which includes an increased number of moderate-to-strong earthquakes recorded in epicentral area, new source metadata and intensity measures, comprising spectral displacements and fling-step amplitudes retrieved from the extended BASeline COrrection (eBASCO) processing of velocity time series. The resulting dataset consists of 81 events with moment magnitude larger or equal than 5.5 and hypocentral depth shallower than 40 km, corresponding to 1189 three-component waveforms, which are selected to have a maximum source-to-site distance within one fault length. Details on the flat-files, metadata and ground-motion parameters, processing scheme and statistical findings are presented and discussed. The analysis of these data allows to recognize the presence of distinctive features, such as pulse-like waveforms, large vertical components and hanging-wall effects, that can be exploited to assess their impact on near-source seismic motion. As an example, we use the NESS2.0 dataset for calibrating an empirical correction factor of a regional ground-motion model mainly based on far-field records. In this way, we can adjust the median predictions to account for near-source effects not fully captured by the reference model.

The final aim of this work is to promote the use of the NESS2 flat-file as a tool to disseminate qualified and referenced near-source data and metadata in the light of improving the constraints of ground-motion models (both empirical and physics-based) close to the source.

INTRODUCTION

The behaviour of strong-motion records near the fault-rupture reveals peculiar features, whose knowledge and proper modelling is fundamental to quantify the seismic input in structural engineering. Notably, the investigation of near-source effects is a well-known research issue, particularly motivated by the need to constrain Ground-Motion Models (GMMs) near the epicenter, and a key-point to define design spectra that are based on Probabilistic Seismic Hazard Assessment (PSHA). These effects are indeed generally underestimated in current seismic design procedures (Alavi and Krawinkler 2004; Kamai et al., 2014). Also, recent studies carried out in the Italian context have underlined the importance of a right quantification of near-source effects in building codes, as it has been demonstrated that elastic design spectral ordinates can be often exceeded in epicentral area, for moderate-to-high earthquakes (Iervolino et al., 2019; Cito and Iervolino, 2020). Since the publication of the 1997 version of the Uniform Building Code (UBC1997), some international codes have introduced near-source factors in the seismic demand to increase the capacity of buildings against near-fault ground-motions at different structural periods (e.g. the U.S. building code ASCE/SEI 7-16; the New Zealand Standards NZS 1170.5, NZS 2004; the Iranian seismic code IS 2800-05; the Taiwan's Building Seismic Design Code CPAMI 2011); see Grimaz and Malisan (2018) and Akkar et al. (2018) for a more detailed discussion. These seismic norms typically introduce simple distance- and period-dependent correction factors that modify the acceleration design spectrum (5%-damped) as a function of the slip-rate and maximum moment magnitude of the nearest seismically active sources. Such correction factors usually account for global near-source impacts or for augmented vertical to horizontal spectral ratio at short-periods (e.g. ASCE 7-16; Eurocode 8-EC8; NZS2004), while a few of them consider explicitly specific effects, such as those related to directivity (NZS2004, ASCE 7-16).

67 The extension of seismic monitoring networks worldwide and the increasing availability
68 of high-quality recordings close to the sources offer the opportunity to investigate these
69 aspects more deeply, with the main aim of including near-source impacts on ground-motion
70 modelling. This is currently a matter of debate among researchers and practitioners by
71 several decades (Sommerville 1998; Abrahamson, 2001; Sommerville, 2003; Bray and
72 Rodriguez-Marek, 2004; Chopra and Chintanapakdee, 2001; Mavroeidis and
73 Papageourgiou, 2003, among the others). A key-part of this discussion deals with the
74 definition of the near-source boundary, for which there is not yet a shared definition. In
75 general, the near-source zone is typically considered to be within a distance of about 20-30
76 km from the fault rupture or even greater, often proposed in an arbitrary way. For instance,
77 Campbell and Bozorgnia (2013) for NGA-West 2 project, used rupture distances less than
78 80 km as near-source, while Farajpour et al. (2018) used a distance less than 120 km for
79 the Iranian catalog of near-source strong-motion records. Different definitions were also
80 given by other authors, such as Maniatakis et al. 2008, who used a set of intensity measures
81 of ground-motion to identify records capable of causing a given level of Modified Mercalli
82 Intensities (MMI), and Martinez-Pereira and Bommer (1998), as well as Spyrakos et al.
83 (2008).

84 Another key-issue is that the majority of the previous studies refer to data of regional
85 networks or to simulated-based or hybrid datasets, which include the most typical intensity
86 measures of engineering interest, whereas displacement-based and integral parameters as
87 well as fling-step measures are still unusual.

88 Moreover, datasets specifically compiled to collect near-source parameters are not
89 common (some examples are the Iranian catalog by Farajpour et al., 2018 and the database
90 SIMBAD by Smerzini and Paolucci, 2013, the latter collecting near-source records for
91 displacement-based assessment and design). More frequently, near-source datasets are

simply extracted on the basis of a distance threshold from more extended datasets including far-field data. Examples of adopted regional or global datasets are the Southern California Seismic Network (SCEDC), the Consortium of Organizations for Strong Motion Observation Systems (COSMOS) database, which archives strong-motion data from the U.S. Geological Survey and California Geological Survey, the K-NET (Kyoshin Network) and KiK-net (Kiban Kyoshin Network) of Japan, the Taiwan network, as well as other strong-motion arrays worldwide (see *Dataset and Resources*). A recent database especially reached in near-source parameters from KiK-net records is provided by Bahrampouri et al. (2021).

An important effort to collect ground-motion data and metadata for the development of GMMs including near-source effects was performed in 2008 by the PEER (Pacific Earthquake Engineering Center), in the framework of the Next Generation Attenuation (NGA) of Ground Motions Projects (see *Dataset and Resources*), i.e. NGA-West2 and NGA-East. This research program developed new robust GMMs including typical near-fault features (Spudich and Chiou 2008; Rowshandel 2010; Shahi and Baker 2011), which, however, required the specification of source parameters in addition to standard predictive ones (such as slip direction, focal mechanism parameters etc.) that are scarcely available in current public repositories of strong-motion records.

To address all the mentioned issues, in 2018 the NEar-Source Strong motion flat-file (NESS1; Pacor et al., 2018a; see *Dataset and Resources*) was published. NESS1 aimed to provide a comprehensive basis of metadata and ground-motion parameters of uniformly processed waveforms recorded exclusively in epicentral area from moderate-to-strong worldwide events. Also, the selection of the data was based on a rational description of the near-source boundary, which is defined according to seismological scaling laws depending on fault size and stress drop (see Pacor et al., 2018b for details).

Evidence of near-source effects was recognized in the NESS1 flat-file, such as velocity pulses, large vertical ground-motions, directional and hanging-wall amplification and fling step, which allowed improving the calibration of ground-motion models and scaling properties in epicentral area (Pacor et al., 2018b; D'Amico et al., 2020; Baltzopoulos et al., 2020; Sgobba et al., 2021a; Schiappapietra et al., 2021). A fundamental step in the compilation of the NESS1 dataset was to retrieve adequate information about event source geometries and rupture mechanisms, which are usually poorly represented in worldwide databases, as well as to calculate different distance metrics from referenced finite-fault models.

Here we present a new expanded and updated version of the NESS1 flat-file (the NESS version 2.0, hereinafter NESS2), which includes both an increased number of earthquake records and new source metadata. A relevant novelty in the new NESS version is the publication of a companion flat-file called NESS2-eBASCO, which provides displacement spectra and fling-step amplitudes retrieved from a piecewise linear detrending of velocity time series (D'Amico et al., 2019; Schiappapietra et al., 2021). This processing scheme, named eBASCO (extended BASeline COrrection), does not apply any filtering to remove the low-frequency content of the signal, thus preserving specific properties of near-source ground-motion featured by one-side pulse in the velocity trace, and the offset at the end of the displacement traces.

In the following, we first present the main structure of the flat-file and discuss the statistics on data, then we perform the calibration of new empirical relationships as a set of tools useful for near-source GMMs calibration and for developing empirical correction factors of attenuation models.

In this view, NESS2 may be considered a dataset suitable to increase the knowledge on the near-source features of ground motion and a tool for constraining the existing regional and global GMMs near the epicenter.

DATASET AND COMPILATION CRITERIA

We compiled NESS2 according to the same criteria adopted to produce the flat-file NESS1 (for more complete description refer to Pacor et al., 2018b), which are the following:

1) data relative to events characterised by moment magnitude $M_w \geq 5.5$ distributed worldwide (Figure 1);

2) hypocentral depth ≤ 40 km;

3) fully-referenced information about the finite-fault source model;

4) maximum source-to-site distance defined through seismological considerations introduced by Pacor et al. (2018b) by assuming the parameter *k* (i.e. number of fault length) equal to unity and average stress drop 10 MPa; the latter implies, for instance, a cut on Joyner-Boore distances of about 15 km for M_w6.0 and 40 km for M_w7.0 (Figure 2a).

The latter point is crucial in order to assume a unique criterion for the definition of the "near-source" regime, that was based on a physics-informed metric instead of an arbitrary one.

The resulting dataset is composed of 1189 manually processed waveforms recorded by 1001 accelerometric stations (mainly installed in free-field conditions) and relative to 81 events occurred in the period time 1933 – 2020. The majority of the events are located in the United States (20 events) and Italy (20 events). The remaining earthquakes are

161 distributed among Japan (10), Turkey (7), Greece (6), Iran (5), New Zealand (5),
162 Montenegro (2), Mexico (2), Taiwan (2), Nepal (1), Uzbekistan (1), and Chile (1).

163 Compared to NESS1, we collected additional source data and waveforms related to 7
164 events and 421 records, which are listed in Table 1: the new events mainly contribute to
165 magnitudes larger than 7.0 (more than 65% with respect to NESS1) and thrust-fault
166 mechanisms, due to the relevant contribution of records corresponding to the event of Chi-
167 Chi, Taiwan 1999. It can be noted the increase of records is distributed at roughly all the
168 distances (Figure 2b). The final dataset has the advantage to include also recent
169 earthquakes, as for example those related to the seismic sequences 2016-2017 of Central
170 Italy as well as the Ridgecrest event on 2019-07-06, providing new contributions compared
171 to what has been published in global datasets. In particular, there are 36 events that are not
172 incorporated in the NGA-West2 database (about 44% over the total events in NESS2) and
173 that correspond to 425 records, the 30% of which occurred after 2014. Moreover, the
174 majority of European earthquakes, particularly those occurring in Italy, are densely sampled
175 thanks to the availability of records from temporary stations installed during the seismic
176 sequences.

177 Namely, the European records and metadata in NESS2 come from the constant updates
178 of strong-motion waveforms collected in the Engineering Strong-Motion Database (ESM)
179 database (Luzi et al., 2016, Luzi et al. 2020), while data from other regions come from
180 different authoritative sources (see *Data and Resources* for a complete list). Earthquake
181 fault models are retrieved from published studies or from regional and international
182 databases, consulting different sources (see Pacor et al., 2018b). In detail, all the data refer
183 to validated metadata and intensity measures (IMs) of waveforms downloaded from different
184 worldwide repositories in raw version and then manually corrected by using the standard
185 procedures described later.

OVERVIEW ON NESS2 STRUCTURE

Data and metadata in NESS2 are distributed as tables provided in two comma-separated files, which can be downloaded from the NESS web-page (see *Dataset and Resources*). A schematic representation of NESS2 flat-files is given in Figure 3. The two tables rely on the same dataset and have in common the main metadata related to event, station and waveforms, whereas additional information in terms of source metadata and IMs are provided separately in the two tables, as outlined in the following sections.

The NESS2 flat-file

The 'NESS2_flat-file.csv' (Sgobba et al., 2021b) is a parametric table containing information about event and source parameters, such as epicenter coordinates, focal depth, nucleation depth, depth to the top of the rupture fault, fault dimensions, style of faulting and fault plane parameters (i.e. strike, dip and rake), moment magnitude, as well as stations and waveforms metadata. In the new version, the following fields are also included: i) the seismic moment M_0 (according to two kinds of estimates: the first computed from the moment tensor solution consistent with the focal mechanism reported in the table; the other obtained from verified source models available in the literature, which are fully referenced in a specific field of the table), and ii) the stress drop $\Delta\sigma$, according to the estimates by Bindi and Kotha (2020) for Europe events, and by Oth et al (2014; 2017) for New Zealand and Japan events. Both these parameters are of great utility for developing new scaling laws due to their predictable effect on the radiated seismic signal (Brune, 1970).

The source of each event in the dataset is fully-referenced with information about the finite-fault source model. Smaller events missing a fault model are featured by a *virtual* fault that is calculated by modifying the procedure by Kaklamanos et al. 2011 on the basis of the

magnitude M_w, the strike and dip of the fault-plane solutions of moment tensor, and the hypocentral depth.

The table also provides different proxies of site response: the average shear wave velocity $V_{S,30}$ [m/s] in the uppermost 30 m is provided for all the records in the flat-file along with the corresponding estimation method (e.g. from Array, Multichannel Analysis of Surface Waves, surface geology etc.; a complete list of methods is given in the dictionary attached to the flat-file). When direct in-situ measurements of the S-wave velocity profile is not available, an estimate of $V_{S,30}$ is provided from empirical correlation with the topographic slope according to Wald and Allen (2007) on the basis of slope measurements from 90 m digital elevation model (DEM, provided by the Shuttle Radar Topography Mission). The soil categories of the National Earthquake Hazards Reduction Program (NEHRP) (BSSC, 2003) and the EC8 (Comité Européen de Normalisation [CEN], 2004) are also included. Namely, the soil classes are assigned on the basis of measured $V_{S,30}$ values, and secondly from other proxies, such as surface geological information for EC8 classification. A complete description of all the site proxies is provided in the user manual (see *Data and Resources*).

The IMs are supplied for each waveform component (two horizontals and vertical) in terms of acceleration elastic spectral ordinates SA (5%-damped) for 36 ordinates in the natural vibration period range 0.01–10s, along with peak ground acceleration (PGA), peak ground velocity (PGV), and peak ground displacement (PGD), integral IMs (Housner, Arias intensities) and significant duration (T90), i.e. the time interval between the points of 5% and 95% of the total energy (Trifunac and Brady, 1975), as well as a newly added field on Cumulative Absolute Velocity (CAV).

Other parameters are related to impulsive motions such as the pulse-like flag, the pulse period and the pulse azimuth retrieved from the identification algorithm developed after Baker (2009) and Shahi and Baker (2011) (more details can be found in Sgobba et al., 2021a). As for NESS1, the IMs of the horizontal components are reported in terms of the

fault-normal (FN) and fault-parallel (FP) components rotated with respect to the fault strike, the maximum (D100), the minimum (D00), and the median (D50) values of the ground-motion parameters that are obtained rotating the time series over all possible orientations (Boore, 2010).

All the data in the table come from accelerometric waveforms processed with a second-order acausal time-domain Butterworth filter to the zero-padded time series and zero-pad removal to make acceleration and displacement consistent after double integration (Paolucci et al., 2011). The corner frequencies of the bandpass Butterworth filter are also reported.

The NESS2-eBASCO flat-file

The 'NESS2_flat-file_eBASCO.csv' (D'Amico et al., 2021b) contains the main events, stations and waveforms metadata, which are common fields with the NESS2 flat-file. The reported IMs are 36 elastic spectral displacement ordinates (SD) in the natural vibration period range 0.01–10 s and permanent displacements of ground-motion (PDS) retrieved by the eBASCO piecewise linear detrending processing of velocity waveforms described in Schiappapietra et al. (2021). Differently from the standard processing scheme (Paolucci et al., 2011) adopted to retrieve IMs in the NESS2flat-file and described in the previous section, eBASCO does not apply any filtering to remove the low-frequency content and does not force both velocity and displacement to return to zero, thus preserving long-period spectral displacements of near-fault waveforms affected by one-side pulses in the velocity trace and static offsets at the end of the displacement. The NESS2_flat-file_eBASCO is composed of 597 three-component waveforms of 65 events, over the total earthquakes included in the NESS2 flat-file (i.e. 81). A more comprehensive analysis on SD from eBASCO processing

versus SD from the standard processing technique can be found in Schiappapietra et al. (2021).

For sake of brevity, we limit here to point out that PGDs from NESS2-eBASCO are in general larger than PGD estimated from standard processing (i.e. the PGD values contained in NESS2 flat-file), depending on the style of faulting (greater bias occurs in dip-slip mechanisms than strike-slip) and PGD value. This is also consistent with the results by Sung et al. (2021), who found that PGD obtained including the permanent tectonic displacement can be five times larger than the bandlimited PGD. Conversely, there are not significant differences in terms of PGAs as such values are related to higher frequencies, while small biases can be found on PGV. In terms of SD, the eBASCO processing scheme provides larger spectral ordinates compared to the standard approach.

We show in Figure 4 some examples of ground displacement time-series obtained with both the processing schemes (i.e. eBASCO vs NESS2 standard algorithm) compared to coseismic displacements recorded by nearby Global Positioning Systems (GPS), as well as related SD (5% damped). Waveforms corrected with eBASCO algorithm diverge substantially from the traces corrected with standard processing, being more consistent with the geodetic measurements, both in terms of amplitude and sense of movement. The corresponding spectra may be also significantly different, especially for periods longer than ~4 s (Schiappapietra et al., 2021), depending both on magnitude and faulting mechanism, thus demonstrating that the spectral displacement estimation is highly sensitive to the processing method in the long-period range.

To test the effectiveness of the eBASCO estimates, we also compare fling-displacements against nearby measurements from GPS or simulated data. Results and reference studies used for this comparison are available in the table of supplement S1. The latter includes coseismic displacements from GPS data recorded at a maximum distance of around 10 km

from the accelerometric stations. Further comparisons are shown in the supplementary file (supplement S2).

Source-to-site distances

Since information on fault geometries in NESS2 is available for each event, either from source models or virtual faults, it was possible to compute the most used source-to-site distance metrics for GMMs and distribute them in the flat-file. The considered metrics are: 1) the epicentral distance (R_{EPI}); 2) the closest distance to the surface projection of the fault rupture plane, Joyner-Boore distance (R_{JB}); 3) the closest distance to the fault rupture plane (R_{RUP}); 4) the hanging/footwall distances (R_X and R_{Y0}), being R_X the horizontal distance measured perpendicularly to the fault strike, from the top edge of rupture plane and R_{Y0} the horizontal distance off the surface projection of rupture plane measured parallel to the fault strike; 5) the distance from the nucleation point (R_{NP}), being the nucleation point (NP) the starting point of the rupture on the fault plane (this may be different from the hypocenter when the instrumental localization provided by the catalogs fall beyond the edges of the proposed fault geometry, especially for older events; if the source model does not provide information, we locate the nucleation point at mid-length and at 2/3 of the width of the fault assuming a bilateral rupture propagation), and 6) the shorter distance from the surface projection of the top edge of rupture plane (R_{LINE}). A sketch showing the different distance metrics is provided in the user manual (see *Data and Resources*).

A novel set of information is also contained in NESS2 flat-file, which refers to the inclusion of the effective fault dimensions (i.e. the effective length and width of the rupture plane derived from the autocorrelation of the slip function; Mai and Beroza, 2000) and corresponding distances. The effective dimensions indeed may differ from the traditional estimates of the source (i.e. derived from geological evidence of rupture or from the spatial extent of early aftershocks, Mai and Beroza, 2000), due to heterogeneous distribution of the

312 fault-slip. The effective faults are taken from the online database of slip distributions of the
313 finite-fault rupture SRCMOD (Mai and Thingbaijam, 2014; see *Dataset and Resources*),
314 which is the most relevant open access archive of earthquake rupture models, distributed in
315 a unified representation. We found 8 fault models in NESS2 that significantly deviate from
316 the geological dimensions, mainly referred to events with magnitude larger than 7.0, which
317 thus are more likely affected by asymmetry in slip distribution, particularly near large
318 asperities on the rupture plane. In these cases, we calculate and report in the table the
319 corresponding distances R_{EPI_eff}, R_{JB_eff}, R_{RUP_eff} and R_{LINE_eff}, along with the effective fault
320 dimensions and the reference source model.

321 In light of the large availability of information on distance metrics in the flat-file, we explore
322 the relationship between the point-source metrics (i.e. epicentral R_{epi} and hypocentral
323 distance R_{hypo}; the latter computed as the root mean square of the sum of squares of
324 epicentral distance and focal depth) and the corresponding finite-fault ones (R_{JB}, R_{RUP},
325 R_{LINE}). This is done with the aim of providing relationships useful for converting different
326 types of input metrics, as sometimes required when combining sets of not consistent GMMs
327 in the framework of PSHA. Similar relations were also proposed by other authors (Sherbaum
328 et al. 2004; Gupta 2006).

329 In line with that studies, we develop a set of simple equations from the differences
330 between the point-source distances and the finite-fault ones available in NESS2 (e.g. R_{epi} -
331 R_{JB}; R_{epi} - $R_{LINE;}$ R_{hypo} - R_{RUP}), being the latter type of metrics always smaller or equal to the
332 former. The computed differences (i.e. residuals) mainly depend on distance and magnitude
333 (Figure 5a for the case of R_{epi} and R_{JB}), with a deviation that increases as the fault size
334 increases. A regression analysis is performed on the residuals vs magnitude to derive
335 distance conversion equations, according to the generic functional form reported in Equation
336 [1]. The best-fitting model refers to the conversion R_{epi} - R_{JB} while the most scattered is the

337 R_{epi} - R_{LINE}; more details on all the developed models are reported in the **APPENDIX** (Figure
338 1A).
339
340 $$R_{finite\text{-}fault} = R_{point\text{-}source} - [a*\exp(b*M_w)] \quad [1]$$
341
342 Figure 5b shows some fitting relations for R_{epi} vs R_{JB} along with the data points, marked
343 differently according to the corresponding magnitude range. The obtained curves show an
344 average trend in agreement with the ones obtained by Sherbaum et al. (2004), although
345 these latter tend to overestimate finite-fault distances compared to this study. The observed
346 bias may be due to the fact that Sherbaum et al. calibrated conversions on simulated data,
347 thus their results may be more affected by model assumptions compared to real
348 observations.
349 However, it must be noted a huge scatter of the data at higher magnitudes, thus
350 suggesting that for very large ruptures different scaling laws could hold. Indeed, at large-
351 magnitudes the effect of the style of faulting becomes relevant, as exhibited by the rhombus
352 markers in the Figure. Indeed, as expected by geometrical considerations, large-magnitude
353 events with strike-slip mechanisms are more scattered as in general they are associated
354 with a Joyner-Boore distance that is smaller (for sites located along the strike of the fault) or
355 greater (for sites located perpendicularly to the strike of the fault), with respect to dip-slip
356 faults originated in the same epicenter and with the same magnitude. Therefore, in such
357 cases (M>7.5), we note that the proposed empirical relations may be inadequate for practical
358 applications and that an explicit dependence on the style of faulting should be considered.
359 Similar curves are obtained for the other two conversion relations (i.e. R_{epi} - $R_{LINE;}$ R_{hypo} -
360 R_{RUP} - figures not reported). The other conversion relationships calibrated on finite-fault
361 metrics are reported in Table 1A in the **APPENDIX**.
362

NEAR-SOURCE FEATURES IN NESS2

NESS1 flat-file has shown some peculiar features of ground-motion in near-source regime, as discussed in Pacor et al (2018b); we repeat here those analyses on NESS2 in order to verify the trends already observed, on the extended dataset.

Figure 6a) and b) show that the logarithmic ratios of vertical components (V) to RotD50 (D50) for PGA are often larger than 0 (i.e. V/D50>1) at R_{JB} roughly equal to zero, thus over the surface projection; then, the ratio decreases as the distance increases. As expected, the vertical ground-motion presents most of its energy in a narrow high-frequency band (Collier and Elnashai, 2001) as evidenced by the comparison between the trend at PGA (a) and at long-period SA(T=3s). In the latter case, indeed, the trend is the opposite, so the ratio increases with distance, likely as a consequence of surface waves generation.

Regarding ground motion polarization effects, we observe that the logarithmic FN/FP ratio is almost constant near 0 at all distances as shown in Figure 6c for PGA, while is larger at T=3s (Figure 6d) for very short distances (R_{JB}<1 km), probably due to the rupture-directivity effects, which typically occur at long-periods, although the scatter is quite large.

Directionality effects are also investigated by means of the index D100/D50 (computed as the ratios between the spectral acceleration SA maximum component RotD100 and the median RotD50, of each observation) vs R_{JB} distance to estimate the deviation of the exceptional values with respect to the median (Figure 6e-f). It can be seen that the maximum SA value over all possible directions at a given period is greater than the median of about 1.2-1.3 times on average, with higher values at long-periods (T=3s; Figure 6f). The ratio is almost constant at all distances but may reach larger values under 1-2 km.

The systematic increase in ground-motion for sites on the Hanging-Wall (HW) side of the fault rupture when compared to sites on the Footwall (FW) side at equal source-to-site distance (the so-called “Hanging-Wall effect”) is evidenced by an asymmetric scatterplot of the NESS2 data with R_X, for dip-slip events (Figure 6g-h). Larger values of motion are

389 recorded on the HW side ($R_X \geq 0$), even more than 25% in the range 0-5 km, with respect to

390 the FW side ($R_X < 0$) at the same distance on PGA (Figure 6g). In contrast with the results

391 obtained on NESS1, we note that the effect persists also at longer distances (about 40% of
392 the increase is still observed at 40 km). This scaling behaviour may depend on the fact that
393 the longest distances in the plot correspond to near-source distances of events with
394 magnitude greater than 6.5, according to the seismological criterion applied to select the
395 data in NESS2 so that a scaling with distance is difficult to appreciate. The HW effect
396 disappears at long-periods (e.g. T=3s, Figure 6h). A similar trend is also observed for the
397 vertical component, not reported for brevity.
398 All the considered parameters in Figure 6 do not show a relevant dependence on the style
399 of faulting. We only note a modest increment of the main trend in Figure 6b in terms of V/D50
400 ratios for TF type at long-periods (circular marker), which lie in the upper bound of the whole
401 dataset distribution. This evidence is consistent with other studies on vertical to horizontal
402 ratio (V/H), which found more relevant effects on TF mechanisms (Bozorgnia and Campbell,
403 2016; Ramadan et al., 2021). The observed trend is also consistent with the predictive V/H
404 models proposed by Bozorgnia and Campbell (2004) and Gülerce and Abrahamson (2011),
405 which suggest the dominance of normal or thrust V/H ratios in the long-period range.

406 Finally, we observe that the empirical cumulative density functions of the ground-motion
407 parameters in NESS2 follow substantially the same distributions reported by Pacor et al.
408 (2018b) for NESS1 (Figure 7a - b). In detail, the values corresponding to the 98th percentile
409 of RotD100 in NESS2 are 982 cm/s^2 and 111 cm/s, respectively for PGA and PGV, which
410 are less than 10% different with respect to the corresponding values in NESS1. Ground
411 motion levels however appear slightly smaller than the corresponding ones in NESS1, likely
412 due to the addition of a dominant number of records from the Chi-Chi (Taiwan) 1999-09-20
413 earthquake, whose PGA were between 0.5 and 0.7 times the worldwide average for soil and

414 rock sites, respectively (D'Amico et al., 2012). On the contrary, the observed horizontal PGV
415 values for this earthquake were found relatively high when compared with predictions for an
416 earthquake of the same magnitude (Tsai and Huang 2000). Nonetheless, we find that the
417 addition of the 352 records of Chi-Chi, while being substantial, does not introduce a
418 significant bias in the distribution of the ground-motion parameters (Figures 7, c-d) as it can
419 be seen that the RotD100 distribution of the Chi-Chi records for PGA (Figure 7c) and SA(1s)
420 (Figure 7d) does not affect the whole distribution. Indeed, the marked points lie in the main
421 trend of the rest of the data and fully contained in the whole variability, so we can consider
422 the NESS2 dataset sufficiently stable and suitable for the calibration of GMMs and correction
423 factors.
424
425 **Pulse-like waveforms**
426 The identification of pulses in the velocity traces is a key-point to fully assess the implications
427 of near-source ground-motions on engineering applications. In fact, pulse-like waveforms
428 exhibit acceleration response spectra with larger narrow-band amplitudes and unusual
429 shapes that strongly affect the seismic response of structures (Bray and Rodriguez-Marek,
430 2004; among the others). For the investigation of such effects within the NESS2 dataset, we
431 apply the identification algorithm modified after Shahi and Baker (2011) (Iervolino and
432 Baltzopoulos, 2020). The latter allows identifying the full-cycle pulses in the velocity time
433 series mostly due to directivity-induced effects, although the pulse-like flag is attributed
434 solely on the basis of the signal characteristics and it is not intended to distinguish among
435 causal effects. The procedure is largely adopted in the literature and is based on the use of
436 a continuous wavelet transform of a single-component's velocity time-history to extract the
437 pulse from the horizontal orientations of each station's record. A discussion on the algorithm
438 details and implementation within the NESS data is given by Sgobba et al. (2021a).

The systematic application of the identification procedure to NESS2 waveforms led to a subset of 312 records tagged as impulsive (corresponding to about 26% over the whole set of processed waveforms), along with the corresponding pulse period Tp (i.e., the pseudo-period of the wavelet extracted from the velocity signal) and pulse azimuthal orientation α, which are included in the flat-file.

Analysis and comments made on the pulse-like waveforms and related spectra for NESS1 data by Sgobba et al. (2021a) remain valid also on the extended dataset NESS2. However, we update the magnitude vs pulse period relationship (Equation [2], Figure 8a) and compare it with the available relations by Sommerville (2003) and Baker (2007; 2009), as well as with the relation calibrated on some earthquakes with finite-fault data in NGA-West2 database (Spudich et al., 2013).

$$T_p = \exp(0.9212*M - 5.244) \qquad [2]$$

The Tp-M relation here developed is similar to those from previous studies in the literature, although closer to the predictions provided by Baker (2007), likely because the methods used to extract the pulses from records are practically the same. The variability of Tp is still high, as for the previous studies, especially for large magnitudes. A relevant source of variability indeed may be related to the methods adopted to extract the pulses, which generally do not distinguish between fling and directivity-related pulses. This evidence may also suggest that the magnitude alone is not enough in constraining the predictive relationships of Tp.

Eq. [2] shows a comparable statistical coefficient of determination (R-squared) with respect to the model previously calibrated on NESS1 and reported by Sgobba et al. (2021) (i.e. 52% vs 49%), indicating that the provided relationship is stable.

Further, we find a positive correlation between the pulse period and the ratio PGD/PGV computed with eBASCO processing (Equation 3, Figure 8b).

$$T_p = 4.8494*(PGD/PGV) \quad [3]$$

The model in Eq. [3] has a comparable goodness of fit (R-squared = 53%) with respect to Tp-M model. The PGD/PGV ratio thus appears a good proxy to predict Tp as it links to physical features of the rupture process, such as directivity, although, it also shows a considerable scatter at large ratios.

Permanent displacements

As previously discussed, data on tectonic permanent displacement (PDS) are computed in NESS2-eBASCO on a subset of the NESS2 waveforms. As before stated, in some cases it was possible to compare the PDSs to nearby GPS measurements of different moderate-to-large worldwide earthquakes to evaluate the effectiveness of eBASCO, finding a general agreement between accelerometric records and geodetic measurements.

The distribution of PDSs in NESS2-eBASCO with magnitude and style of faulting (SOF)is depicted in Figure 9a for RotD100 components. It appears that a bi-linear relationship holds between the PDS and magnitude that increases with magnitude increasing up to about Mw 6.5 and decreases for larger magnitudes. Distance scaling with R_{JB} (Figure 9b) in terms of RotD100 indicates a generally rapid decay of PDS at all magnitude ranges, but with a large variability particularly in the range 5.5-6.0. Note however that the PDSs smaller than about 4-5 cm may be more affected by uncertainty in the estimation processing due to high noise levels, thus contributing to a more scattered trend.

The main findings by Schiappapietra et al (2021) on the same dataset also show a dependence of PDS scaling on the focal mechanism, fault-dip and distance from the fault

plane R_X, which are useful observations to develop a novel attenuation model of fling-displacements.

CALIBRATION OF A NEAR-SOURCE CORRECTION FACTOR FOR SA

The goal of this section is to propose an example of use of the NESS2 data to calibrate a correction factor of the median predictions of a GMM. Indeed, a possible strategy to account for near-source effects in GMMs is to introduce empirical or parametric adjustment in their functional forms. The approach here adopted implements the method proposed by Atkinson (2008; 2010) known as “Referenced Empirical Approach”, by performing a regression analysis on the residuals of the reference GMM with respect to NESS2 data.

To this aim, we select the regional model developed by Lanzano et al. (2019) for shallow active crustal earthquakes in Italy, shortly named ITA18, which predicts acceleration linear response spectra SA in the period range 0-10s. The choice of a regional model is not arbitrary, since these models are typically constrained by few records of strong earthquakes in near-source conditions. As a matter of fact, ITA18 dataset is composed of less than 7% of records at distances shorter than 20 km. On this model, we apply a residual analysis by selecting the data of NESS2 with magnitude up to 7.5 for consistency purposes with the range of validity of the italian model.

Figure 10 represents the residuals of ITA18 against NESS2 (i.e. 10-base logarithmic difference between NESS2 observations and ITA18 predictions) with reference to PGA, as a function of M_W (Figure 10a), R_{JB} (Figure 10b) focal mechanism (Figure 10c) and shear-wave velocity, V_{S30} (Figure 10d). It can be noted a clear dependence of the residuals by all the considered parameters, which represent the explanatory variables usually considered for the calibration of empirical GMMs. The same plots for SA(T=1s) are available in the supplement S3. It is worth to be noted that the observed patterns follow on from the behaviour of the reference model adopted; indeed they just reflect the difference between

516 the observed NESS2 data and ITA18 and thus should not be interpreted in the way we
517 usually expect the scaling with the explanatory variables (distance, magnitude, focal
518 mechanisms, Vs30). As a matter of fact, the magnitude scaling of ITA18 shows a more
519 pronounced saturation at magnitudes larger than 6.7 with respect to that observed on
520 NESS2 data, both for short and long-periods (Lanzano et al., 2019). Conversely, at
521 magnitudes smaller than 6.7, the model tends to underestimate the observed ground motion.
522 This clearly indicates the need to introduce a different slope of the correction in the two
523 magnitude ranges and to adopt a bi-linear scaling in the functional form of the correction
524 term with an hinge magnitude equal to 6.7.

525 For similar reasons, the distance scaling is featured by large positive residuals,
526 particularly in the range 10-20 km, while decreasing at longer distances. In terms of site
527 dependency, the residuals versus $V_{s,30}$ at short-periods (Figure 10d) show a slightly negative
528 bias (observations lower than ITA18 predictions) for $V_{s,30}$ lower than 200 m/s where besides,
529 few data are available. For larger $V_{s,30}$ the average trend increases very weakly.

530 In light of above and in order to not distort the intrinsic features of the original model (ITA18)
531 and its predictive power in ranges that are not sufficiently sampled by the NESS2 data, we
532 adopt the strategy of simplifying the dependencies on the explanatory variables as much as
533 possible. Therefore, we include in the regression only the dependence on magnitude and
534 distance, thus neglecting the small bias observed with $V_{s,30}$ for the scope of this work. As a
535 result, we set the following functional form to model the residuals:

536

537 $$\delta c(T) = a_R + F_M(M_W, SOF) + F_D(M_W, R) + dBe + dS2S + dW_0 \quad [4]$$

538

539 where a_R is the offset, while F_M and F_D are the magnitude and distance scaling
540 respectively, which are computed as:

$$F_M(M_W, SOF) = b_R(M_W - M_{ref}) + f_{jR}\, SOF_{jR} \quad [5]$$

$$F_D(M_W, R) = c_R \log_{10}(R_{JB}^2 + h^2_R)^{0.5} \quad [6]$$

In equation [5], b_R is a positive value for M<=6.7 while is zero for M>6.7. This means that the magnitude scaling term is forced to be zero when negative, so that the ITA18 model corrected with NESS2 (NESS-ITA18 hereafter) always amplifies the predictions with respect to the uncorrected ITA18, thus in order to provide safe predictions over all scenarios in near-source conditions. SOF_{js} are dummy variables for SS (j=1), TF (j=2), and NF (j=3); M_{ref} is set to 6.7. The terms b_R, c_R, f_{jR} are the model coefficients, which are provided in the supplementary material (S4), for each intensity measure (PGA, PGV and SA at periods from 0.01 to 10s) along with the corresponding statistical significance (e.g. coefficients with p-Values <0.05 are considered to be significant, while coefficients with p-Values >0.05 are considered to be less significant).

We perform the regression in two steps: in the first one, we compute the earthquake pseudo-depth through a nonlinear regression; in the second step, we perform a linear ordinary least-squares mixed-effect regression (Bates et al., 2015) to obtain the model coefficients and the corresponding uncertainty. In the latter step, the random-terms dBe, dS2S and dW_0 of equation [4], which represent respectively, the between-event, the site-to-site and the remaining aleatory residuals, according to Al-Atik et al. (2010) are estimated, together with the corresponding standard deviations denoted by τ, ψ and ϕ_0.

The obtained relation allows to correct the ground-motion model for NESS2 data, by logarithmically combining the ITA18 median predictions (SA_{ITA18}) with the correction factor δ_c, according to the following equation:

$$\log_{10}(SA_{ITA18\text{-}NESS}) = \log_{10}(SA_{ITA18}) + \delta_c \quad [7]$$

For conservative purposes, the factor delta δ_c is used to correct ITA18 only if positive.

Figure 11a shows the final prediction $SA_{ITA18\text{-}NESS}$ along with SA_{ITA18} for different magnitude levels (M_w 5.5 - 6.5 - 7.5) and R_{JB} distances (5 -15 -30) km in Figure 11b, as well as a function of the SOF (Figure 11c) for M_w7.0.

It can be noted that near-source correction provides a more significant contribution to the spectral amplification with increasing magnitude and decreasing distances, mainly at short-medium periods. A negligible effect is observable for M_w=5.5 and distances equal or larger than 30 km, rather the introduction of the correction factor to ITA18 appears to be more relevant for TF and NF.

Hence, the correction effect is not meaningful at all scenarios and periods, likely because the near-source features contained in NESS2 arise as smeared and counterbalanced effects in the correction, resulting in a slight average amplification.

The scenario-dependence of ITA18-correction in terms of SA can be noted in Figure 12 which shows contour plots in terms of M_w and R_{JB} for SA at the period T=0.2s, where the effect of the correction is maximized. As expected, higher levels correspond to large-magnitudes and short-distances, and to TF mechanisms for this period (Figure 12b). One can also note that the increasing of the spectral ordinates may be relevant: as an example, a scenario event TF-type with magnitude 7.0 above the fault projection should be corrected with an additional acceleration demand larger than 250 cm/s^2 with respect to ITA18 median prediction. This effect decreases with distance, wiping out at about 30 km from the fault. Slightly lower corrections are necessary for NF types (Figure 12a), while SS faulting types no need corrections of ITA18 median predictions, at all spectral periods; indeed where the correction is negative (as for SS faults, see Figure 10c) we forced the values to be zero for conservative purposes.

Linked to these trends of the median correction, we also observe a relatively high level of the total standard deviation, which has almost the same order of magnitude of the ITA18 one (a table of the standard deviations at each period is provided in the supplement S5). We interpret this evidence as the expression of all the near-source effects not fully captured by the median correction and that reflect in the various components of variability. In light of this, we explore the uncertainty terms of the correction model depending on magnitude and distance, thus investigating potential trends.

Figure 13 shows the remaining aleatory uncertainty ϕ_0 of ITA18-NESS as a function of distance R_{JB} for different spectral ordinates, as well as the standard deviation of the between-event term τ versus M_w. The latter shows a significant trend (Figure 13a), which increases with magnitude increasing. This finding is consistent with the expected higher variability of strong-events, potentially related to the increasing complexity of large rupture sources, although it opposes to previous findings focused on far-field data (e.g. Abrahamson and Silva, 2008; Campbell and Bozorgnia, 2014; Lanzano et al., 2019). These studies found indeed that the values of τ for small-events are generally greater than for large-events.

The uncertainty of the remaining residuals ϕ_0 does not show any particular trend with distance (Figure 13b) at the considered spectral periods, confirming that the cutting on distance in NESS2 is effective and that all the included records result in a near-field regime, exhibiting homogeneous variability features. As a general remark, we observe that though the aleatory uncertainty ϕ_0 is relatively low, the various contributions to the total standard deviation σ in ITA18-NESS have a comparable order of magnitude of the regional model ITA18, that increases with period increasing (Figs. 13c-d).

CONCLUSIONS AND FINAL REMARKS

We have presented a new version of the worldwide NESS flat-file, which includes additional data, covering a wide range of magnitude and distances, as well as further metadata related to events and sources, such as stress-drop and seismic moment. The flat-file has also been added with new IMs related to displacement spectral ordinates and fling-step amplitudes provided in a novel add-on table of the parameters retrieved from piecewise linear detrending of the velocity time series (eBASCO). The fling displacements have been compared against GPS measurements, wherever available from nearby stations or from other sources of data, such as simulations of coseismic displacements, demonstrating that, for the set of comparisons, eBASCO scheme is effective in retrieving the tectonic fling-step.

On this renewed and updated base of data, we investigated some major statistics, leading to recognize the most relevant properties of near-source records. The main observations are the following:

- differences with NESS1 in terms of ground motion distribution are modest (less than 10% in terms of 98° percentiles in RotD100 for PGA and PGV). In general NESS2 shows cumulative distributions of ground motion intensity measures that are similar to those of NESS1, even if they appear generally lower at short-periods. This is due to the addition of a dominant number of records from Chi-Chi (Taiwan) 1999-09-20 event (352 records), which however do not introduce a bias in ground motion distribution and therefore does not affect the stability of the whole dataset;

- About 30% of waveforms are featured by pulses in the accelerometric time series. The pulse period was found not only magnitude-dependent, as pointed out by previous studies (Sgobba et al., 2021a), but also positively correlated to the ratio between PGD and PGV;

639 ● The ratio between the vertical components to the RotD50 horizontal
640 components is close to 1 over the rupture fault at short periods, whereas it decreases
641 to about the half at long-periods;
642 ● No relevant differences between FN and FP components are detected at
643 short-periods. Rather, predominant FN components are observed at longer periods
644 for very short distances (R_{JB}<1 km), probably due to the rupture-directivity effects;
645 ● The maximum horizontal acceleration spectral ordinates (RotD100) over all
646 possible directions is greater than the median (RotD50) of about 1.2-1.3 times on
647 average;
648 ● The data exhibit Hanging-Wall effects on dip-slip events. Larger acceleration
649 spectral ordinates are recorded on the HW side in the range 0-5 km, with respect to
650 the FW side (more than 25%) at the same distance at short-periods. The effect
651 persists at longer distances (about 40% of increase at 40 km) while vanishes at long-
652 periods;
653 ● The eBASCO scheme is suitable to preserve the low-frequency content of the
654 displacement time-series and thus it enables to recover the tectonic flings effectively.
655 Instead, the application of a high-pass filter could lead to a reduction of the spectral
656 amplitudes mainly from periods longer than ~4 s, depending both on the magnitude
657 and faulting mechanism;
658 ● eBASCO estimates of permanent displacements show a correlation with
659 magnitude and a scaling with the distance from the fault plane that depends both on
660 the focal mechanism and fault-dip, as also observed by Schiappapietra et al (2021)
661 on the same dataset.

662 As a more general remark, we point out that the attempt to compiling and maintaining the
663 NESS2 flat-files is intended to increase the knowledge of ground-motion behaviour in near-

source, not only just through the dissemination of new data in epicentral area, but also via a qualified and homogeneous base to developing new products and models for design purposes. Indeed, our main idea is to provide uniformly processed and high-quality data and metadata to allow calibration of a novel set of scaling laws and predictive models to serve the seismological and engineering community with fully consistent and reproducible relationships.

As a result, the efforts made so far on NESS have produced some developments in near-source ground-motion research, such as the empirical correction term of a GMM to account for the pulse-like records (Sgobba et al., 2021a), the adjustment factor for the Vertical-to-Horizontal (V/H) spectral ratios of crustal shallow earthquakes in Italy (Ramadan et al., 2021), as well as the validation of a dataset of broadband ground-motions (BB-SPEED set) from 3D physics-based simulations (Paolucci et al., 2021).

In line with those studies, we herein provided an example of calibration of a correction factor useful to adjust an existing well-constrained italian attenuation model (ITA18). We have chosen to follow the strategy of developing a correction factor of an existing model instead of a new one, specifically calibrated on NESS2 data, to avoid the issue of introducing an over-complexity in the functional form, due to the inclusion of additional parameters, typically related to source-site geometry and rupture process (Mai, 2009), that are often scarcely available or affected by large uncertainty. Also, the calibration of an *ad-hoc* model would have suffered from being poorly constrained at longer distances, owing to the dataset cutoff. In this respect, a potential strategy would be also to supplement the worldwide dataset with NESS2 data, in order to provide a better constraint at higher magnitudes and near-fault distances in both empirical and physics-based ground motion modelling.

Finally, we point out that the correction here proposed does not aim to model each near-source peculiar phenomenon like directivity, pulse-like etc., although it incorporates all these

effects in the model variability. Instead, it should be intended as an *a-posteriori* empirical adjustment to use for increasing the median predictions of specific GMMs that are affected by a bias due to paucity of large-magnitude and short-distance records in their calibration dataset. In such cases, the correction enables to improve the predictive power of the biased attenuation model through a conservative increment that covers the global near-source effects recognised in NESS2. In our opinion this approach has an application-oriented potential such as for deterministic scenario's studies, particularly at short-medium periods and for scenarios dominated by large-magnitudes and short-distances from known source zones.

Ongoing research aims to complement this correction model also with an adjustment factor of the displacement spectra predicted by existing equations, in order to account for the effect of eBASCO processing on the long-period range, as well as to calibrate a new model of the fling-step amplitudes.

Data and Resources

The NEar-Source Strong-motion flat-file v2.0 (NESS2) and version v1.0 (NESS1) are available at this URL: http://ness.mi.ingv.it. The NESS2 user manual can be downloaded at this link: http://ness.mi.ingv.it/download/NESS_flat-file_2020_user_manual.pdf. The reference dataset for the finite-source rupture model is the database SRCMOD (Mai et al., 2014; http://equake-rc.info/SRCMOD/). Accelerometric time series were obtained from different online databases: the European Integrated Data Archive (EIDA, https://www.orfeus-eu.org/data/eida/), the GeoNet seismic catalogue (https://www.geonet.org.nz/) for New Zealand, the strong-motion seismograph networks of the National Research Institute for Earth Science and Disaster Resilience

713 (http://www.kyoshin.bosai.go.jp/) for Japan, the Unified HEllenic Accelerogram Database
714 (HEAD, http://www.itsak.gr/en/head or http://accelnet.gein.noa.gr) for Greece, the ITalian
715 ACcelerometric Archive (ITACA, http://itaca.mi.ingv.it) for Italy, the National strong-motion
716 Network of Turkey (TR-NSMN, http://kyhdata.deprem.gov.tr/2K/kyhdata_v4.php), the
717 Strong-motion virtual data center (http://strongmotioncenter.org/), the U.S. Geological
718 Survey (https://earthquake.usgs.gov/), the California Geological Survey
719 (http://www.quake.ca.gov/), and the Engineering strong-motion database (ESM,
720 https://esm-db.eu). Cited dataset including near-fault data are: the Southern California
721 Earthquake Data Center (SCEDC) database https://scedc.caltech.edu/. International
722 Federation of Digital Seismograph Networks. https://doi.org/10.7914/SN/CI); the
723 Consortium of Organizations for Strong Motion Observation Systems (COSMOS)
724 http://db.cosmos-eq.org, the K-NET (Kyoshin Network) and KiK-net (Kiban Kyoshin
725 Network) of Japan https://www.kyoshin.bosai.go.jp/ and the Taiwan network at
726 https://scweb.cwb.gov.tw/en-us/earthquake/data, the Next Generation Attenuation
727 relationships for Western U.S. database (Next Generation Attenuation-West2 Project are
728 available at https://peer.berkeley.edu/thrust-areas/data-sciences/databases). All websites
729 were last accessed in February 2021.

730 A supplemental material is provided for this article that includes a table of the permanent
731 displacements recovered from NESS2-eBASCO dataset and coseismic displacements from
732 GPS data (supplement-S1), further comparisons from other validation data (supplement-
733 S2), some figures related to the total residuals of NESS2 with reference to ITA18 model at
734 SA(T=1s) in supplement-S3, as well as some tables containing the regression coefficients
735 of the correction model for ITA18 (supplement-S4) and the corresponding standard
736 deviations (supplement-S5).

737

Acknowledgements

This study has been developed within the research programs INGV-ReLUIS (Rete dei Laboratori Universitari di Ingegneria Sismica) in the framework of DPC-ReLUIS Agreement 2019–2021 “Contributi Normativi—Progetto Azione Sismica (CONPAS)” WP18, funded by the Presidenza del Consiglio dei Ministri—Dipartimento della Protezione Civile (DPC), Italy. The authors are grateful to Georgios Baltzopoulos and Iunio Iervolino (Università degli Studi Federico II di Napoli, Italy) for having identified the pulse-like records within the NESS2 dataset. The authors also thank the Project Coordinator Roberto Paolucci for supporting and encouraging the development of this work.

This study has been partially funded by INGV-DPC agreement B1 2019-2021, with the aim of promoting research activities in the field of seismic hazard in Italy. Lucia Luzi is acknowledged for her support in this research and for the coordination of the activities related to the european accelerometric database.

References

Abrahamson N.A. (2001). Incorporating effects of near fault tectonic deformation into design ground-motions. Presentation sponsored by the Earthquake Engineering Research Institute Visiting Professional Program, hosted by the State University of New York at Buffalo, 26 Oct. 2001, available at http://mceer.buffalo.edu/outreach/pr/ abrahamson.asp.

Abrahamson, N., and W. Silva (2008). Summary of the Abrahamson & Silva NGA ground motion relation*. Earthquake Spectra*. **24**:1, 67-97.

Akkar, S., S., Moghimi, and Y. Arıcı, (2018). A study on major seismological and fault-site parameters affecting near-fault directivity ground-motion demands for strike-slip faulting

761 for their possible inclusion in seismic design codes, *Soil Dynamics and Earthquake*
762 *Engineering*, **104**, 88-105, ISSN 0267-7261.

763 Al-Atik, L., N. A. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn
764 (2010). The variability of ground-motion prediction models and its components, *Seismol.*
765 *Res. Lett.* **81**, no. 5, 794 – 801.

766 Alavi B. and H. Krawinkler (2004). Strengthening of moment-resisting frame structures
767 against near-fault ground motion effects. *Earthq. Eng. Struct. Dyn.*, **33**, 707-722.

768 ASCE-7.16: Minimum design loads for buildings and other structures. American Society
769 of Civil Engineers.

770 ATC 1996: Seismic evaluation and retrofit of concrete buildings. Report ATC-40, Applied
771 Technology Council, Redwood City, U.S.A. (also Report SSC 96-01, Seismic Safety
772 Commission, State of California, Sacramento, U.S.A.).

773 Atkinson, G. (2008). Alternative ground-motion prediction equations for eastern North
774 America from a Referenced Empirical Approach: Implications for epistemic uncertainty. *Bull.*
775 *Seism. Soc. Am.*, **98**, 1304-1318

776 Atkinson, G.M. (2010). Ground-motion prediction equations for Hawaii from a referenced
777 empirical approach. *Bull. Seismol. Soc. Am.* **100** (2), 751-761.

778 Bahrampouri M, A. Rodriguez-Marek, S. Shahi, and H. Dawood (2021). An updated
779 database for ground motion parameters for KiK-net records. *Earthquake Spectra*; **37**(1):505-
780 522. doi:10.1177/8755293020952447.

781 Baker J.W. (2007) Quantitative classification of near-fault ground motions using wavelet
782 analysis, *Bulletin of the Seismological Society of America*, **97** (5), 1486- 1501.

Baker, J.W. (2009). Identification of Near-Fault Velocity Pulses and Prediction of Resulting Response Spectra. Proc. Geotech. *Earthq. Eng. Struct. Dyn*. 40975, 1–10.

Baltzopoulos G., L. Luzi, and I. Iervolino (2020). Analysis of near-source ground motion from the 2019 Ridgecrest earthquake sequence. *Bulletin of the Seismological Society of America*, **110**(4): 1495-1505.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4, *J. Stat. Software* **67**, no. 1, 1 – 48.

BHRC – PN 253 (2005) "Iranian code of practice for seismic resistant design of building," Iranian building codes and standards, third revision, (in Persian).

Bindi, D., S.R. Kotha, (2020). Spectral decomposition of the Engineering Strong Motion (ESM) flat file: regional attenuation, source scaling and Arias stress drop. *Bull Earthquake Eng* **18**, 2581–2606. https://doi.org/10.1007/s10518-020-00796-1.

Boore D. M. (2010). Orientation-independent, non geometric-mean measures of seismic intensity from two horizontal components of motion, *Bull. Seismol. Soc. Am.* **100**, 1830–1835.

Bozorgnia Y., and K.W. Campbell (2004). The vertical-to-horizontal spectral ratio and tentative procedures for developing simplified V/H and vertical design spectra, *J. Earthq. Eng.* **4**, 539-561.

Bozorgnia Y., and K.W., Campbell (2016). Vertical ground motion model for PGA, PGV, and linear response spectra using the NGA-West2 database. *Earthquake Spectra*; **32**(2): 979-1004

804 Bray, J.D., and A. Rodriguez-Marek (2004). Characterization of forward directivity
805 ground motions in the near-fault region. *Soil Dyn. Earthq. Eng.* **24**, 815–828.

806 Brune, J. N. (1970), Tectonic stress and the spectra of seismic shear waves from
807 earthquakes, *J. Geophys. Res.*, **75**(26), 4997– 5009, doi:10.1029/JB075i026p04997.

808 Campbell K.W., and Y. Bozorgnia (2013). NGA-West2 CampbellBozorgnia ground
809 motion model for the horizontal components of PGA, PGV, and 5% damped elastic pseudo-
810 acceleration response spectra for periods ranging from 0.01 to 10 sec, PEER Report
811 2013/06.

812 Campbell, K.W., and Bozorgnia, Y. (2014). NGA-West2 Ground Motion Model for the
813 Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration
814 Response Spectra. *Earthq. Spectr.*, **30**, 1087–1115.

815 Chopra, A.K., and C. Chintanapakdee (2001). Comparing response of SDF systems to
816 near-fault and far-fault earthquake motions in the context of spectral regions. *Earthquake*
817 *Engng. Struct. Dyn.*, **30**: 1769-1789. https://doi.org/10.1002/eqe.92.

818 Cito P., and Iervolino I. (2020). Rarity, proximity, and design actions: mapping strong
819 earthquakes in Italy. *Annals of Geophysics*, **63**(6), SE671.

820 Collier C.J., and A.S. Elnashai (2001). A procedure for combining vertical and horizontal
821 seismic action effects. *J. Earthq. Eng.*, **5**, 521-539.

822 CPA (2011). Seismic Design Code and Commentary for Buildings, 2011 Edition,
823 Construction and Planning Agency, Ministry of Interior Affair, Taipei, Taiwan (in Chinese).

824 D’ Amico, S., A., Akinci, and, L., Malagnini (2012). Predictions of high-frequency ground-
825 motion in Taiwan based on weak motion data, *Geophysical Journal International*, **189**, Issue
826 1, Pages 611–628, https://doi.org/10.1111/j.1365-246X.2012.05367.x

827 D’Amico, M., C., Felicetta, E. Schiappapietra, F., Pacor, F., Gallovič, R., Paolucci, R.,
828 Puglia, G., Lanzano, S., Sgobba, L., Luzi (2018). Fling Effects from Near-Source Strong-
829 Motion Records: Insights from the 2016 Mw 6.5 Norcia, Central Italy, *Earthquake. Seism.*
830 *Res. Lett.* **90**, 659–671.

831 D’Amico, M.C., E., Schiappapietra, C., Felicetta, S., Sgobba, F., Pacor, G., Lanzano, E.,
832 Russo, L., Luzi. (2021). NEar-Source Strong-Motion Flatfile from eBASCO (NESS-
833 eBASCO); Version 2.0; Istituto Nazionale di Geofisica e Vulcanologia (INGV): Milan, Italy,
834 2021. https://doi.org/10.13127/NESS.2.0_EBASCO.

835 D'Amico, M., E., Schiappapietra, G., Lanzano, S., Sgobba, and F., Pacor (2020). Fling-
836 step recovering from near-source waveforms and ground displacement attenuation models,
837 EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-22285,
838 https://doi.org/10.5194/egusphere-egu2020-22285, 2020.

839 Farajpour, Z., M., Zare, S., Pezeshk, A., Ansari, and E., Farzanegan (2018). Near-
840 source strong motion database catalog for Iran. *Arabian Journal of Geosciences*, **11**, 1-16.

841 Grimaz, S., P., and Malisan (2014). Near field domain effects and their consideration in
842 the international and Italian seismic codes. *Boll. Geofis. Teor. Appl.* **55**, 717–738

843 Gülerce Z., and N.A., Abrahamson (2011). Site-specific design spectra for vertical
844 ground motion. *Earthquake Spectra*; **27**(4): 1023-1047.

Gupta ID (2006) Defining source-to-site distances for evaluation of design earthquake ground motion. Proceedings of 13th Symposium on Earthquake Engineering, IIT, Roorkee, 18–20 December 2006; I:295–306

ICBO; 1997 Uniform Building Code: Vol.2. Whittier, California: International Conference of Building Official (ICBO), 2: 34-35.

Iervolino, I., and G., Baltzopoulos (2020) Università Federico II di Napoli, Italy. Brief technical note on pulse-like ground motion identification, available on https://esm-db.eu/#/documents/pulse_like_identification.

Iervolino I., G., Baltzopoulos, E., Chioccarelli, and A., Suzuki (2019) Seismic actions on structures in the near-source region of the 2016 central Italy sequence. *Bulletin of Earthquake Engineering*, **17**: 5429-5447.

Kaklamanos J., L.G., Baise, and D.M., Boore (2011). Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice, *Earthq. Spectra* **27**, 4, 1219–1235.

Kamai, R., N., Abrahamson, and R., Graves (2014). Adding Fling Effects to Processed Ground-Motion Time Histories. *Bull. Seism. Soc. Am.* **104**, 1914–1929

Lanzano G., L., Luzi, F., Pacor, C., Felicetta, R., Puglia. S., Sgobba, M., D'Amico. (2019). A revised ground motion prediction model for shallow crustal earthquakes in Italy. *Bulletin of the Seismological Society of America*, **109**(2), 525-540, DOI: 10.1785/0120180210.

Luzi L., G., Lanzano, C., Felicetta, M., D'Amico, E., Russo, S., Sgobba, F., Pacor, F., and ORFEUS Working Group 5 (2020). Engineering Strong Motion Database (ESM)

867 (Version 2.0). Istituto Nazionale di Geofisica e Vulcanologia (INGV).
868 https://doi.org/10.13127/ESM.2

869 Luzi L., R., Puglia, E., Russo, M., D'Amico, C., Felicetta, F., Pacor, G., Lanzano G., U.,
870 Çeken, J., Clinton, G., Costa, L., Duni et al. (2016). The Engineering Strong-Motion
871 Database: A Platform to Access Pan-European Accelerometric Data. *Seismological*
872 *Research Letters*; **87** (4): 987–997. doi: https://doi.org/10.1785/0220150278

873 Mai P.M. (2009). Ground Motion: Complexity and Scaling in the Near Field of
874 Earthquake Ruptures. In: Meyers R. (eds) Encyclopedia of Complexity and Systems
875 Science. Springer, New York, NY. https://doi.org/10.1007/978-0-387-30440-3_263.

876 Mai, P. M., and Beroza, G. C. (2000). Source scaling properties from finite-fault rupture
877 models*, Bull. Seismol. Soc. Am.,* **90**, 605–614.

878 Mai, P. M., and K. K. S. Thingbaijam (2014). SRCMOD: An online database of finite
879 source rupture models*, Seismol. Res. Lett.* **85**, no. 6, 1348–1357, doi: 10.1785/0220140077.

880 Mavroeidis G. P., and A.S., Papageorgiou (2003). A mathematical representation of
881 near-fault ground motions, *Bull. Seismol. Soc. Am.* **93**, no. 3, 1099–1131.

882 NZS 1170:5: 2004 Structural design actions, Part 5: Earthquake actions – New Zealand.
883 New Zealand Standards, Wellington, IABN 1-86975-018-7.

884 Oth, A. and A.E. Kaiser (2014). Stress release and source scaling of the 2010-2011
885 Canterbury, New Zealand, earthquake sequence from spectral inversion of ground motion
886 data. *Pure Appl. Geophys.*, Topical Volume Earthquake source physics on various scales,
887 **171**(10), 2767-2782, doi: 10.1007/s00024-013-0751-1.

888 Oth, A., H. Miyake and D. Bindi (2017). On the relation of earthquake stress drop and
889 ground motion variability*. J. Geophys. Res.*, **122**, 5474-5492, doi: 10.1002/2017JB014026

890 Pacor, F., C., Felicetta, G., Lanzano, S., Sgobba, R., Puglia, M., D'Amico, L., Luzi; NEar-
891 Source Strong-Motion Flatfile (NESS) (Version 1.0) [Data Set]; (2018a). Istituto Nazionale
892 di Geofisica e Vulcanologia (INGV): Rome, Italy. https://doi.org/10.13127/NESS.1.0

893 Pacor, F., C., Felicetta, G., Lanzano, S., Sgobba, R., Puglia, M., D'Amico, E., Russo,
894 G., Baltzopoulos, I., Iervolino (2018b). NESS1: A Worldwide Collection of Strong-Motion
895 Data to Investigate Near-Source Effects. *Seismol. Res. Lett.*, **89**, 2299–2313.

896 Paolucci, R., F., Pacor, R., Puglia, G., Ameri, C., Cauzzi, M., Massa (2011). Record
897 processing in ITACA, the new Italian strong-motion database. In Earthquake Data in
898 Engineering Seismology: Predictive Models, Data Management and Networks; Akkar, S.,
899 Gülkan, P., van Eck, T., Eds.; Springer: Dordrecht, The Netherlands; pp. 99–113. ISBN 978-
900 94-007-0151-9. (printed version); 978-94-007-0152-6 (e-book version).

901 Paolucci, R., C., Smerzini, M., Vanini (2021) BB-SPEEDset: a dataset of broadband
902 near-source earthquake ground motions from 3D physics-based simulations (submitted for
903 publication to *BSSA Special Section on Fault Displacement and Near-Source Ground*
904 *Motion Models*).

905 Puglia, R., E., Russo, L., Luzi, M., D'Amico, C., Felicetta, F., Pacor, G., Lanzano (2018).
906 Strong-motion processing service: A tool to access and analyse earthquakes strong-motion
907 waveforms. *Bull. Earthq. Eng.* **16**, 2641–2651.

908 Ramadan, F., C., Smerzini, G., Lanzano, F., Pacor (2021). An empirical model for the
909 vertical-to horizontal response spectral ratios for Italian earthquakes including near-source
910 effects (submitted for publication to *Earthquake Engineering and Structural Dynamics*).

911 Rowshandel, B. (2010). Directivity correction for the Next Generation Attenuation (NGA)
912 relations. *Earthq. Spectra.* **26**(2), 525-559, doi 10.1193/1.3381043.

913 Scherbaum F, J., Schmedes, F., Cotton F (2004). On the conversion of source-to-site
914 distance measures for extended earthquake source models. *Bull Seism Soc Am*
915 **94**(3):1053–1069

916 Schiappapietra, E., C., Felicetta, M., D'Amico (2021). Fling-Step Recovering from Near-
917 Source Waveforms Database. *Geosciences*, **11**, 67.
918 https://doi.org/10.3390/geosciences11020067

919 Sgobba, S., G., Lanzano, F., Pacor, C., Felicetta (2021a). An Empirical Model to
920 Account for Spectral Amplification of Pulse-Like Ground Motion Records. Geosciences, **11**,
921 15. https://doi.org/10.3390/geosciences1101001.

922 Sgobba, S., F., Pacor, C., Felicetta, G., Lanzano, M., D'Amico, E., Russo, and L., Luzi
923 (2021b). NEar-Source Strong-motion flatfile (NESS), version 2.0 (Version 2.0) [Data set].
924 Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/NESS.2.0

925 Shahi, S.K., and J.W Baker (2011). An empirically calibrated framework for including the
926 effects of near-fault directivity in probabilistic seismic hazard analysis. *Bull. Seismol. Soc.*
927 *Am.* **101**, 742–755

928 Smerzini, C. and R., Paolucci (2013). SIMBAD: a database with selected input motions
929 for displacement-based assessment and design–3rd release. Research Project DPC-
930 RELUIS 2013, 2010-2013. http://wpage.unina.it/iuniervo/SIMBAD_Database_Polimi.pdf

931 Somerville, P.G. (1998). Development of an improved representation of near fault
932 ground motions. In Proceedings of the SMIP98 Seminar on Utilization of Strong Ground
933 Motion Data, Oakland, CA, USA, 15 September 1998; pp. 1–20.

Somerville, P.G. (2003). Magnitude scaling of the near fault rupture directivity pulse. *Phys. Earth Planet*, **137**, 201–212.

Spudich, P., B.S.J., Chiou (2008). Directivity in NGA earthquake ground motions: Analysis using isochrone theory. *Earthq. Spectra*, **24**, 279–298.

Sung, C.-H., N. A. Abrahamson, and J.-Y. Huang (2021). Conditional Ground-Motion Models for Horizontal Peak Ground Displacement for Active Crustal Regions, Bull. Seismol. Soc. Am. XX, 1–21, doi: 10.1785/0120200299

Trifunac M.D., and A.G. Brady (1975). A study on the duration of strong earthquake ground motion. *Bulletin of the Seismological Society of America*, **65**(3), 581-626.

Tsai Y.B. and M.W., Huang (2000). Strong ground motion characteristics of the Chi-Chi, Taiwan earthquake of September 21, 1999, *Earthq. Eng. Eng. Seism*., **2**, 1–21.

Wald D.J. and, T.I. Allen (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, **97**(5), 1379-1395.

948 **Full mailing address of each author**

949

950 *Sara Sgobba*, Istituto Nazionale di Geofisica e Vulcanologia, Via Alfonso Corti 12, 581
951 20133 Milano, Italy, sara.sgobba@ingv.it

952 *Chiara Felicetta*, Istituto Nazionale di Geofisica e Vulcanologia, Via Alfonso Corti 12, 581
953 20133 Milano, Italy, chiara.felicetta@ingv.it

954 *Giovanni Lanzano*, Istituto Nazionale di Geofisica e Vulcanologia, Via Alfonso Corti 12, 581
955 20133 Milano, Italy, giovanni.lanzano@ingv.it

956 *Fadel Ramadan*, Department of Civil and Environmental Engineering, Politecnico di
957 Milano, Via Colombo, 40 - 20133 Milano, Italy, fadel.ramadan@mail.polimi.it

958 *Maria D'Amico,* Istituto Nazionale di Geofisica e Vulcanologia, Via Alfonso Corti 12, 581
959 20133 Milano, Italy, maria.damico@ingv.it

960 *Francesca Pacor*, Istituto Nazionale di Geofisica e Vulcanologia, Via Alfonso Corti 12, 581
961 20133 Milano, Italy, francesca.pacor@ingv.it

962 **Tables**

963

964 **Table I. New worldwide events included in NESS2**

EVENT ID (event date and time GMT: YYYMMDD HHMMSS)	Event name (Country)	M_w	Style Of Faulting (SOF)	#records
EMSC-20140126_0000046 (2014-01-26 13:55:41)	Kefallonia (Greece)	6.19	SS	1
EMSC-20170118_0000034 (2017-01-18 10:14:12)	Capitignano (Italy)	5.5	NF	7
EMSC-20190706_0000043 (2019-07-06 03:19:57)	Ridgecrest (US)	7.1	SS	24
INT-UT19990920_174715 (1999-09-20 17:47:15)	Chi-Chi (Taiwan)	7.58	TF	352
INT-UT19991022_021856 (1999-10-22 02:18:56)	Chi-Chi (Taiwan)	6.1	TF	30

USGS-iscgem893168 (1952-07-21 11:52:16)	Kern Country (US)	7.3	TF	1
USGS-us20005i1a (2016-04-14 15:03:46)	Kumamoto (Japan)	6.0	SS	6

965

966 **List of figure captions**

967 **Figure 1.** Map of the epicenters of the worldwide earthquakes included in the near-source
968 strong-motion flat-file (NESS2). The color version of this figure is available only in the
969 electronic edition.

970

971 **Figure 2.** (a) Magnitude vs Joyner-Boore distance scatter plot of NESS2 data and (b)
972 breakdown of records within the NESS2 dataset by bins of Joyner-Boore distance.

973

974 **Figure 3.** Schematic view of the NESS2 flat-files.

975

976 **Figure 4.** Displacement waveforms processed with eBASCO (black lines) and NESS2
977 standard algorithm (gray lines) plotted along with GPS (squares) coseismic displacement
978 values (left panels). On the right panels the related displacement spectra are shown with
979 their identification code: a) event EMSC-20161030_0000029 Norcia 2016/10/30 (Italy; CLO
980 station); b) event IT-2009-0009 L'Aquila 2009/04/06 (Italy; AQV station); c) event USGS-
981 us20005iis Kumamoto 2016/04/15 (Japan; KMM18 station).

982

983 **Figure 5.** a) Scatter plot of the residuals (Repi - R_{JB}) vs Magnitude in log-log scale; b) fitting
984 curves for the relation Repi - R_{JB} for magnitude ranges in NESS2 compared with Sherbaum
985 et al. (2004) relations. The color version of this figure is available only in the electronic
986 edition.

987

988 **Figure 6.** Distribution of data in NESS2 with distance at PGA (on the left panel) and SA at
989 period T=3s (on the right panel) with different styles of faulting (SS – Strike-Slip; TF – Thrust-
990 Fault; NF – Normal-Fault): (a-b) ratio V/D50 vs RJB; (c-d) ratio log(FN/FP) vs RJB; (e-f) ratio

991 D100/D50 vs RJB; (g-h) horizontal PGA and spectral amplitudes SA vs Rx on Hanging-Wall
992 (HW) and Footwall (FW) side. The average dashed black curve and the error bar is
993 overlapped to data. The number of records in each bin is reported on the top. The color
994 version of this figure is available only in the electronic edition.

995

996 **Figure 7.** Comparison of empirical cumulative distribution functions of NESS2 and NESS1
997 records for the PGA (solid lines) and PGV (dashed lines) in terms of: (a) maximum of the
998 rotated waveforms RotD100 (D100) and (b) vertical component (V). The reported numbers
999 indicate the peak values at the 98° percentiles; D100 distribution versus Joyner-Boore
1000 distance of Chi-Chi Taiwan 1999-09-20 records compared to NESS2 records: (c) PGA and
1001 (d) SA(T=1s). The color version of this figure is available only in the electronic edition.

1002

1003 **Figure 8.** a) Pulse period versus magnitude Eq. [2] (RMSE = 0.62; R-squared = 0.5181)
1004 compared with previous studies; b) pulse-period versus PGD/PGV ratio for different
1005 magnitude ranges, Eq. [3] (RMSE = 1.81; R-squared = 0.5285), for NESS2 pulse-like
1006 records (eBASCO processed, FN components). The color version of this figure is available
1007 only in the electronic edition.

1008

1009 **Figure 9.** Horizontal permanent displacement PDS (RotD100 components) versus a)
1010 magnitude (dots are marked based on the styles of faulting: SS – Strike-Slip; TF – Thrust-
1011 Fault; NF – Normal-Fault), and b) R_{JB} distance for the magnitude range 5.5–8.1 (dots are
1012 marked based on the magnitude bins). The color version of this figure is available only in
1013 the electronic edition.

1014

Figure 10. Total residuals δc, computed as logarithm difference between NESS2 observations and ITA18 predictions for PGA, as a function a) moment magnitude (M_W); b) Joyner-Boore distance (R_{JB}); c) style of faulting (SOF); d) average shear-wave velocity in the uppermost 30m ($V_{s,30}$).

Figure 11. Horizontal acceleration spectral predictions of ITA18 corrected with NESS2 data ($SA_{ITA18\text{-}NESS}$) compared with uncorrected ITA18 (SA_{ITA18}): sensitivity analysis with: a) Mw (fixed style of faulting TF; $V_{s,30}$=650 m/s and R_{JB} = 15km); b) R_{JB} distance (fixed style of faulting TF; M_w=6.5 and $V_{s,30}$=650 m/s) and c) styles of faulting (fixed Mw 7.0; $V_{s,30}$=650 m/s and R_{JB} = 15km). The color version of this figure is available only in the electronic edition.

Figure 12. Scenario-dependence of ITA18 correction for SA(T=0.2s) with reference to $V_{s,30}$ at 700 m/s and focal mechanisms of a) Normal-Fault type and b) Thrust-Fault type. The color version of this figure is available only in the electronic edition.

Figure 13. Scenario-dependent standard deviations: a) between-event standard deviation versus magnitude; b) event- and site- corrected residuals standard deviation versus Joyner-Boore distance; c) between-event standard deviation versus periods; d) event- and site-corrected residuals standard deviation versus periods. The color version of this figure is available only in the electronic edition.

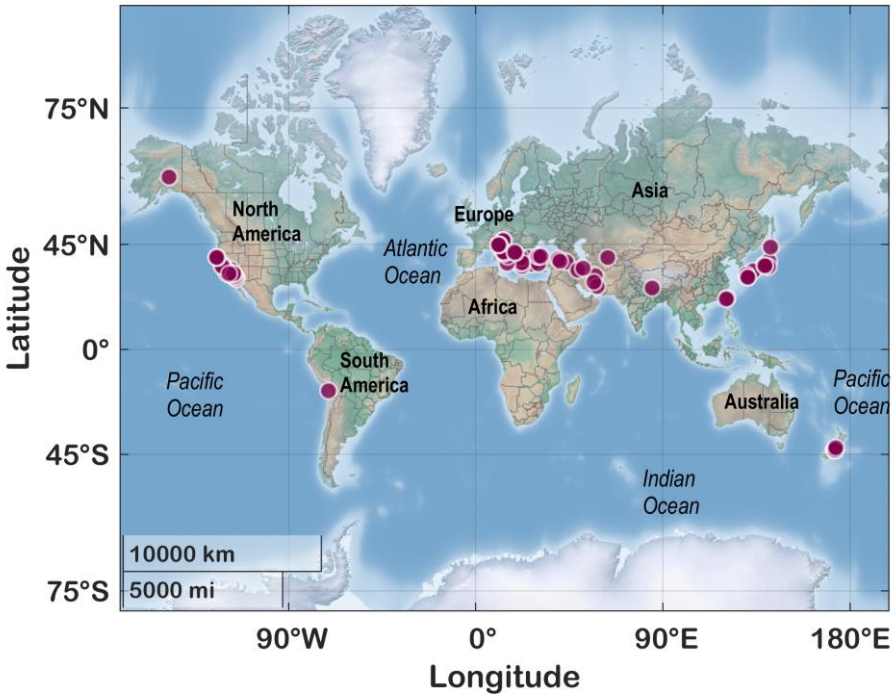
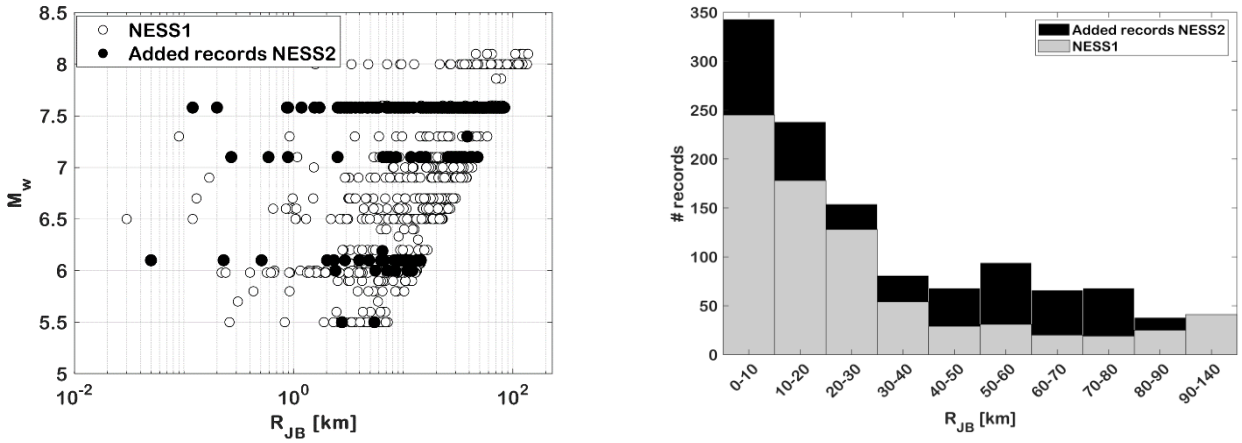


Figure 1. Map of the epicenters of the worldwide earthquakes included in the near-source strong-motion flat-file (NESS2). The color version of this figure is available only in the electronic edition.



(a) **(b)**

1040 **Figure 2. (a) Magnitude vs Joyner-Boore distance scatter plot of NESS2 data and**

1041 **(b) breakdown of records within the NESS2 dataset by bins of Joyner-Boore distance**.

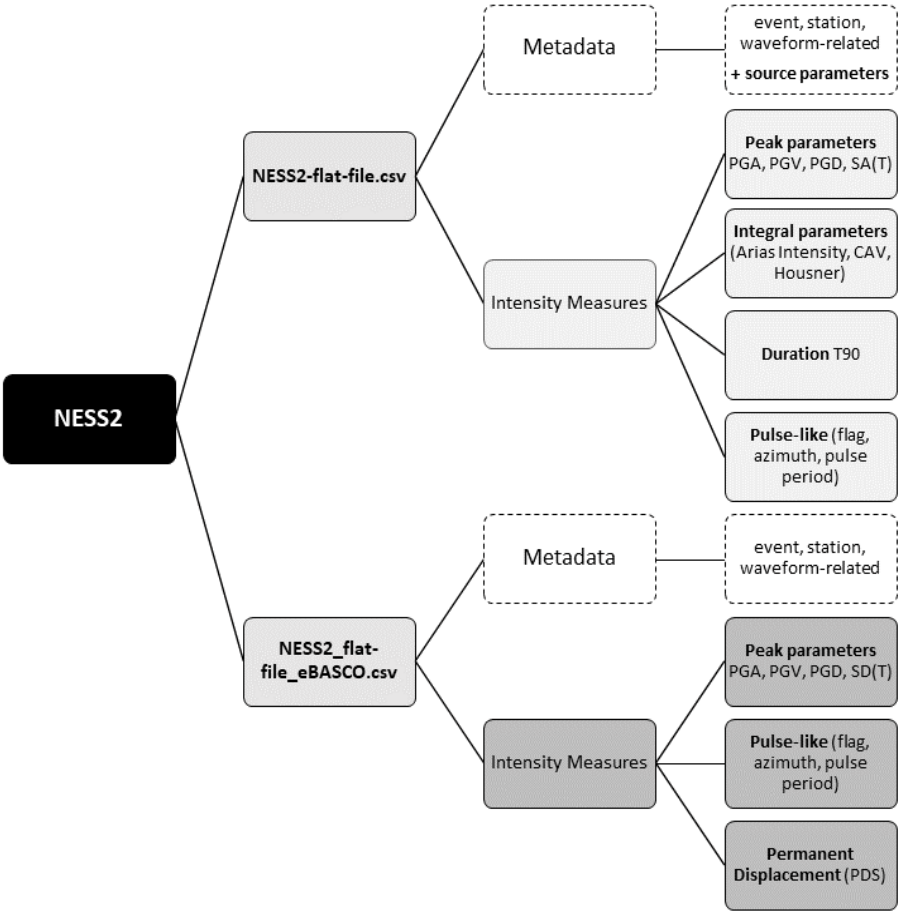
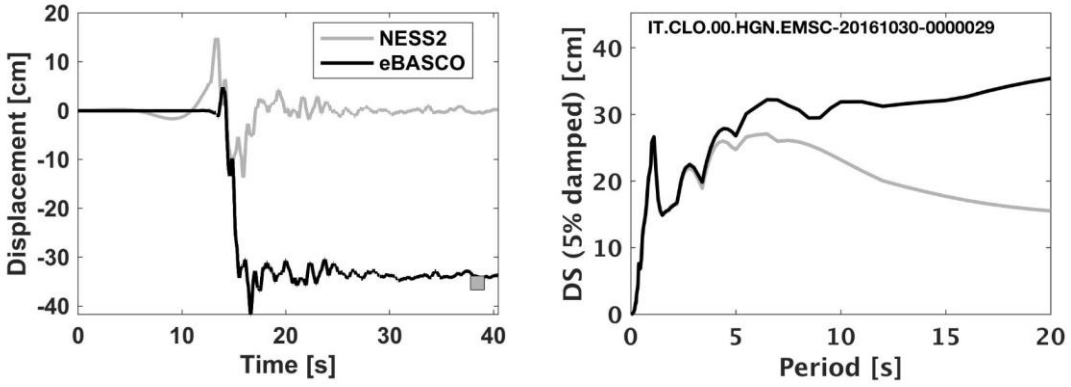
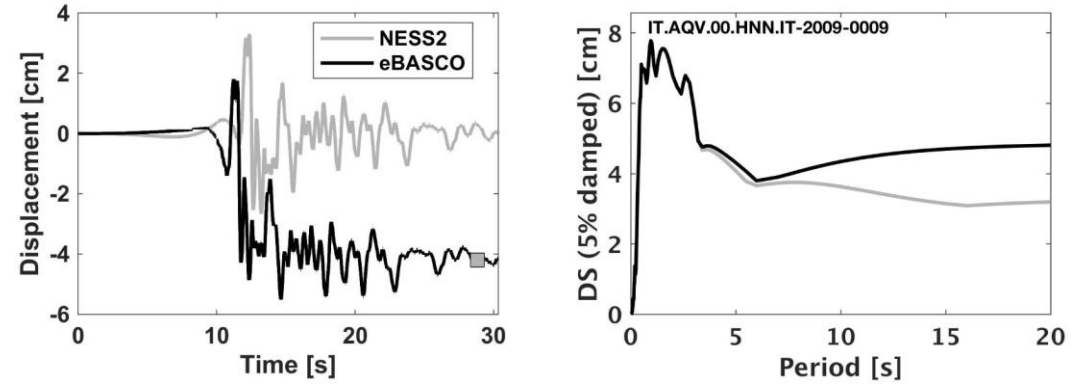


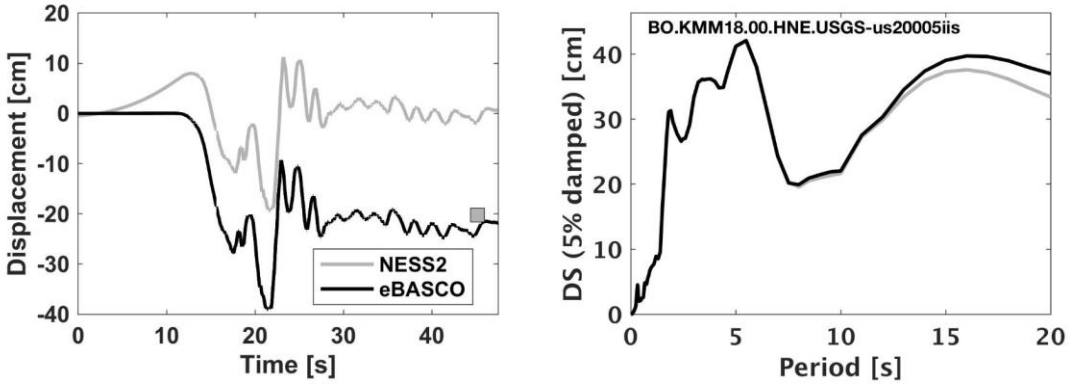
Figure 3. Schematic view of the NESS2 flat-files.



(a)



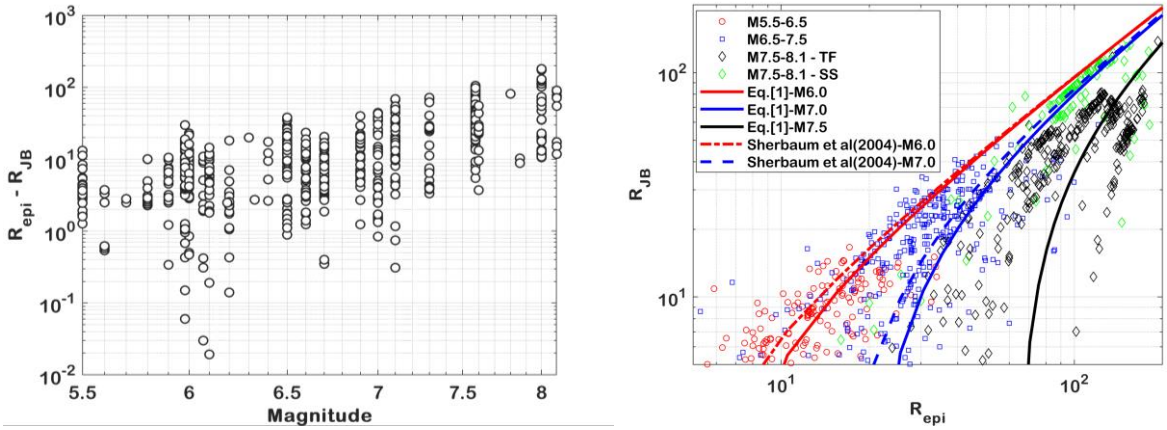
(b)



(c)

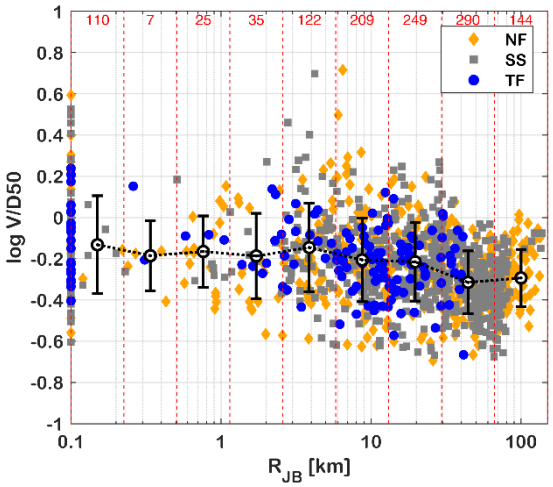
Figure 4. Displacement waveforms processed with eBASCO (black lines) and NESS2 standard algorithm (gray lines) plotted along with GPS (squares) coseismic displacement values (left panels). On the right panels the related displacement spectra are shown with their identification code: a) event EMSC-20161030_0000029 Norcia 2016/10/30 (Italy; CLO station); b) event IT-2009-0009 L'Aquila 2009/04/06 (Italy;

AQV station); c) event USGS-us20005iis Kumamoto 2016/04/15 (Japan; KMM18 station).

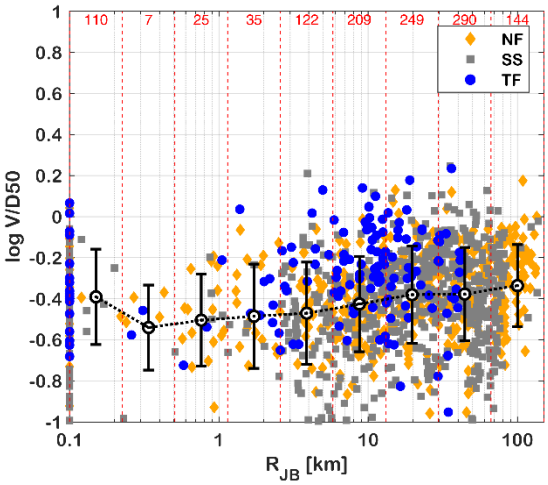


(a) **(b)**

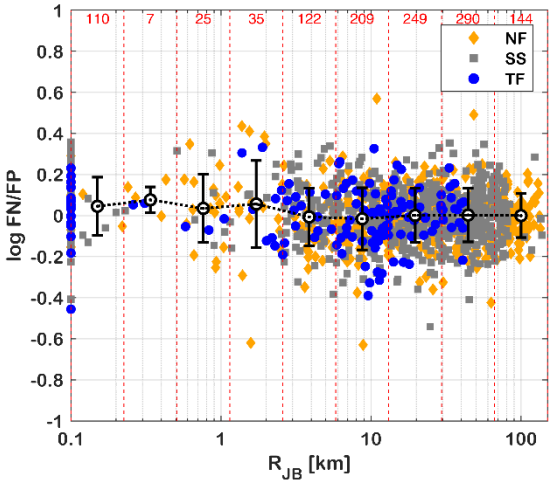
Figure 5. a) Scatter plot of the residuals (R_{epi} - R_{JB}) vs Magnitude in log-log scale; b) fitting curves for the relation R_{epi} - R_{JB} for magnitude ranges in NESS2 compared with Sherbaum et al. (2004) relations. The color version of this figure is available only in the electronic edition.


110 7 25 35 122 209 249 290 144
NF
SS
TF
log V/D50
R_JB [km]

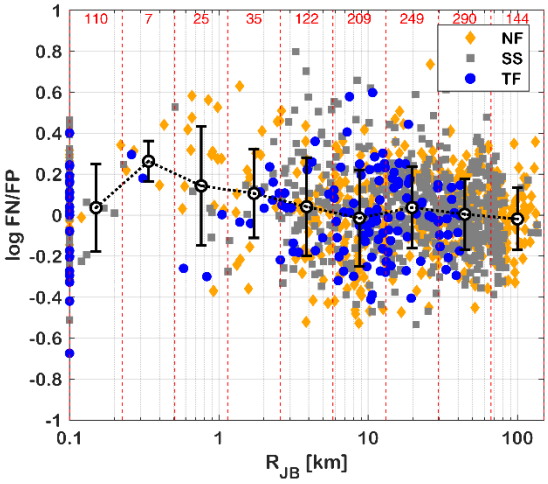
(a)


110 7 25 35 122 209 249 290 144
NF
SS
TF
log V/D50
R_JB [km]

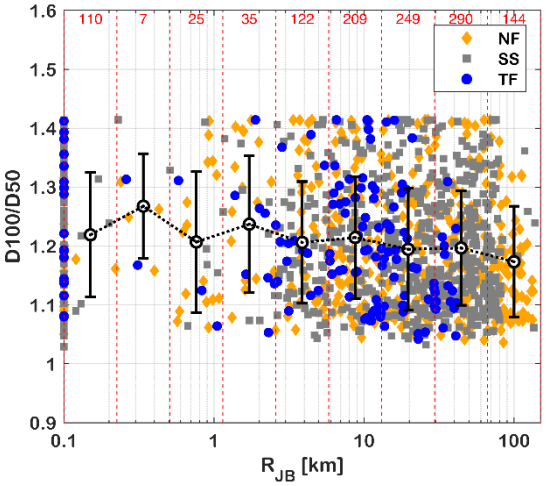
(b)

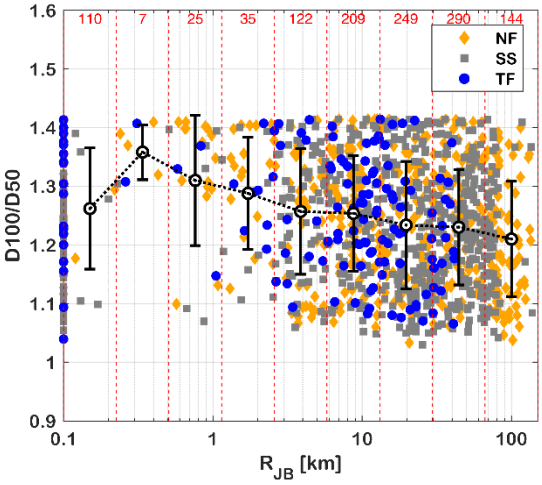

110 7 25 35 122 209 249 290 144
NF
SS
TF
log FN/FP
R_JB [km]

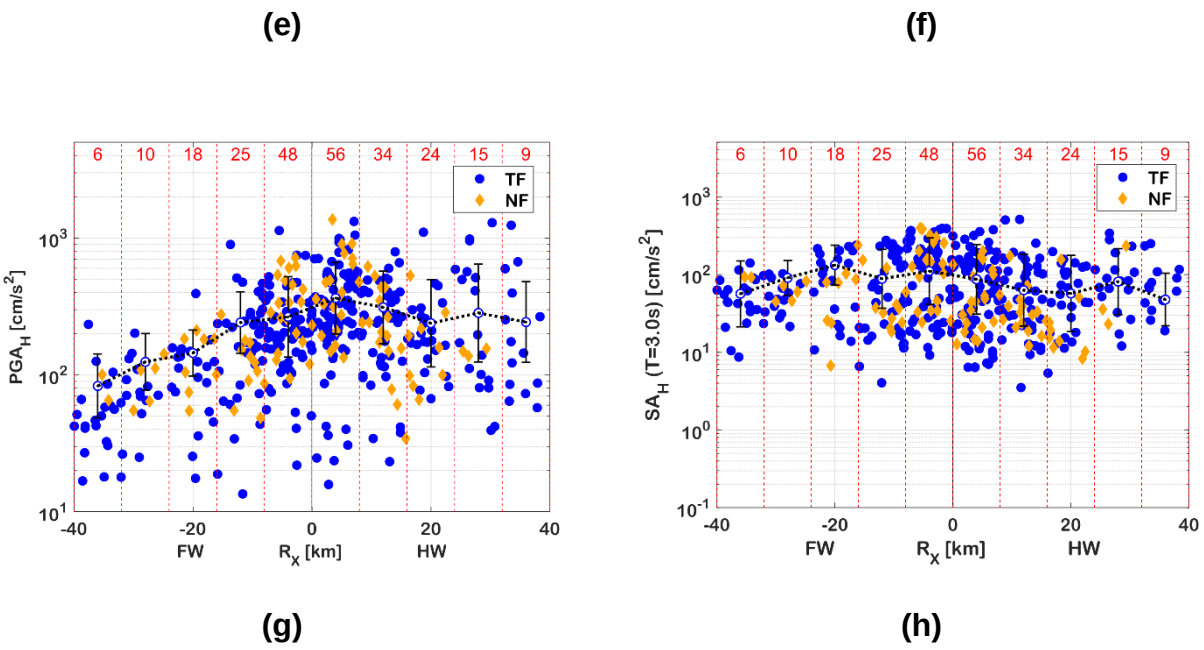
(c)


110 7 25 35 122 209 249 290 144
NF
SS
TF
log FN/FP
R_JB [km]

(d)


110 7 25 35 122 209 249 290 144
NF
SS
TF
D100/D50
R_JB [km]


110 7 25 35 122 209 249 290 144
NF
SS
TF
D100/D50
R_JB [km]



1059 **Figure 6. Distribution of data in NESS2 with distance at PGA (on the left panel) and**
1060 **SA at period T=3s (on the right panel) with different styles of faulting (SS – Strike-**
1061 **Slip; TF – Thrust-Fault; NF – Normal-Fault): (a-b) ratio V/D50 vs R_{JB}; (c-d) ratio**
1062 **log(FN/FP) vs R_{JB}; (e-f) ratio D100/D50 vs R_{JB}; (g-h) horizontal PGA and spectral**
1063 **amplitudes SA vs Rx on Hanging-Wall (HW) and Footwall (FW) side. The average**
1064 **dashed black curve and the error bar is overlapped to data. The number of records in**
1065 **each bin is reported on the top. The color version of this figure is available only in the**
1066 **electronic edition.**

1067

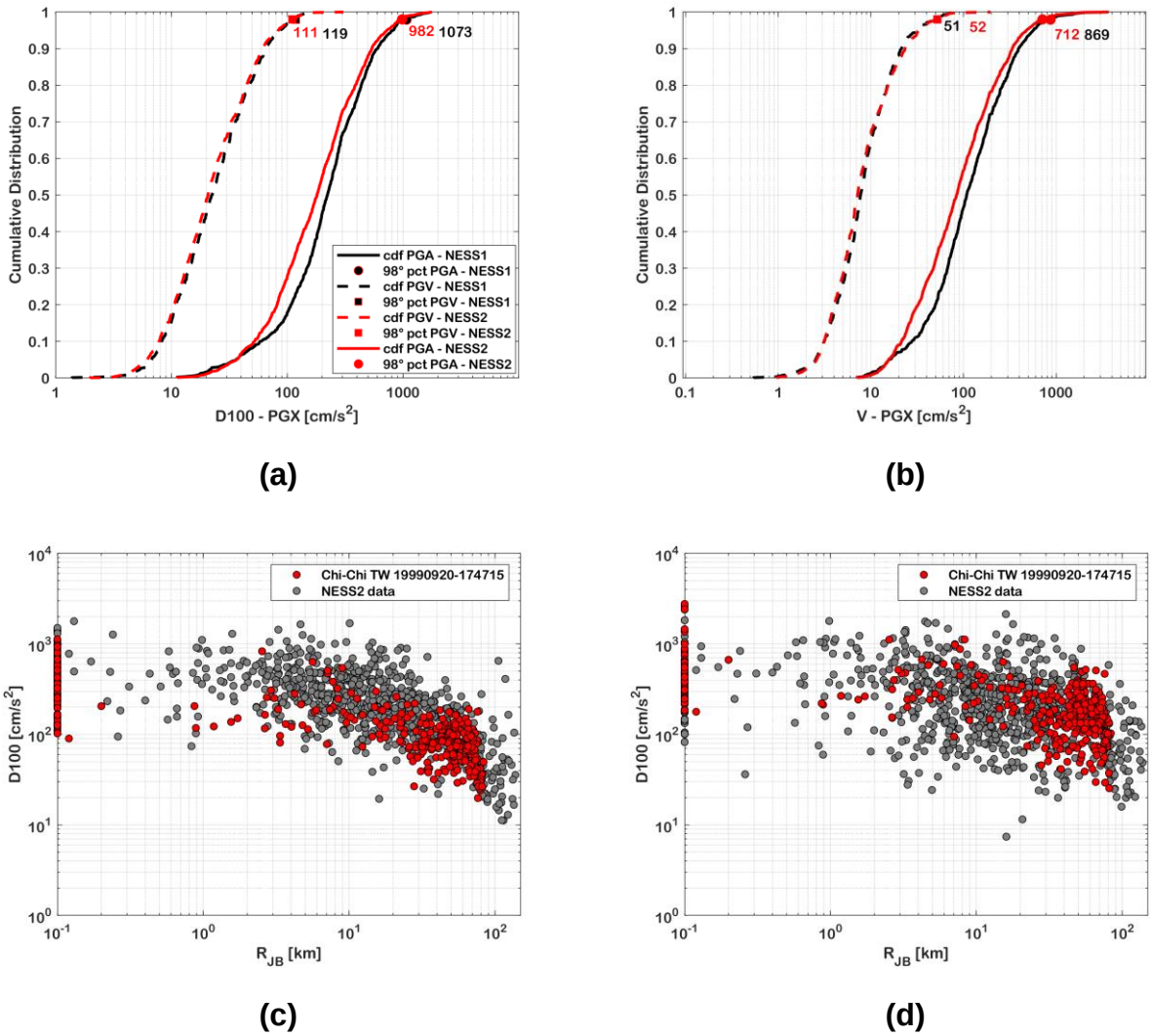
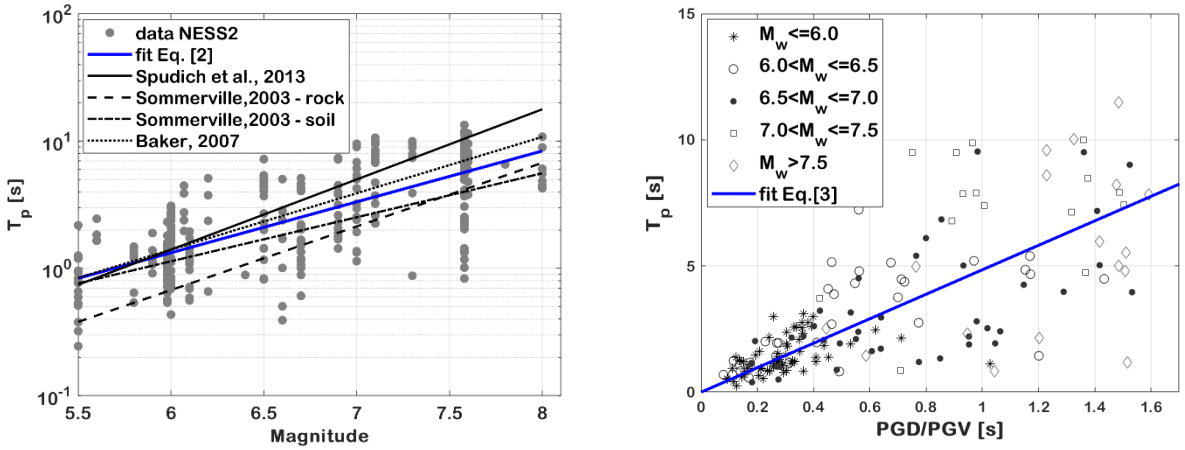
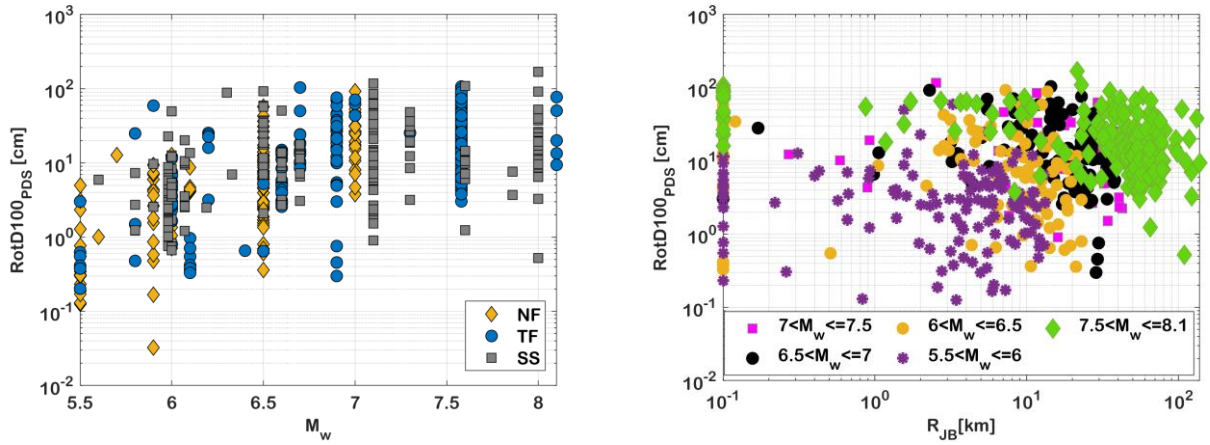


Figure 7. Comparison of empirical cumulative distribution functions of NESS2 and NESS1 records for the PGA (solid lines) and PGV (dashed lines) in terms of: (a) maximum of the rotated waveforms RotD100 (D100) and (b) vertical component (V). The reported numbers indicate the peak values at the 98° percentiles; D100 distribution versus Joyner-Boore distance of Chi-Chi Taiwan 1999-09-20 records compared to NESS2 records: (c) PGA and (d) SA(T=1s). The color version of this figure is available only in the electronic edition.



(a) (b)

Figure 8. a) Pulse period versus magnitude Eq. [2] (RMSE = 0.62; R-squared = 0.5181) compared with previous studies; b) pulse-period versus PGD/PGV ratio for different magnitude ranges, Eq. [3] (RMSE = 1.81; R-squared = 0.5285), for NESS2 pulse-like records (eBASCO processed, FN components). The color version of this figure is available only in the electronic edition.



(a) **(b)**

Figure 9. Horizontal permanent displacement PDS (RotD100 components) versus a) magnitude (dots are marked based on the styles of faulting: SS – Strike-Slip; TF – Thrust-Fault; NF – Normal-Fault), and b) R_{JB} distance for the magnitude range 5.5–8.1 (dots are marked based on the magnitude bins). The color version of this figure is available only in the electronic edition.

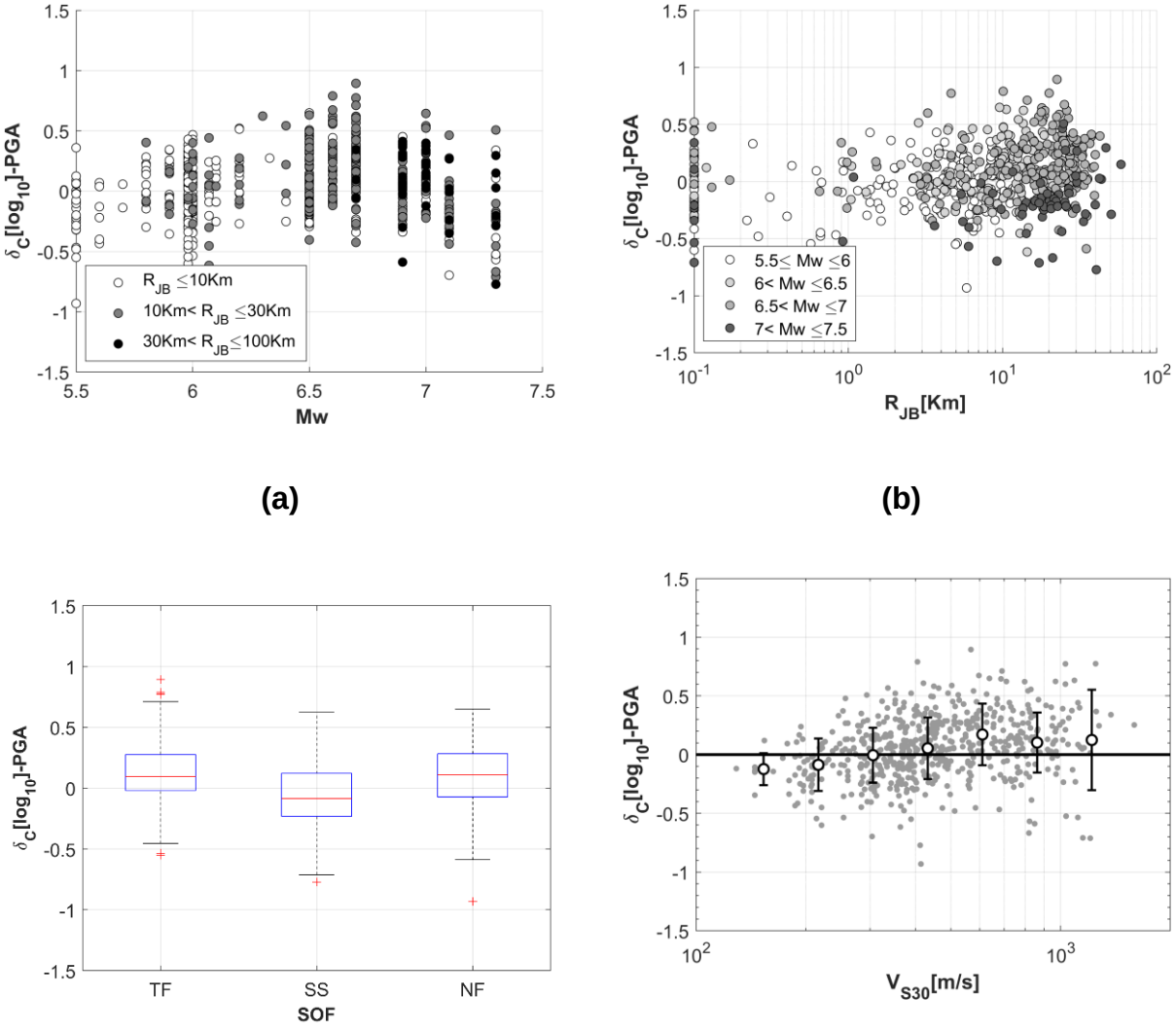


Figure 10. Total residuals δc, computed as logarithm difference between NESS2 observations and ITA18 predictions for PGA, as a function a) moment magnitude (M_W); b) Joyner-Boore distance (R_{JB}); c) style of faulting (SOF); d) average shear-wave velocity in the uppermost 30m ($V_{s,30}$).

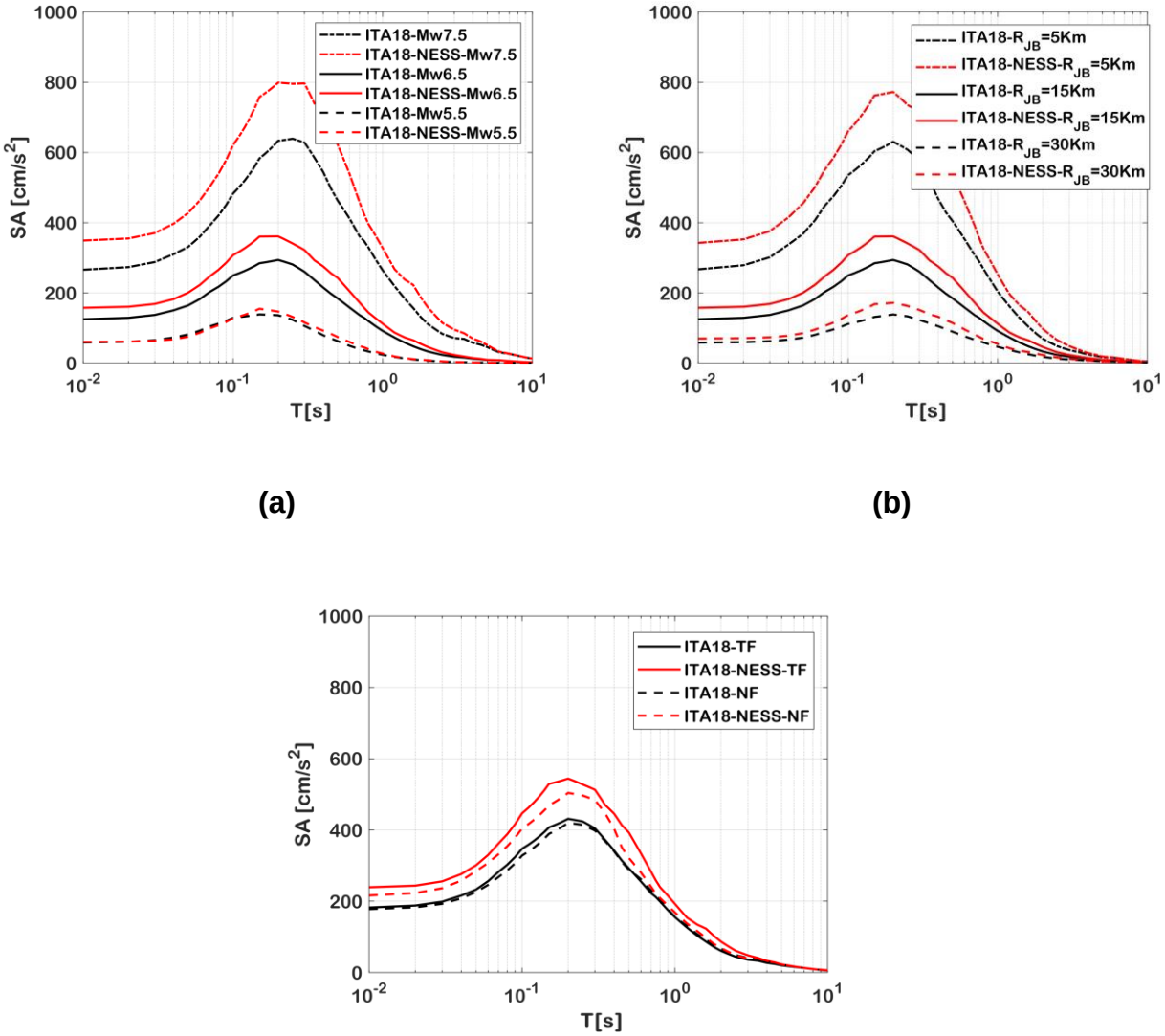
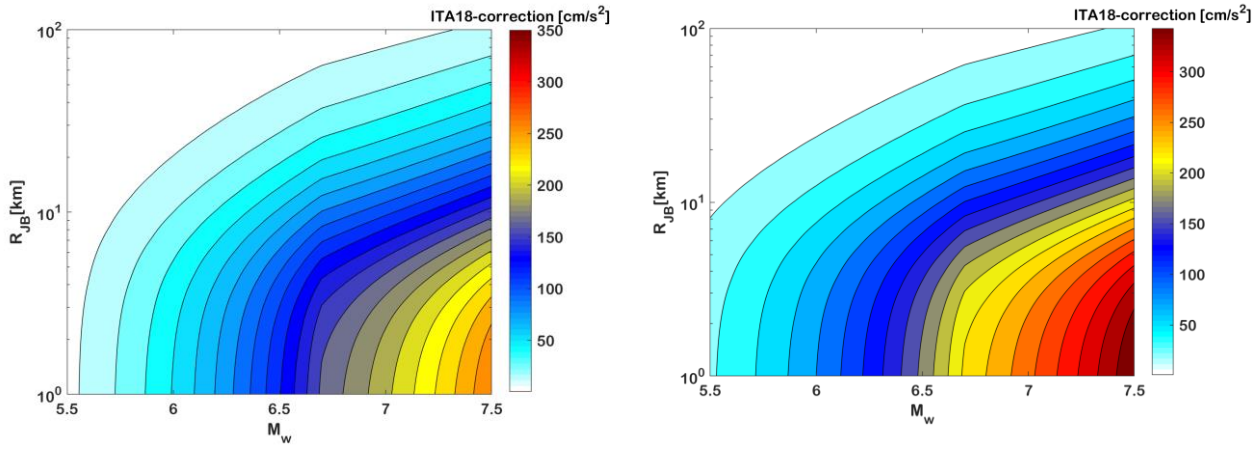


Figure 11. Horizontal acceleration spectral predictions of ITA18 corrected with NESS2 data ($SA_{ITA18\text{-}NESS}$) compared with uncorrected ITA18 (SA_{ITA18}): sensitivity analysis with: a) M_w (fixed style of faulting TF; $V_{s,30}$=650 m/s and R_{JB} = 15km); b) R_{JB} distance (fixed style of faulting TF; Mw=6.5 and $V_{s,30}$=650 m/s) and c) styles of faulting (fixed Mw 7.0; $V_{s,30}$=650 m/s and R_{JB} = 15km). The color version of this figure is available only in the electronic edition.



(a) **(b)**

Figure 12. Scenario-dependence of ITA18 correction for SA(T=0.2s) with reference to $V_{s,30}$ at 700 m/s and focal mechanisms of a) Normal-Fault type and b) Thrust-Fault type. The color version of this figure is available only in the electronic edition.

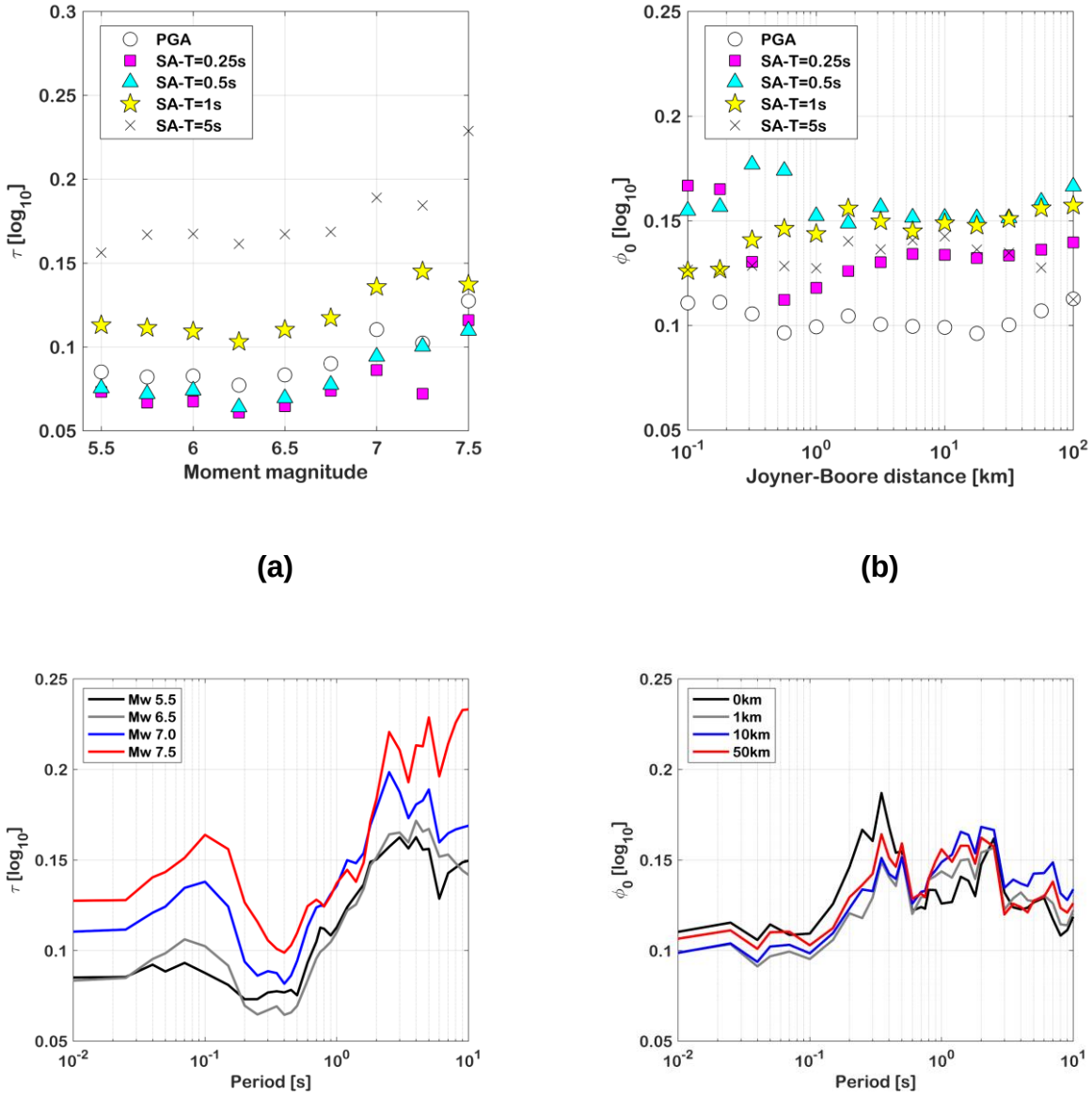


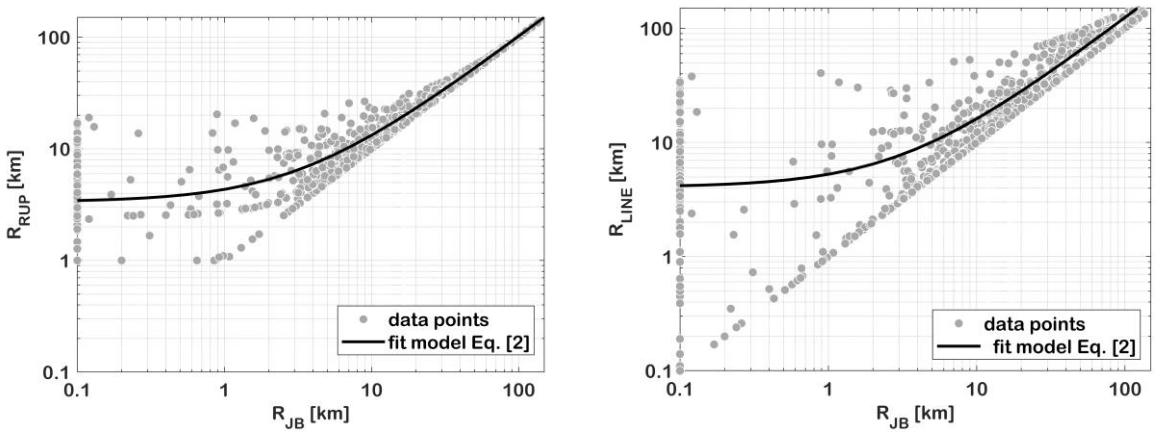
Figure 13. Scenario-dependent standard deviations: a) between-event standard deviation versus magnitude; b) event- and site- corrected residuals standard deviation versus Joyner-Boore distance; c) between-event standard deviation versus periods; d) event- and site- corrected residuals standard deviation versus periods. The color version of this figure is available only in the electronic edition.

APPENDIX

Here are reported conversion relationships among different finite-fault metrics (Y and X) as in equation [1A] and detailed in Table 1A:

$$R_{\text{finite-fault-Y}} = \alpha * R_{\text{finite-fault-X}} + \beta \qquad [1A]$$

Figure 1A a) and b) show the distribution of data along with the fitting relations for the conversion of R_{JB} vs R_{RUP} and R_{JB} vs R_{LINE}, respectively. In these relationships (equation [1A]), the dependence on magnitude is less clear with respect to the relations between point-source and finite-fault metrics of Table 1A (equation [1] in the paper) and thus we derive simple linear models.



(a) **(b)**

Figure 1A. Linear fitting model [1A] with data points for a) R_{JB} vs R_{RUP} and b) R_{JB} vs R_{LINE} (log-log scale).

Table 1A. Main parameters of the fitting models of equation [1] and [1A]

Conversion type	Model	Index of the goodness of fit

R_{epi} vs R_{JB}	Equation [1] a=0.001399; b=1.38	R-square: 0.4116 RMSE: 32.71
R_{epi} vs R_{LINE}	Equation [1] a=0.0002009; b=1.597	R-square: 0.3656 RMSE: 28.81
R_{hypo} vs R_{RUP}	Equation [1] a=0.001403; b=1.374	R-square: 0.4028 RMSE: 31.46
R_{JB} vs R_{RUP}	Equation [1A] α=0.9896; β=3.344	R-square: 0.9858 RMSE: 116.3
R_{JB} vs R_{LINE}	Equation [1A] α=1.207; β=4.062	R-square: 0.8673 RMSE: 462.8

1128

Figure 1

Click here to access/download;Figure;Figure1.pdf

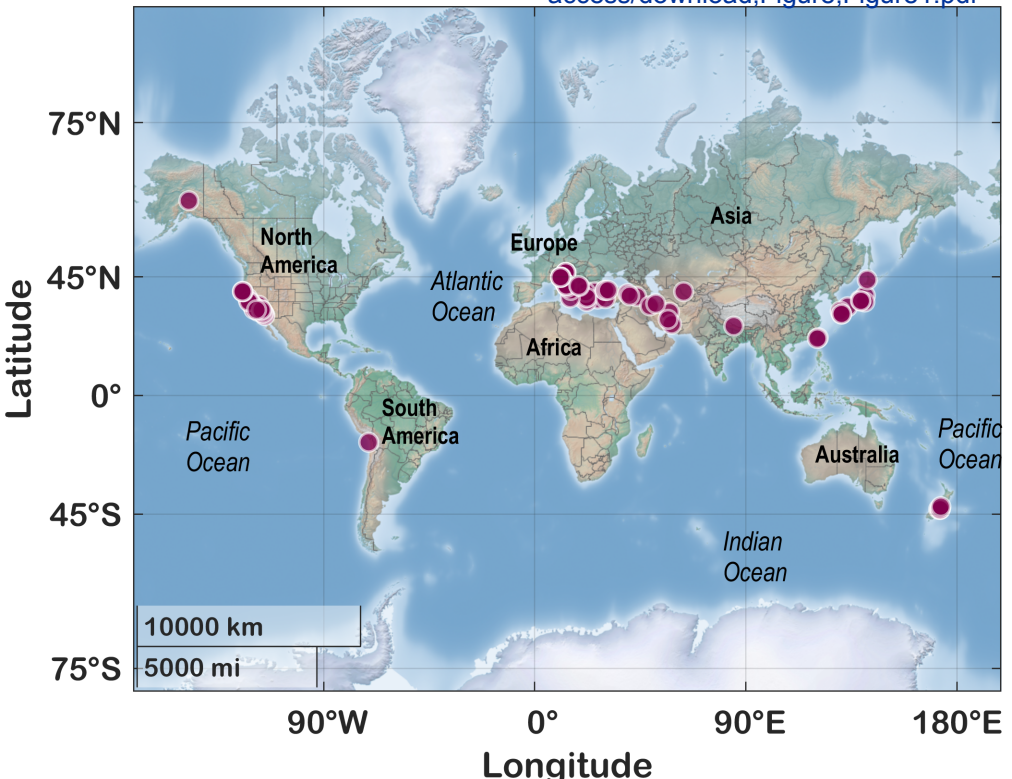


Figure 2a

Click here to access/download;Figure;Figure2a.pdf

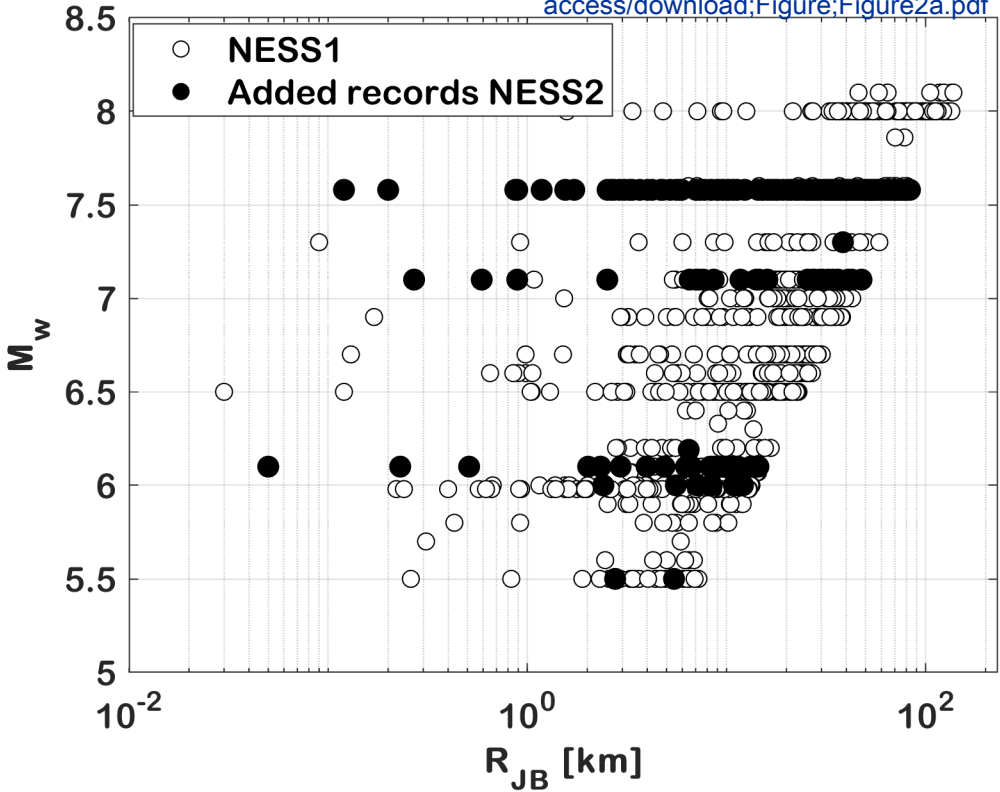


Figure 2b

Click here to access/download;Figure;Figure2b.pdf

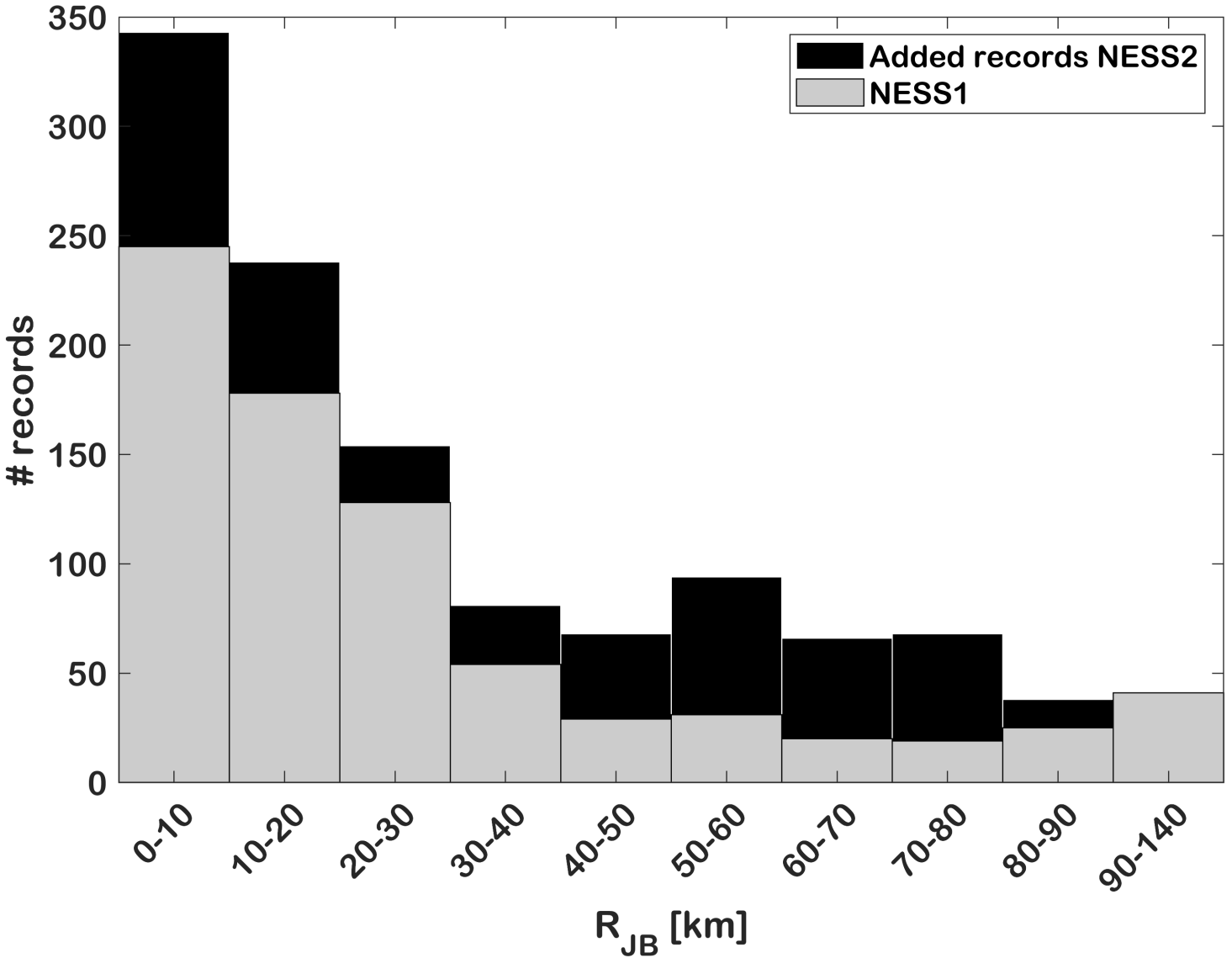


Figure 3

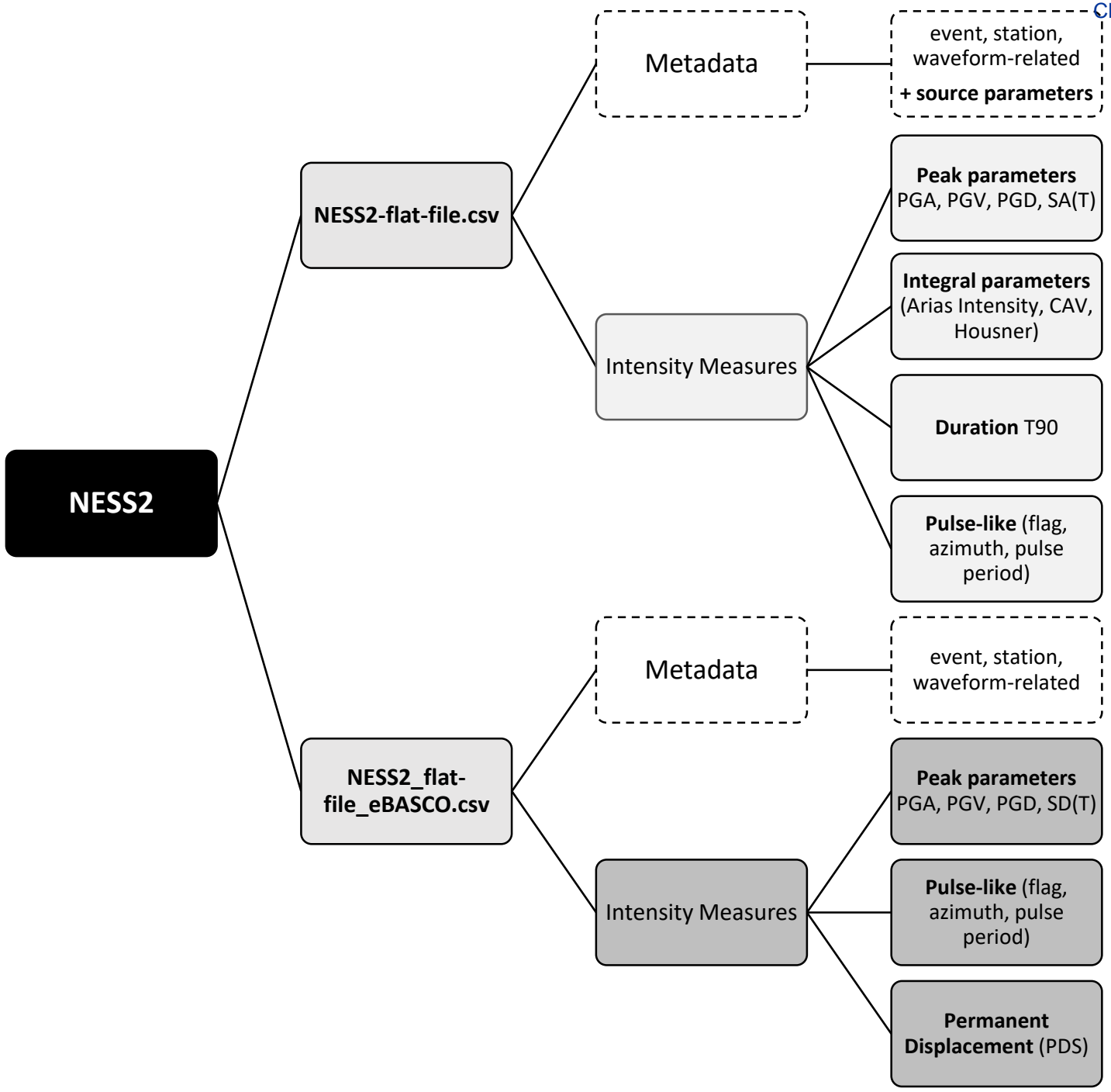


Figure 4a

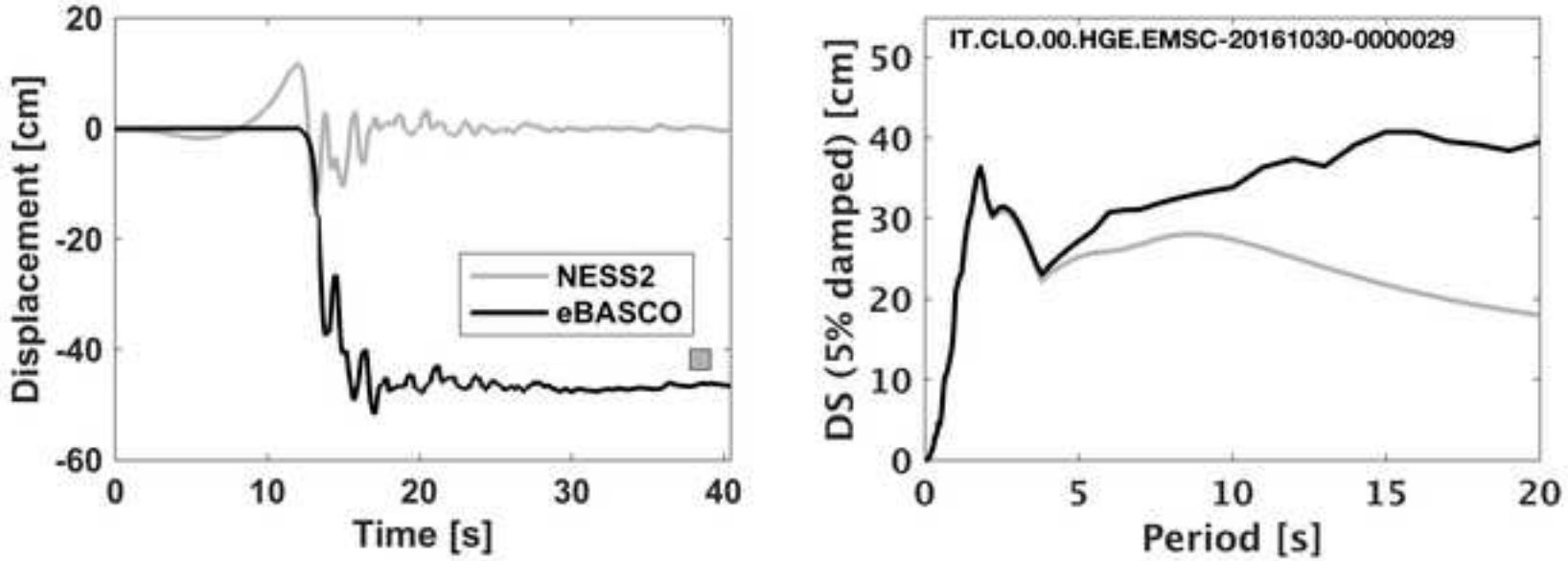


Figure 4b

Click here to access/download;Figure;Figure4b.tiff

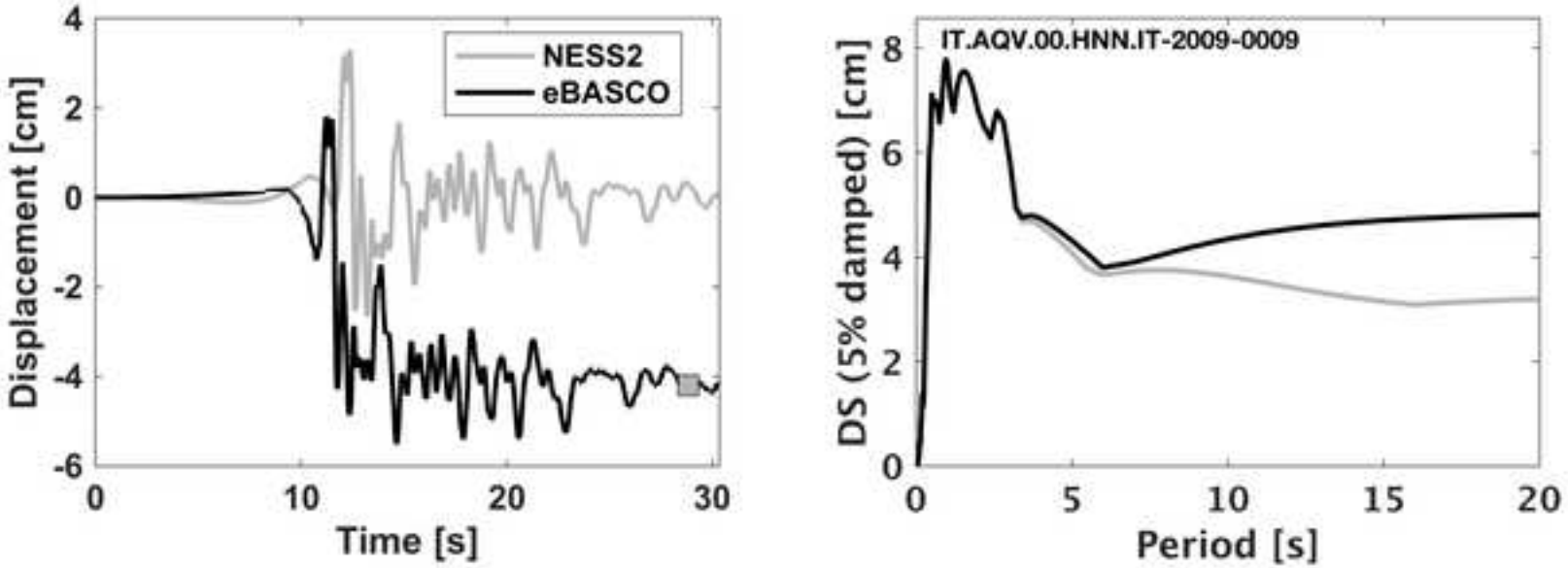


Figure 4c

Click here to access/download;Figure;Figure4c.tiff

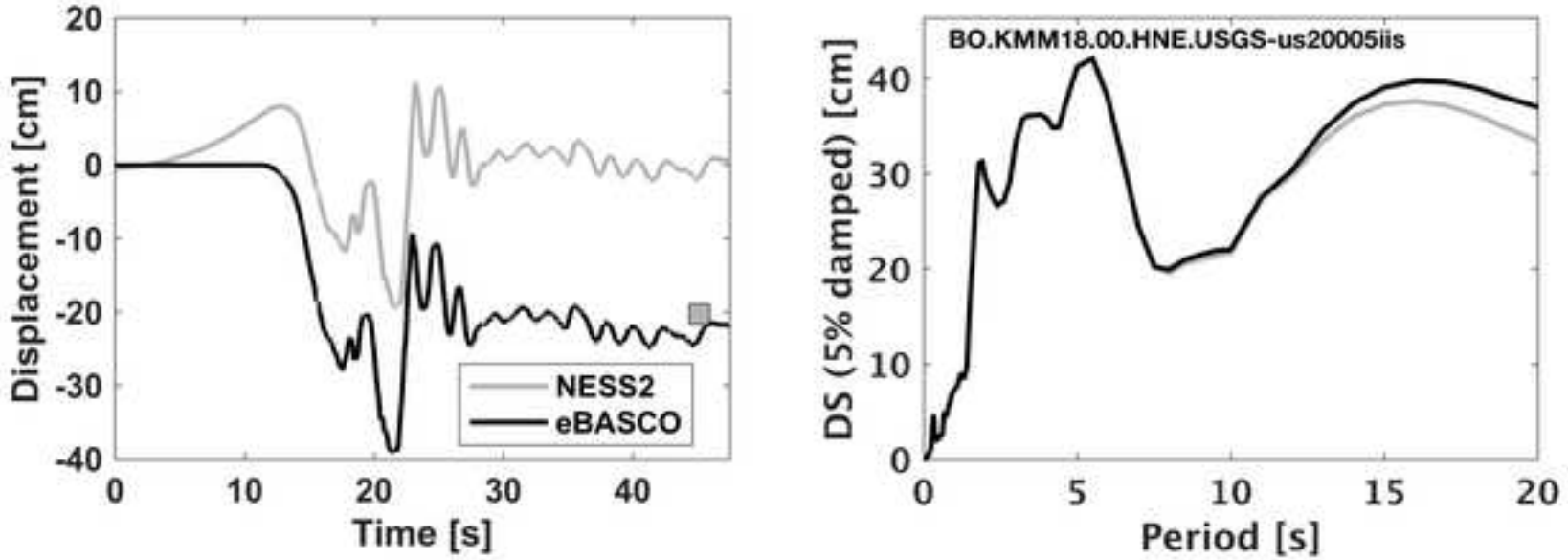


Figure 5a

Click here to access/download;Figure;Figure5a.pdf

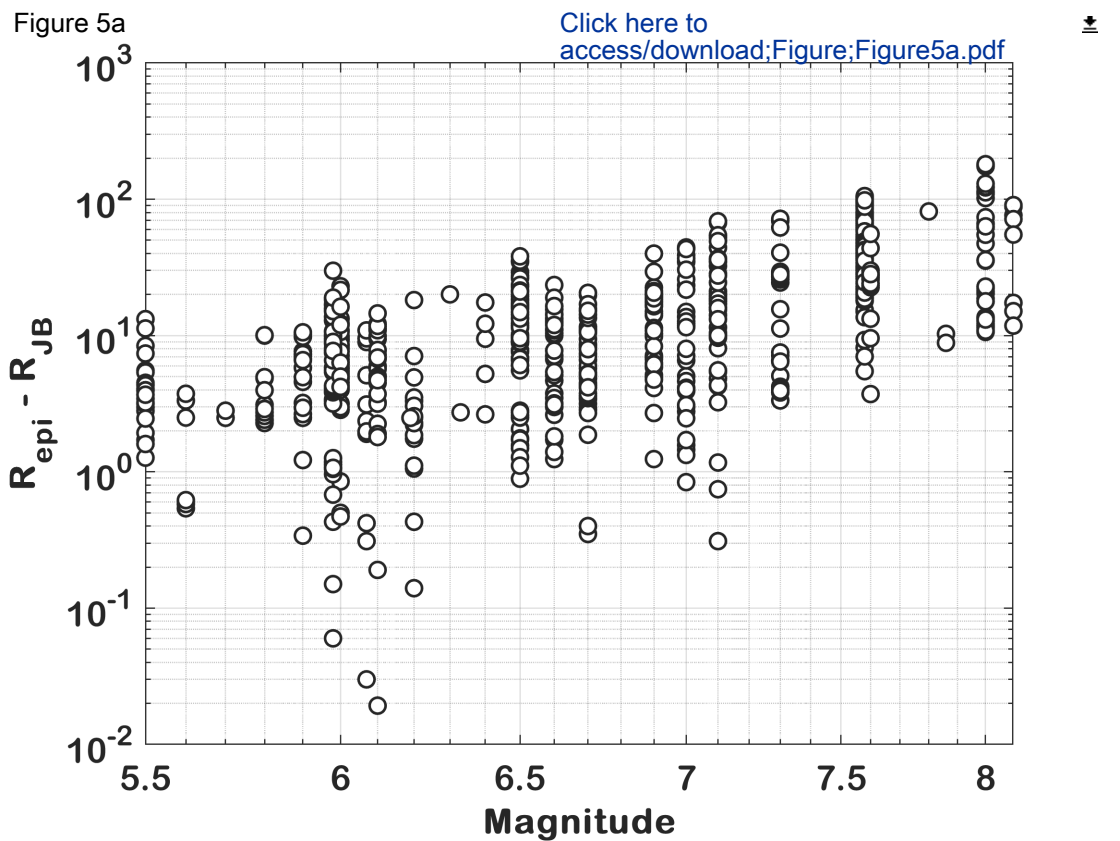


Figure 5b

Click here to access/download;Figure;Figure5b.pdf

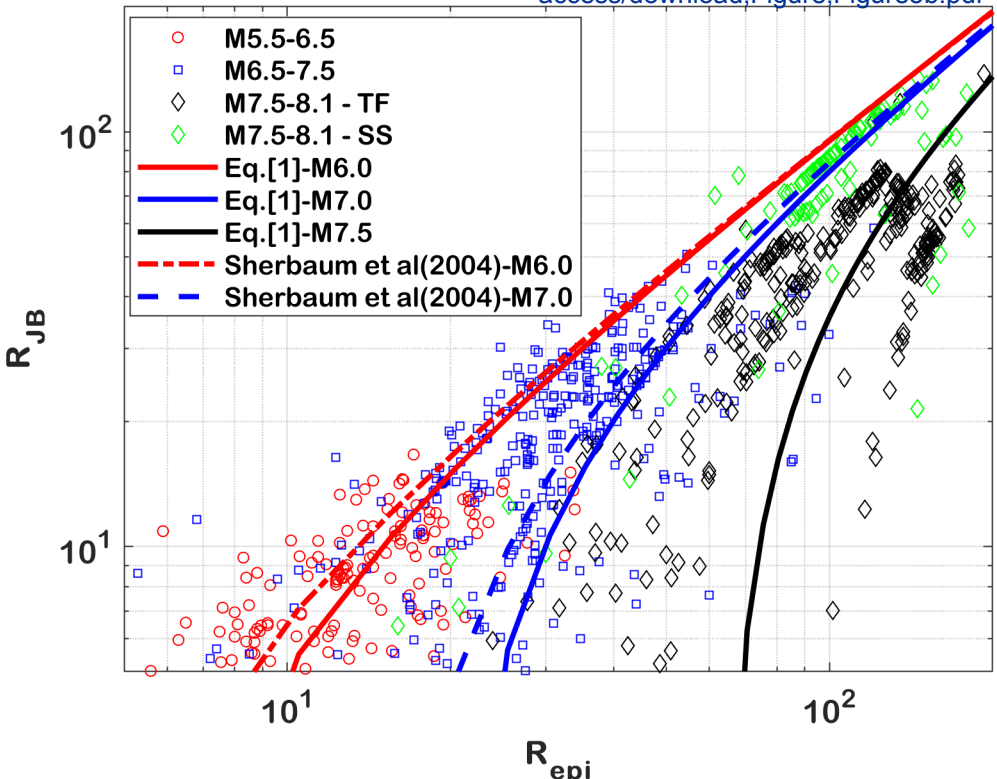


Figure 6a

Click here to access/download;Figure;Figure6a

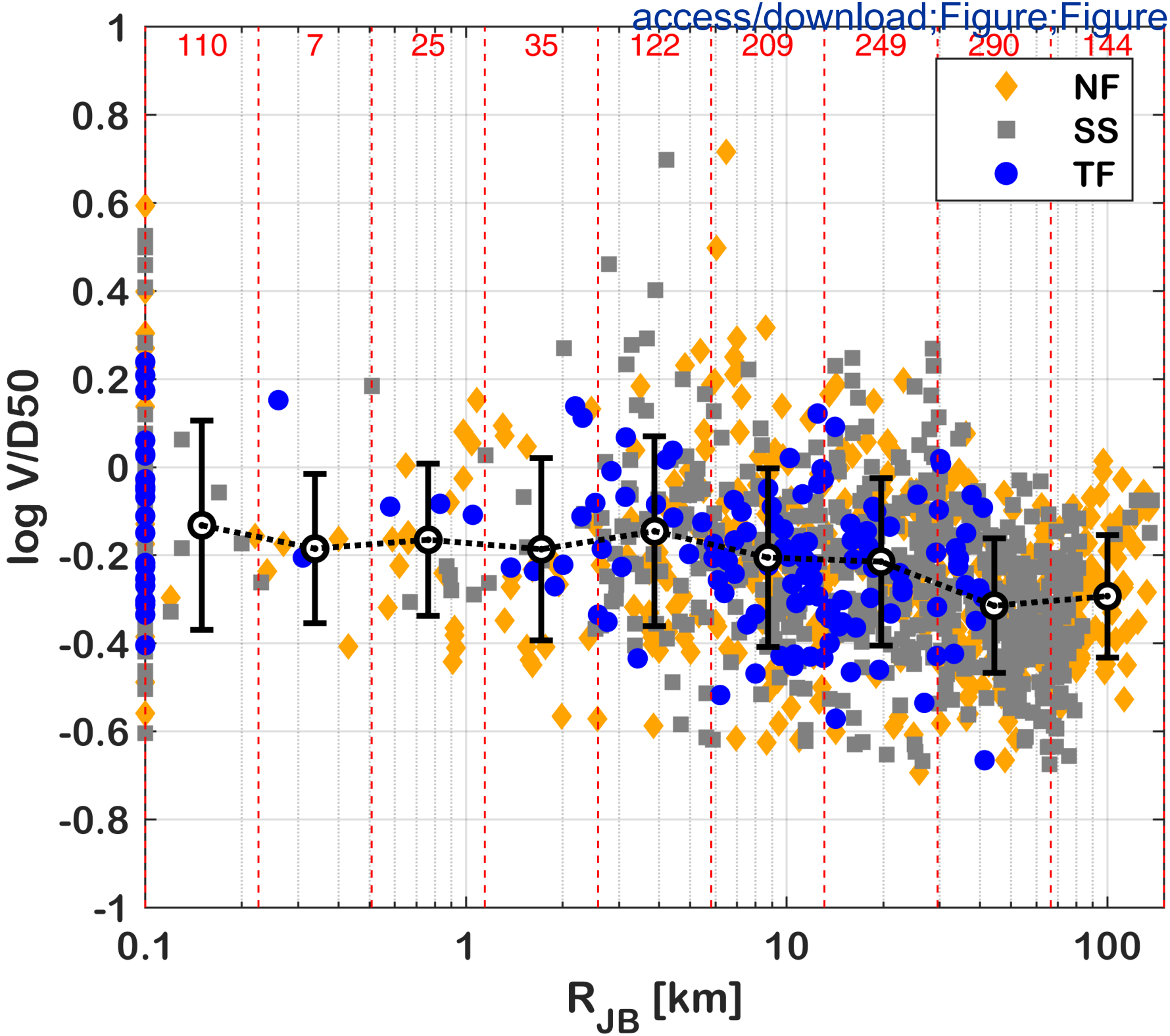


Figure 6b

Click here to access/download;Figure;Figure6b

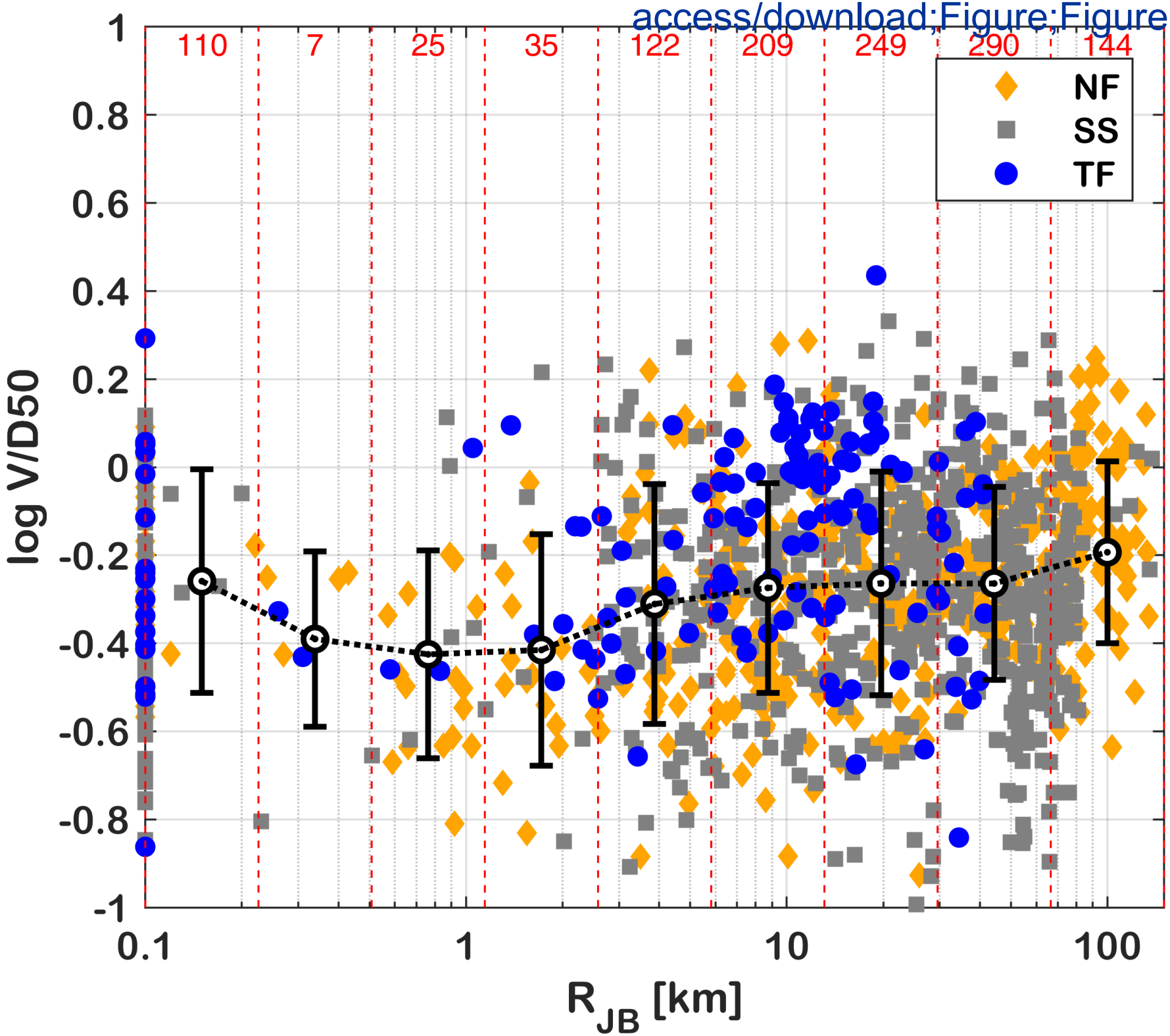


Figure 6c

Click here to access/download;Figure;Figure6c

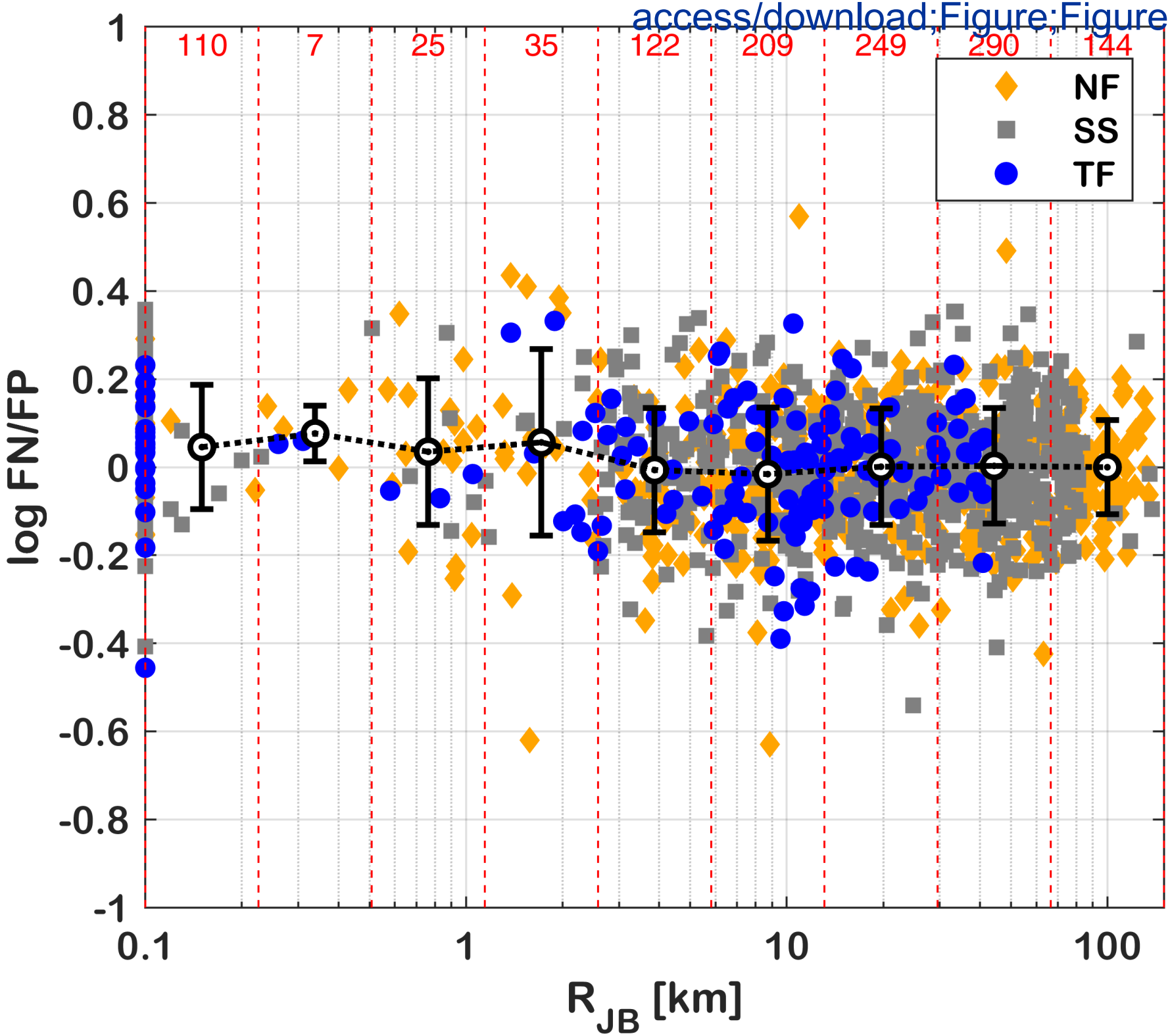


Figure 6d

Click here to access/download;Figure;Figure6d

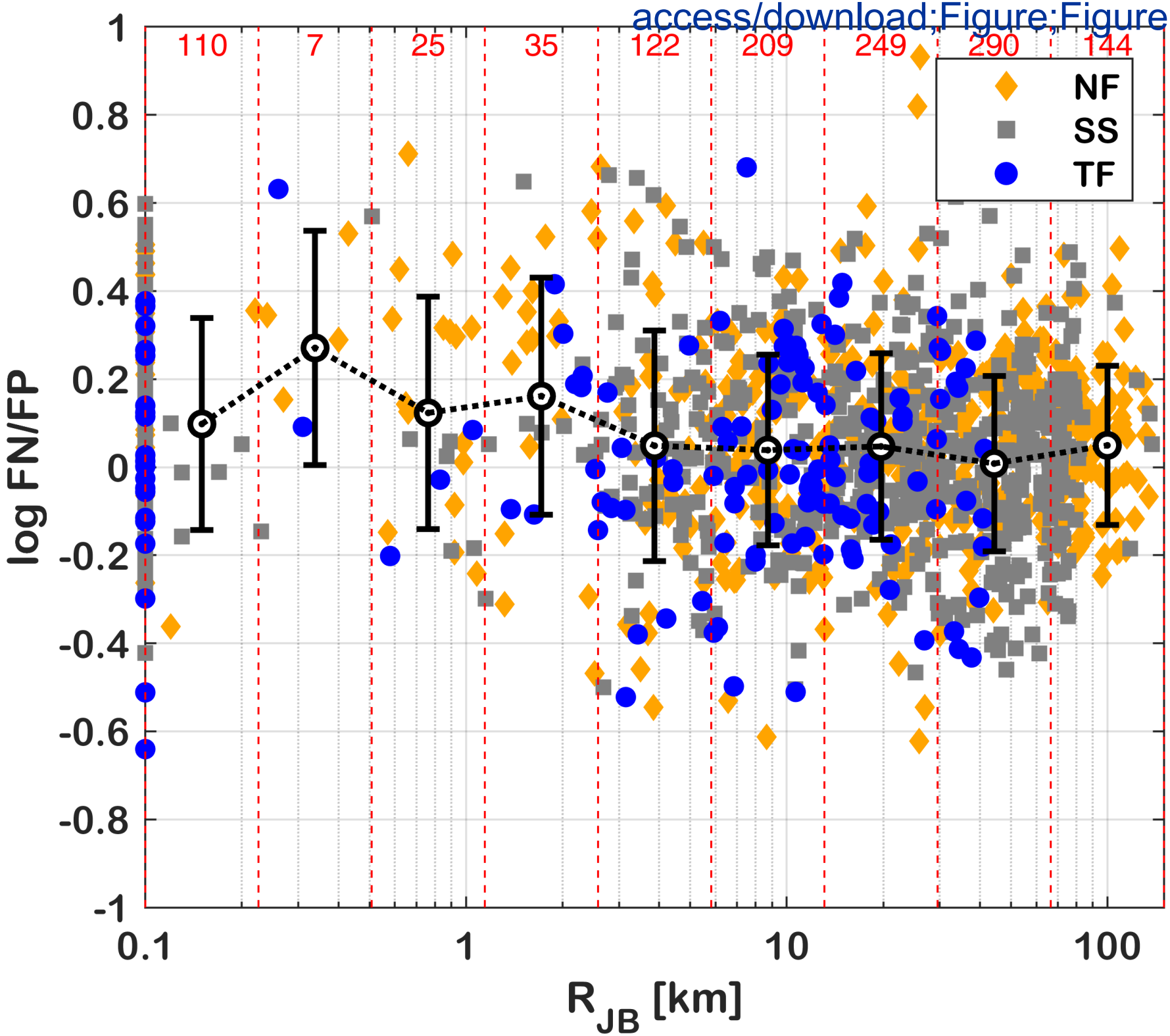


Figure 6e

Click here to access/download;Figure;Figure6e

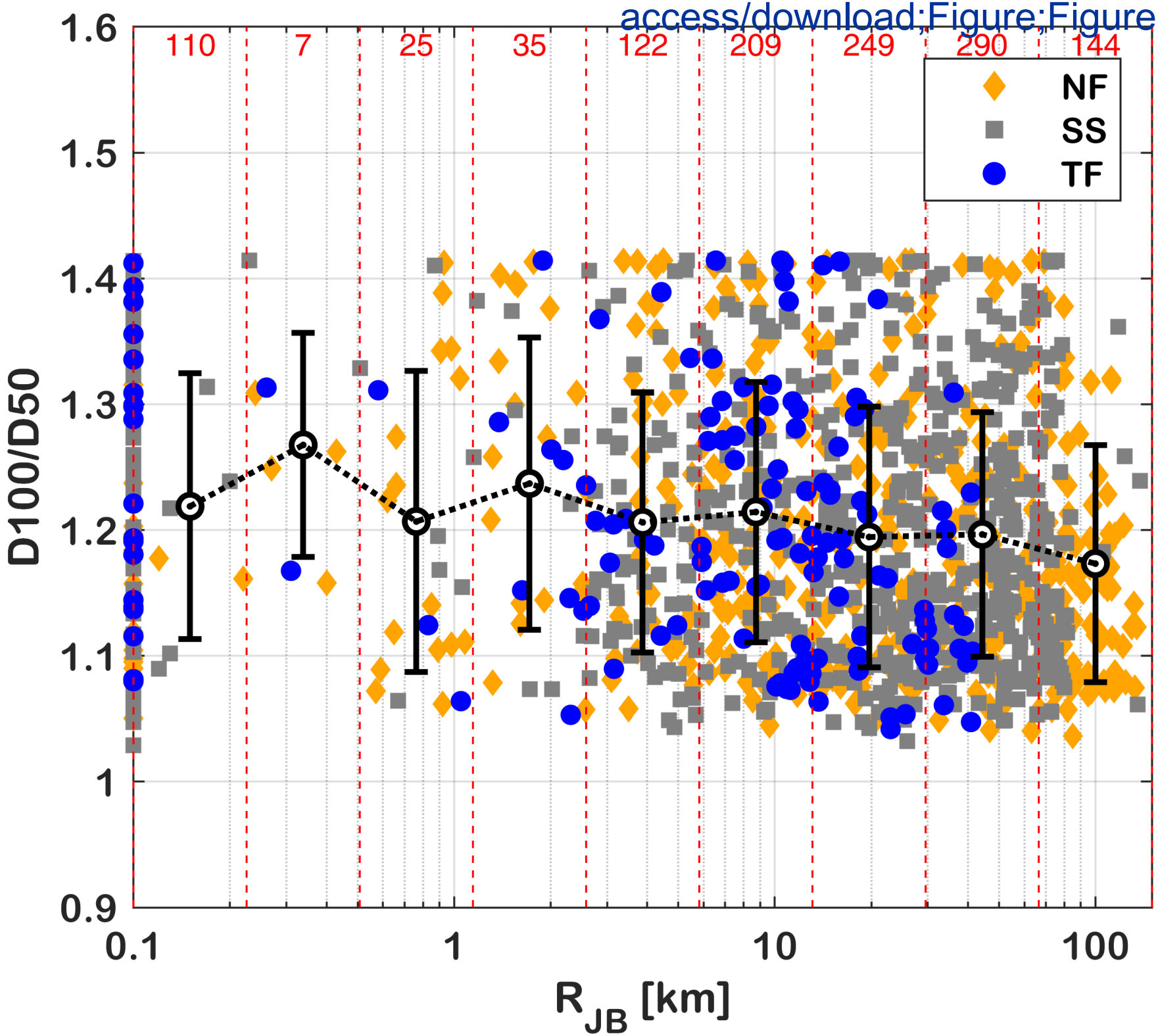


Figure 6f

Click here to access/download;Figure;Figure6f.

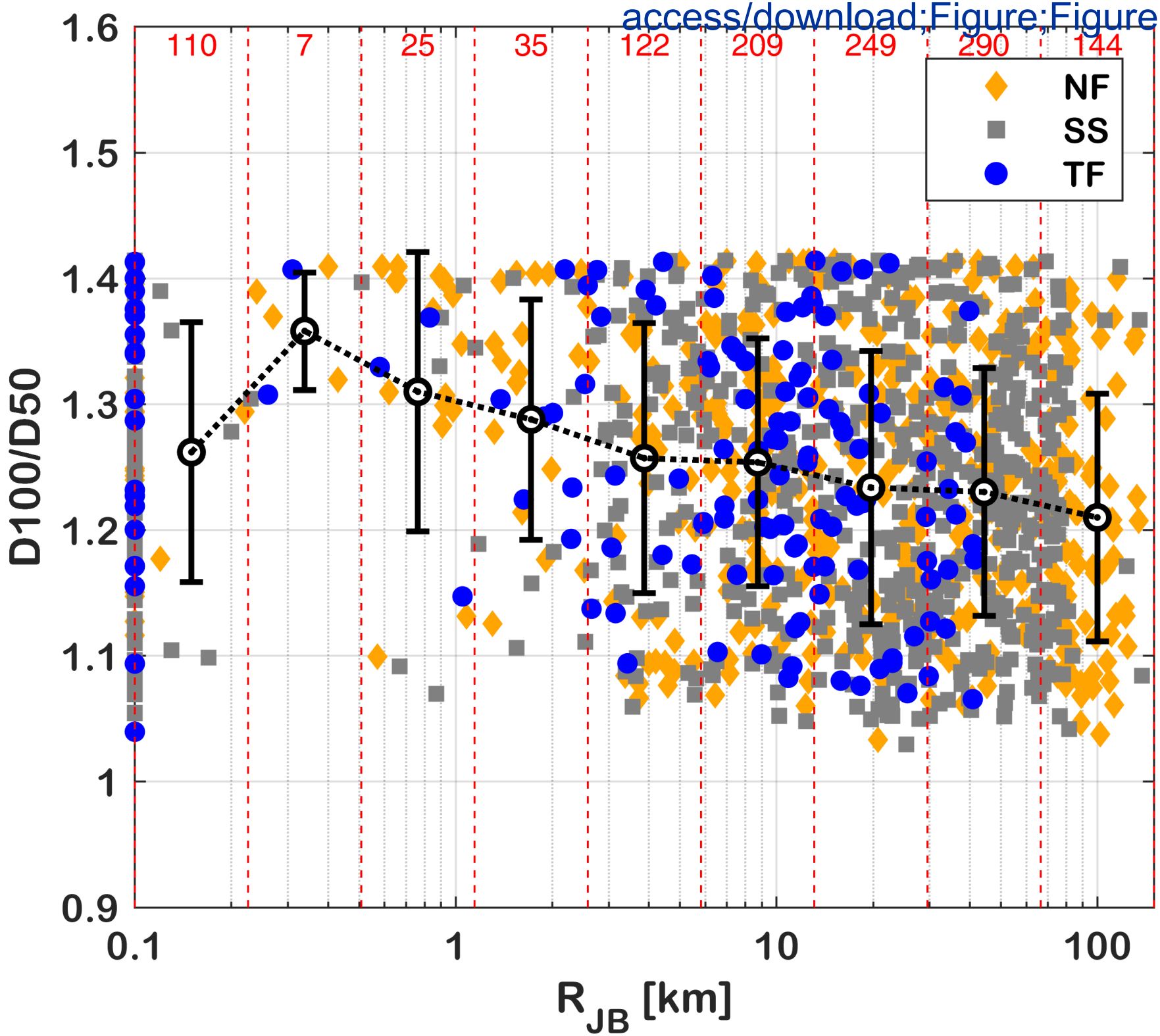


Figure 6g

Click here to access/download;Figure;Figure6g.pdf

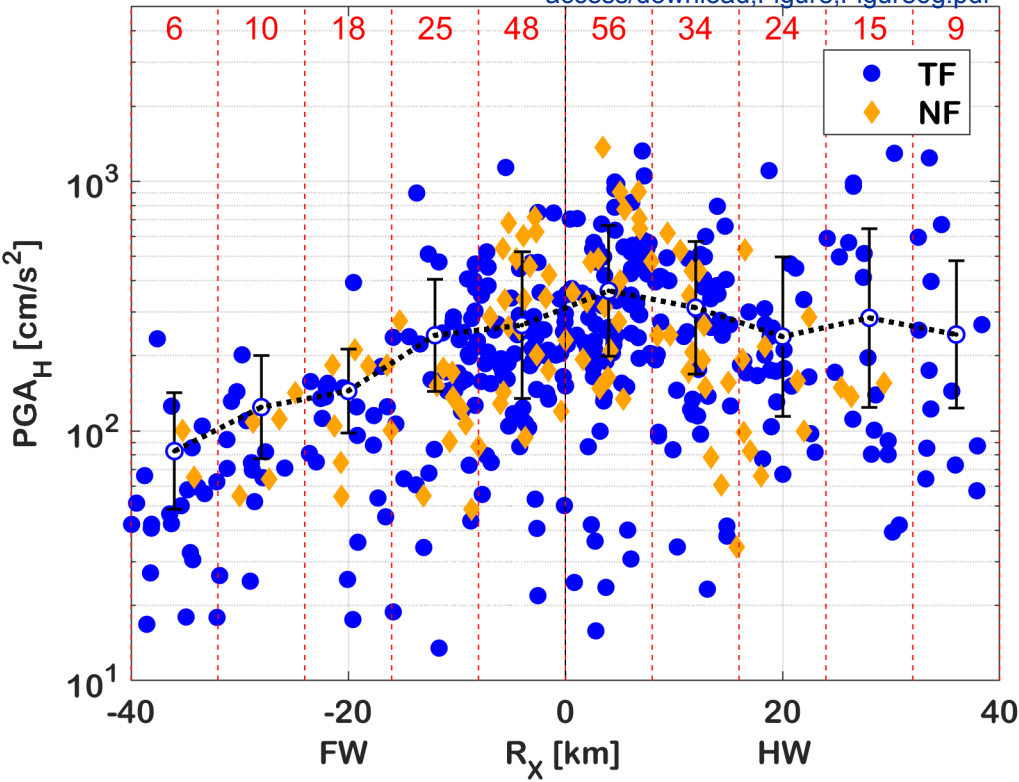


Figure 6h

Click here to access/download;Figure;Figure6h.pdf

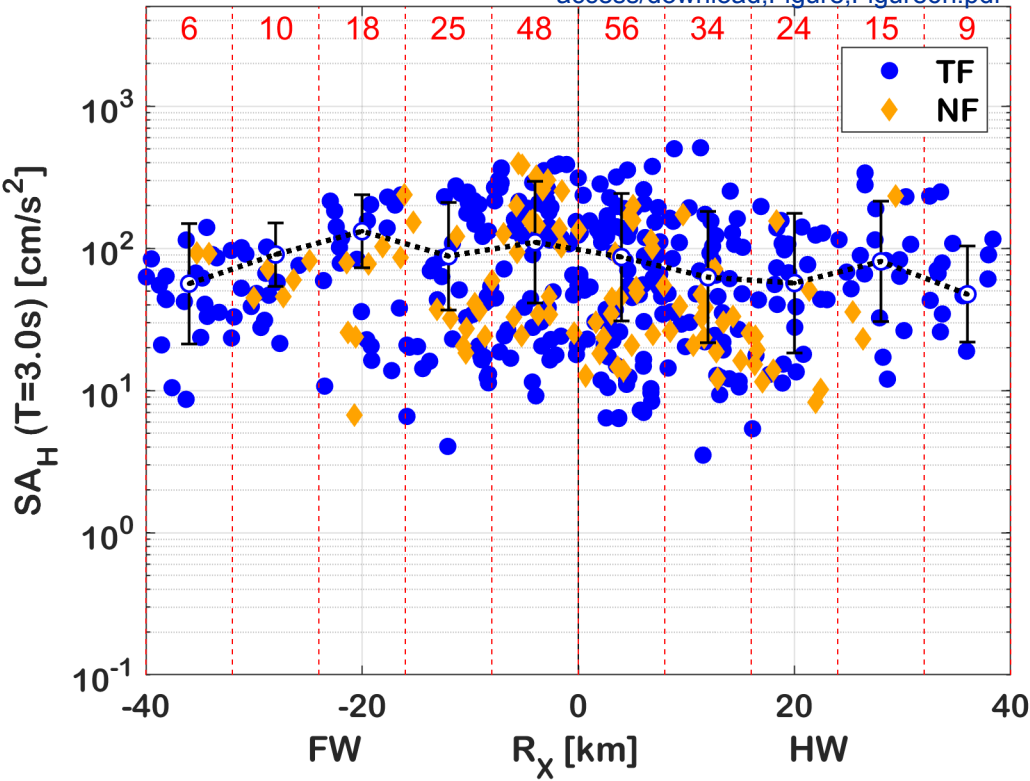


Figure 7a

Click here to access/download;Figure;Figure7a.pdf

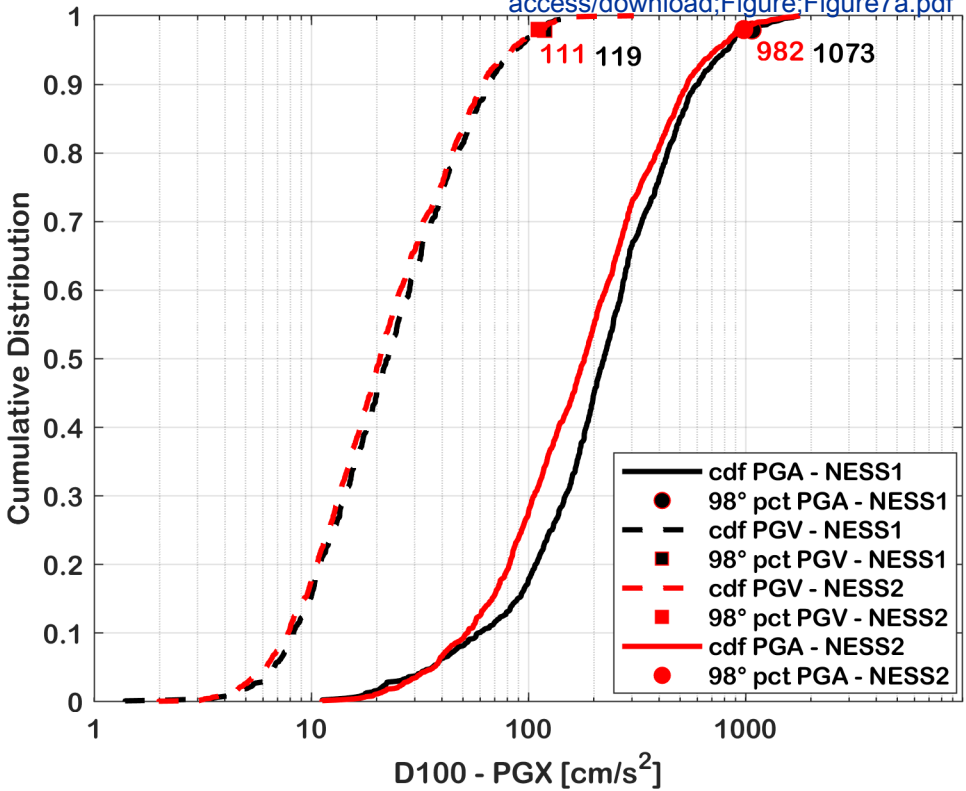


Figure 7b

Click here to access/download;Figure;Figure7b.pdf

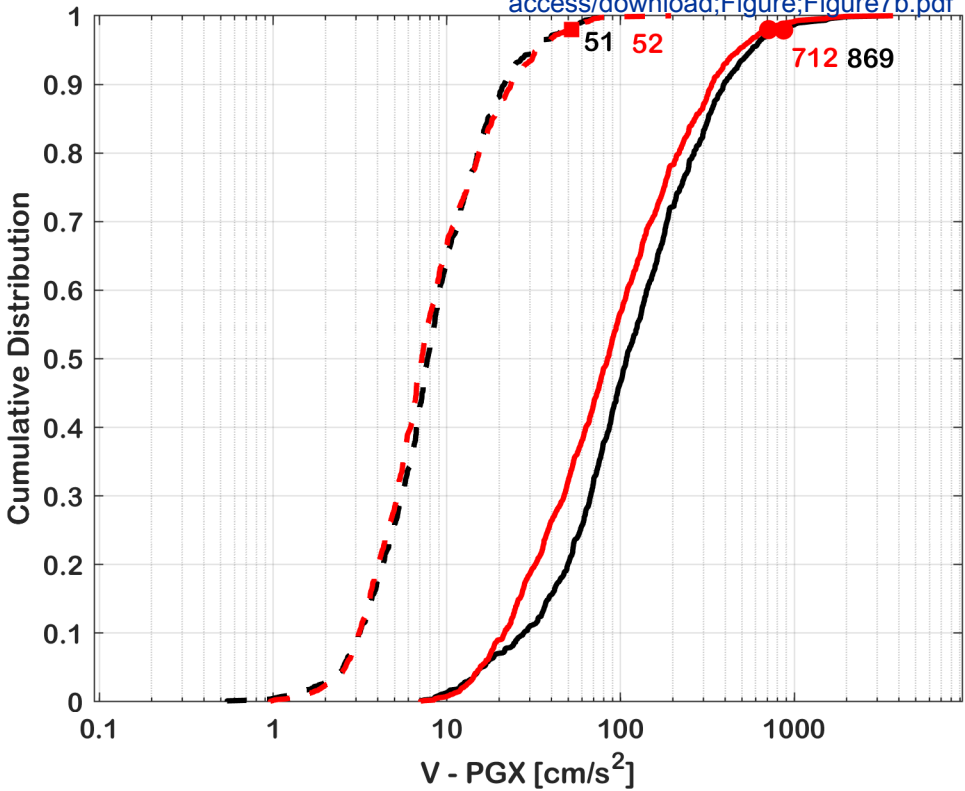


Figure 7c

Click here to access/download;Figure;Figure7c.pdf

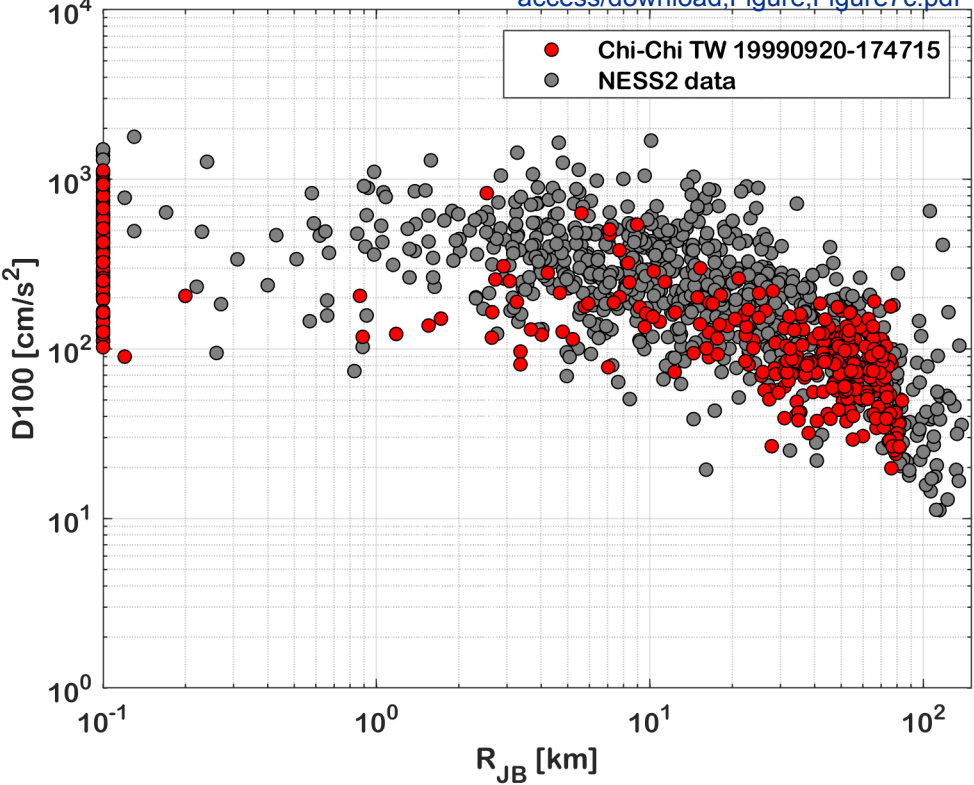


Figure 7d

Click here to access/download;Figure;Figure7d.pdf

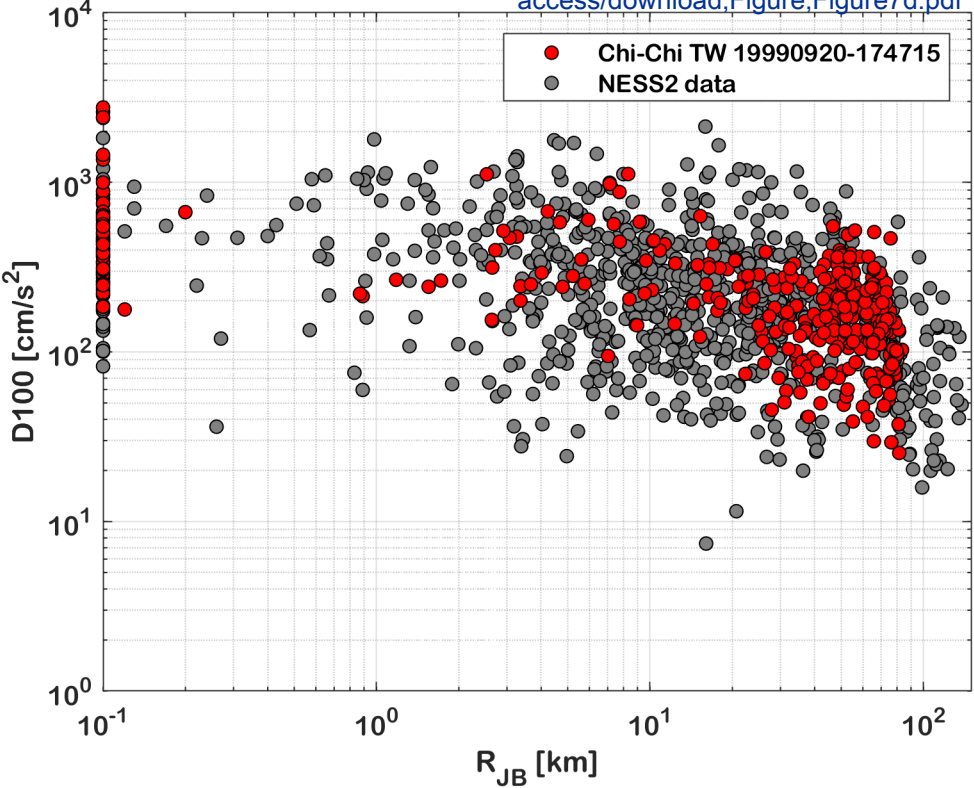


Figure 8a

Click here to access/download;Figure;Figure8a.pdf

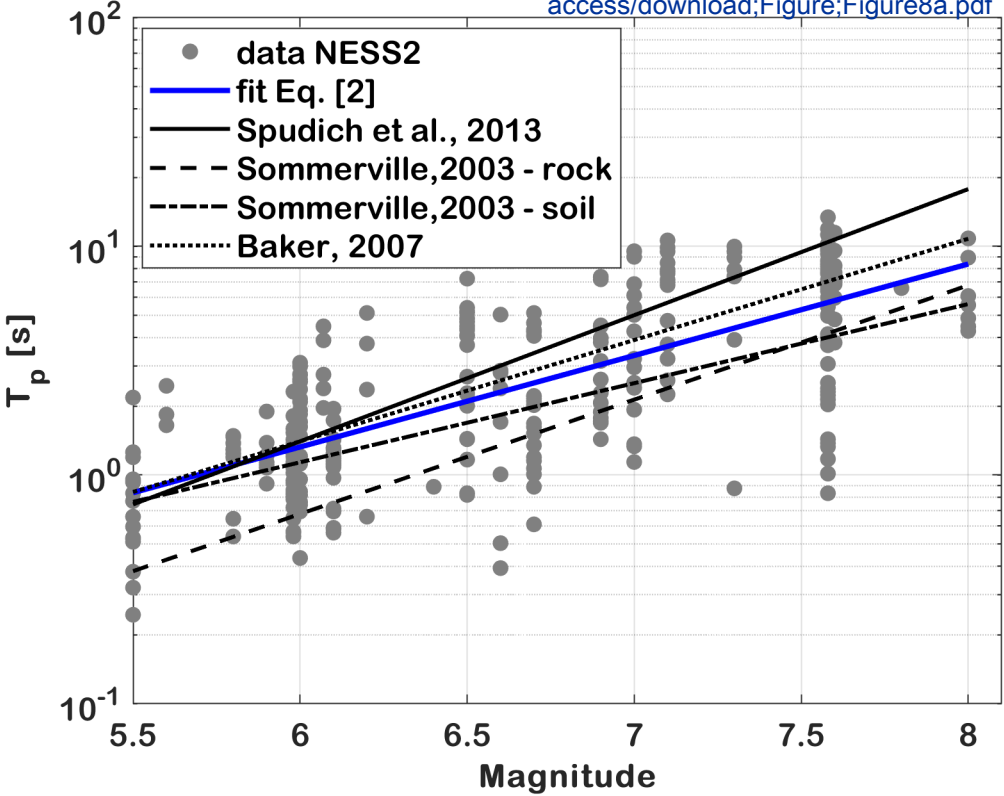


Figure 8b

Click here to access/download;Figure;Figure8b.pdf

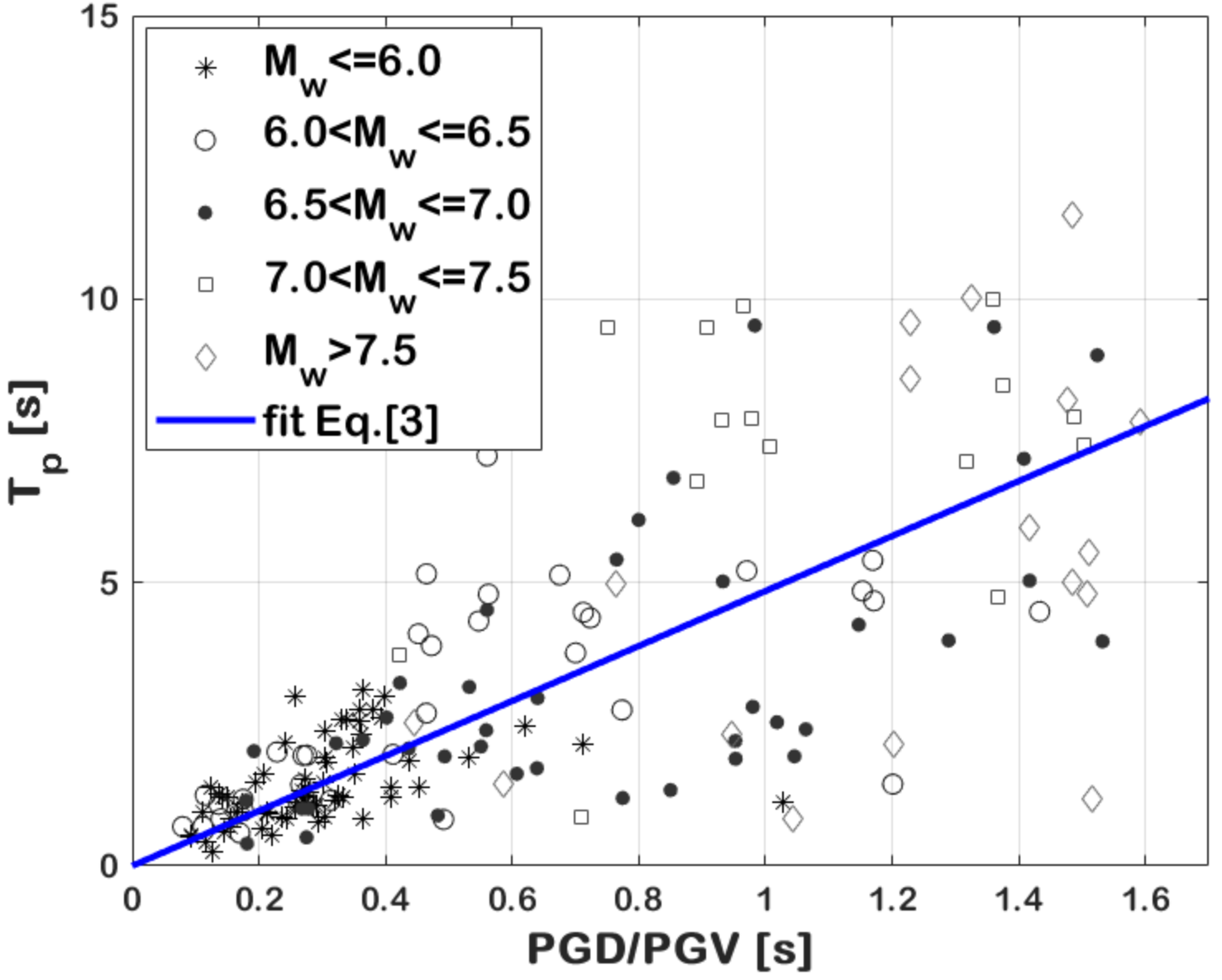


Figure 9a

Click here to access/download;Figure;Figure9a.pdf

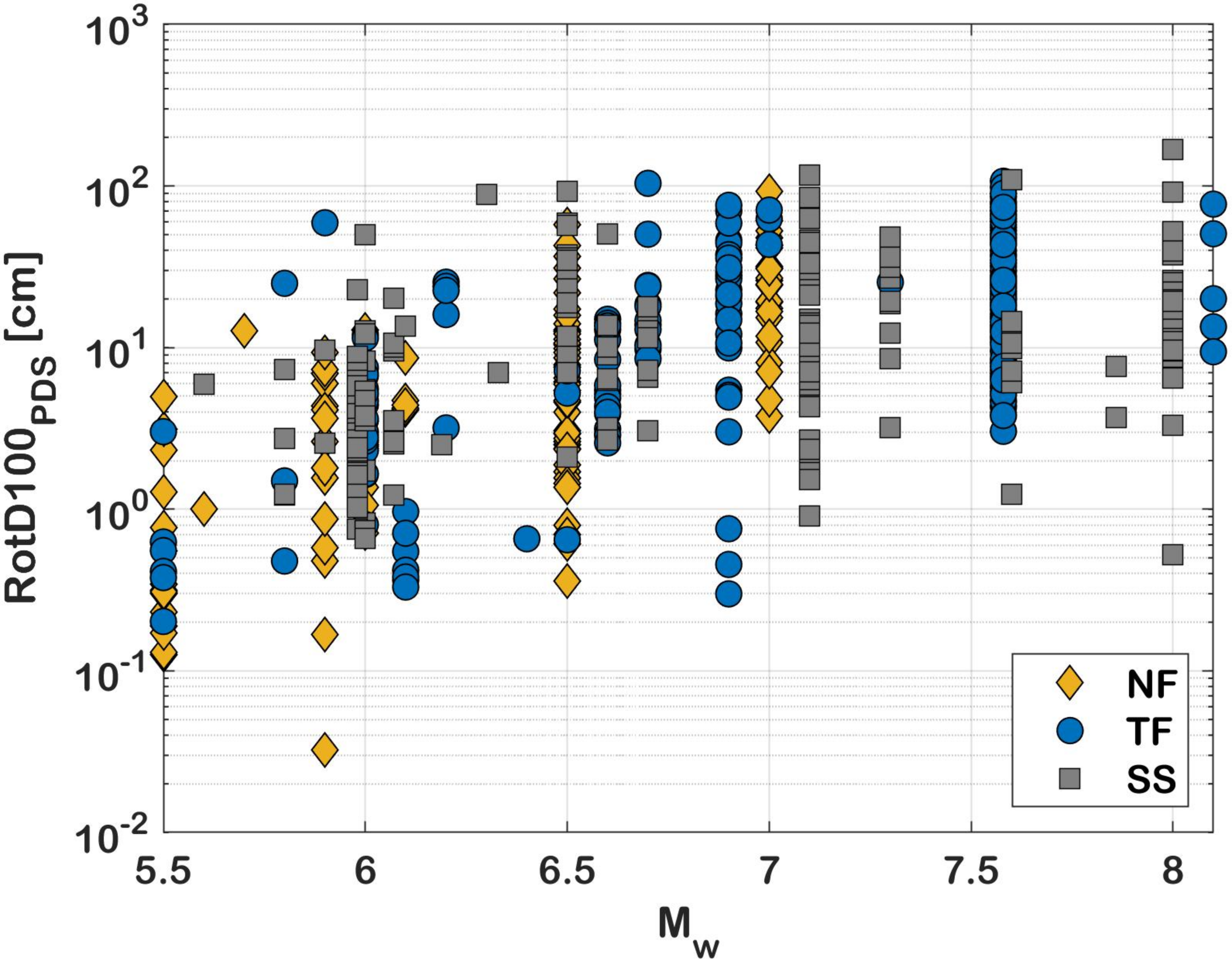


Figure 9b

Click here to access/download;Figure;Figure9b.pdf

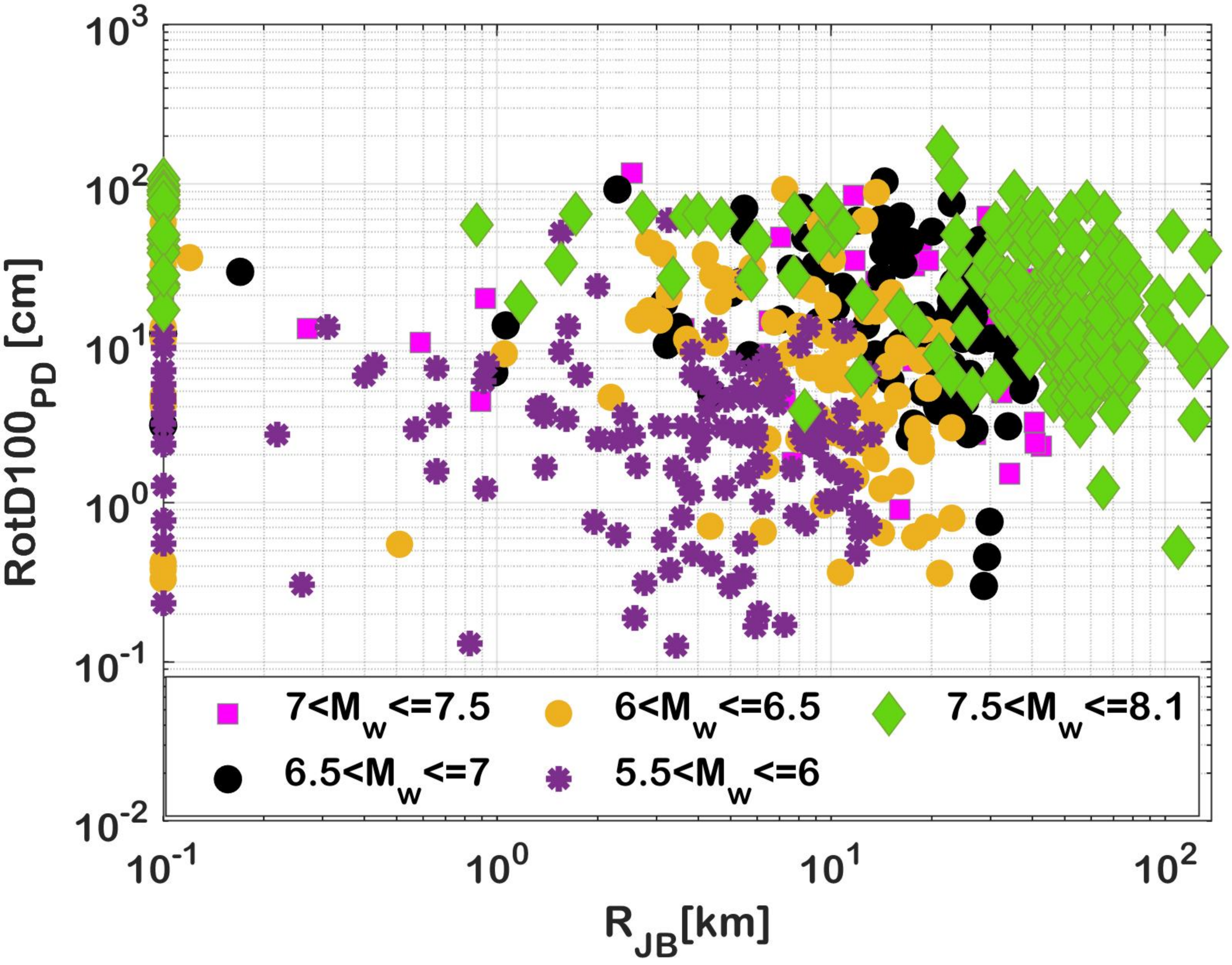


Figure 10a

Click here to access/download;Figure;Figure10a.pdf

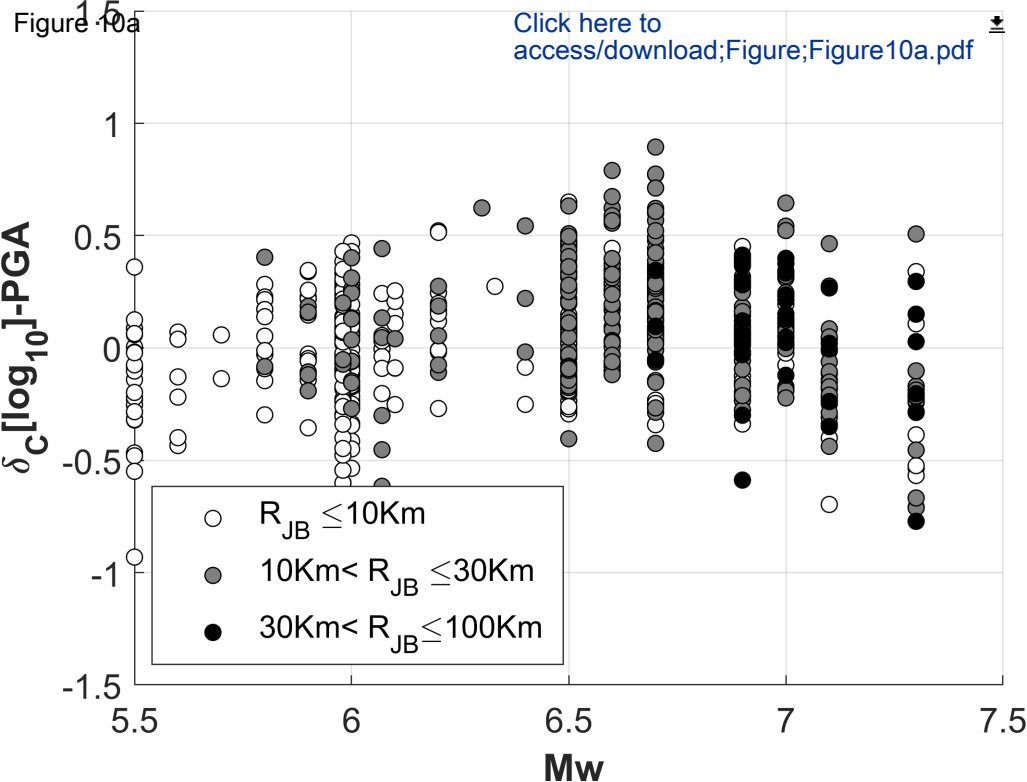


Figure 10b

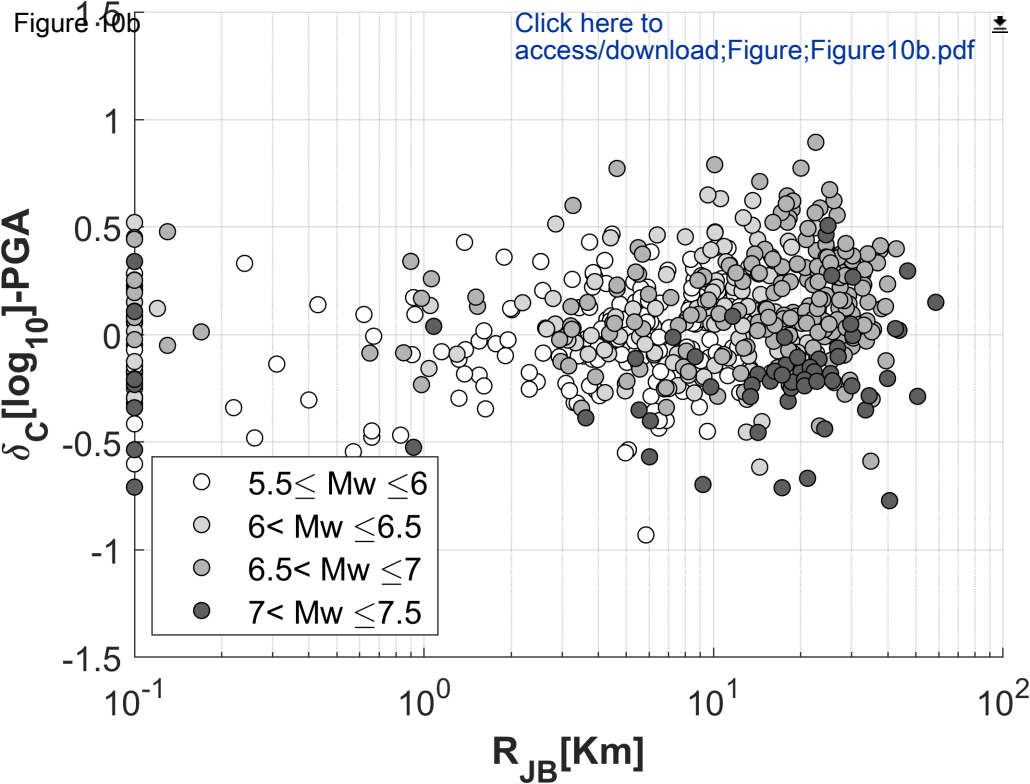


Figure 10c

Click here to access/download;Figure;Figure10c.pdf

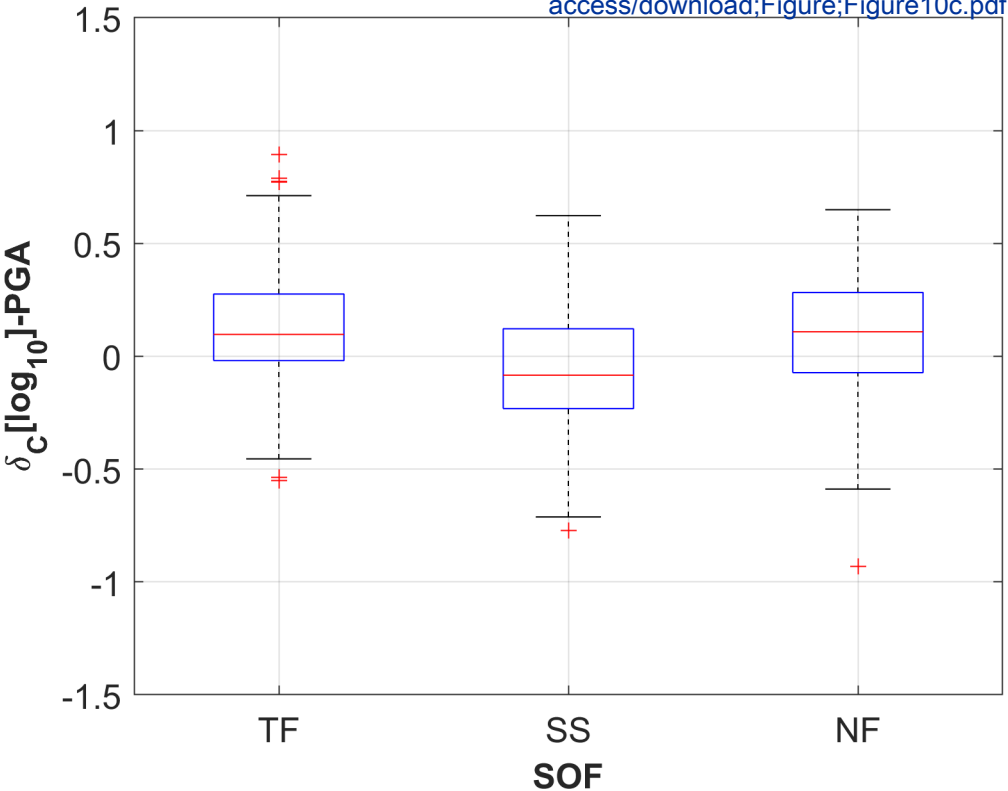


Figure 10d

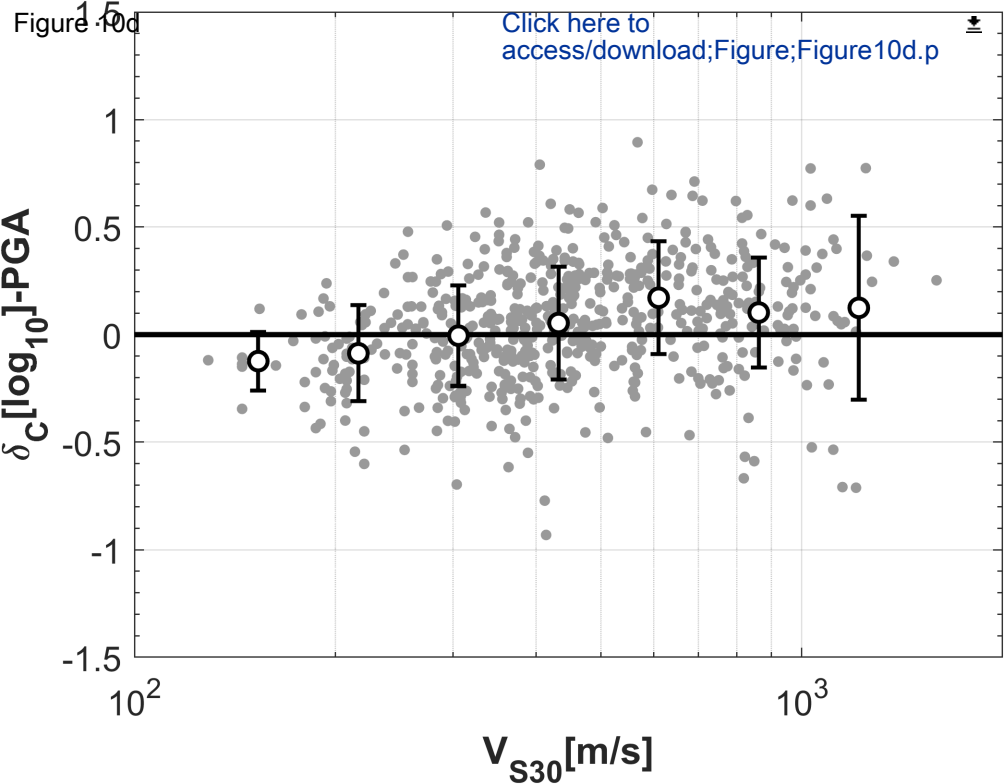


Figure 11a

Click here to access/download;Figure;Figure11a.pdf

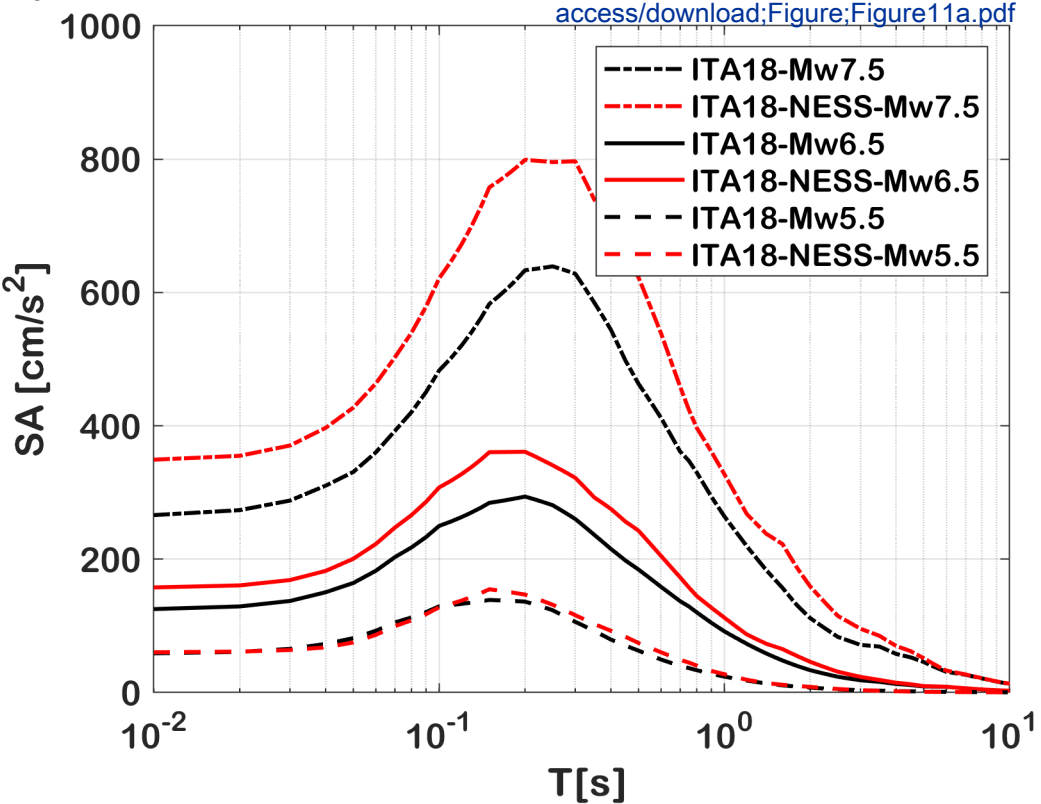


Figure 11b

Click here to access/download;Figure;Figure11b.pdf

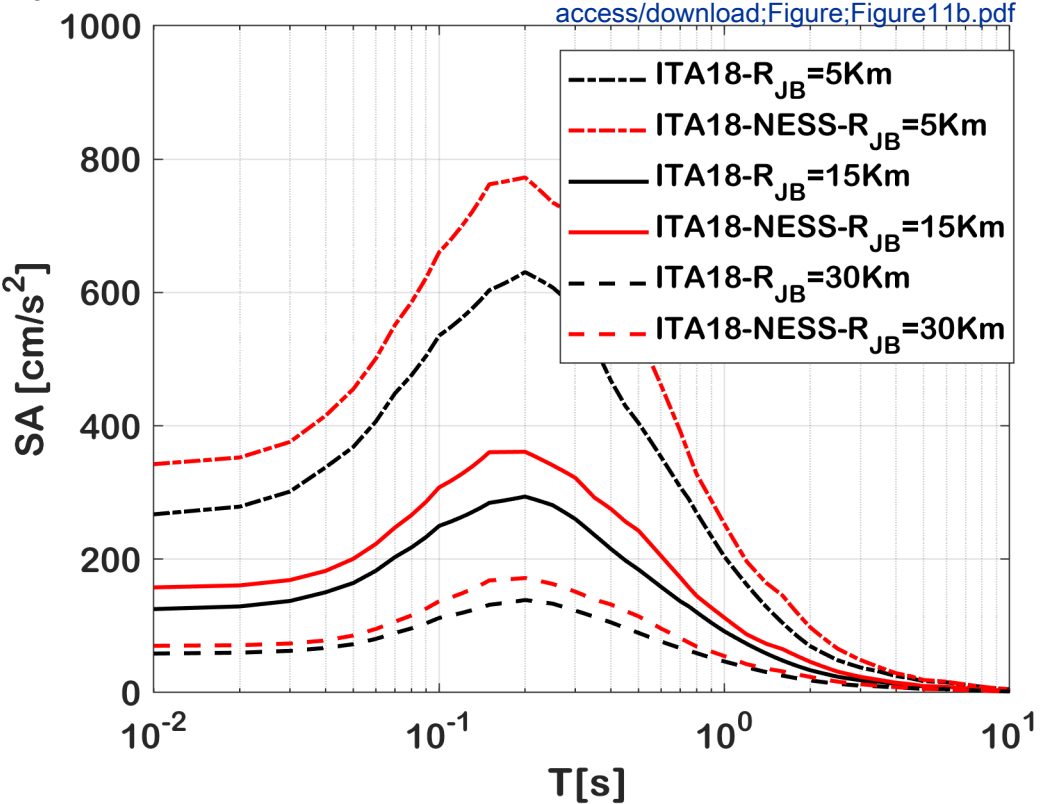


Figure 11c

Click here to access/download;Figure;Figure11c.pdf

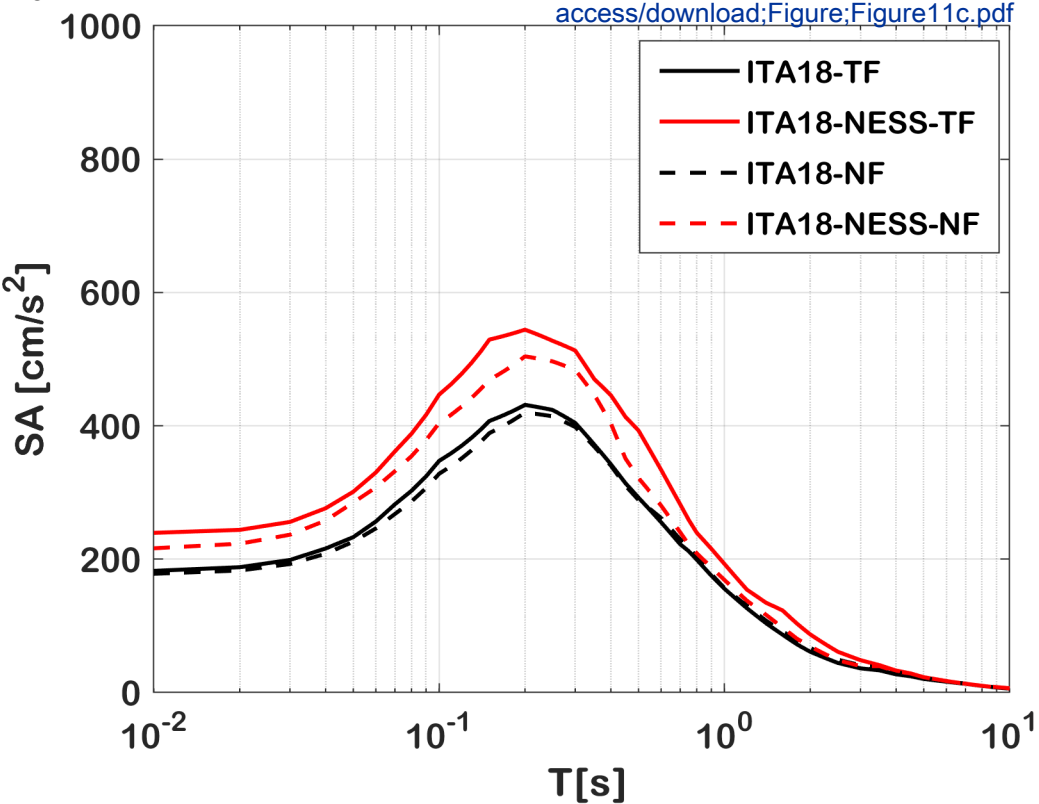


Figure 12a

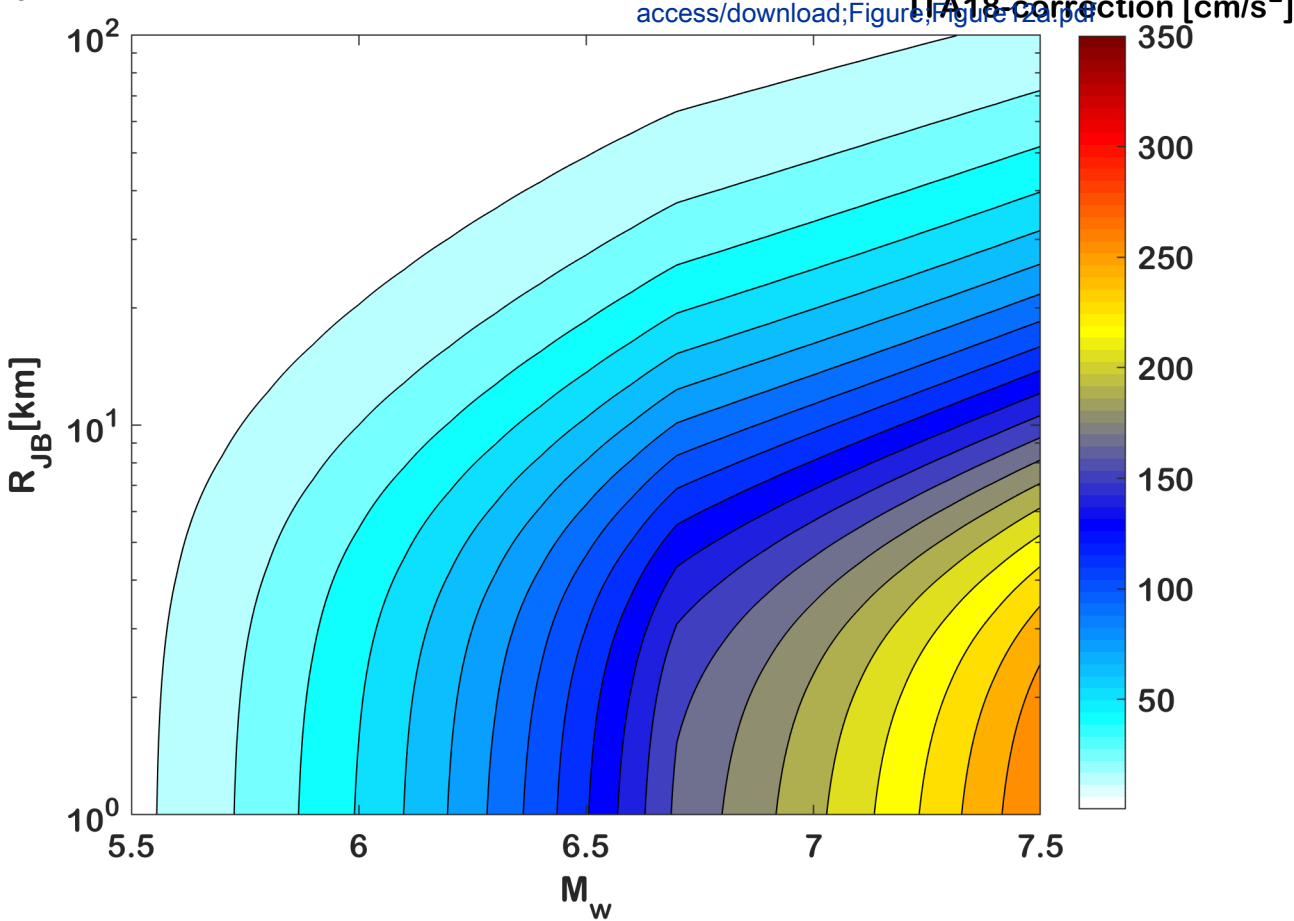


Figure 12b

Click here to access/download:Figure:Figure12b.pdf

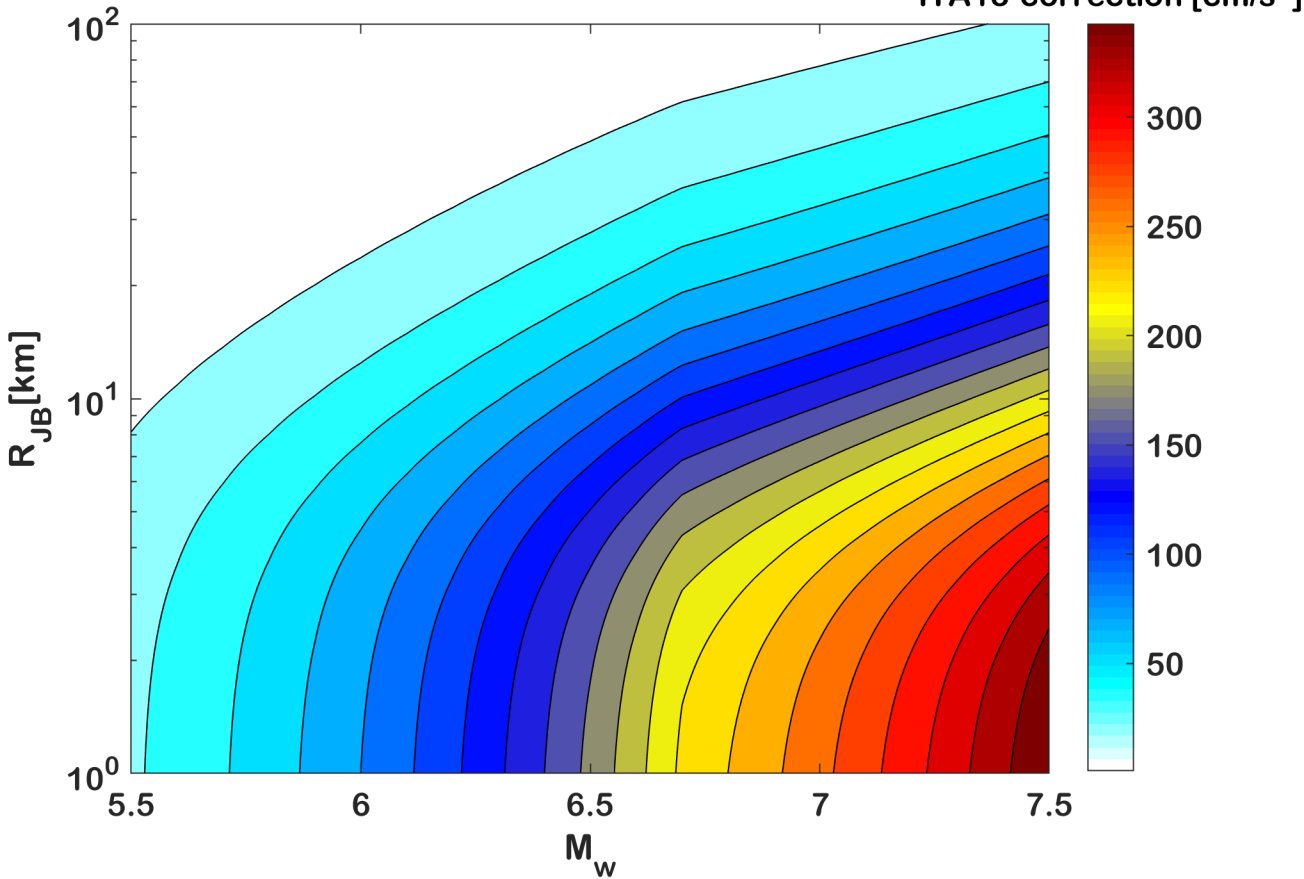


Figure 13a

Click here to access/download;Figure;Figure13a.pdf

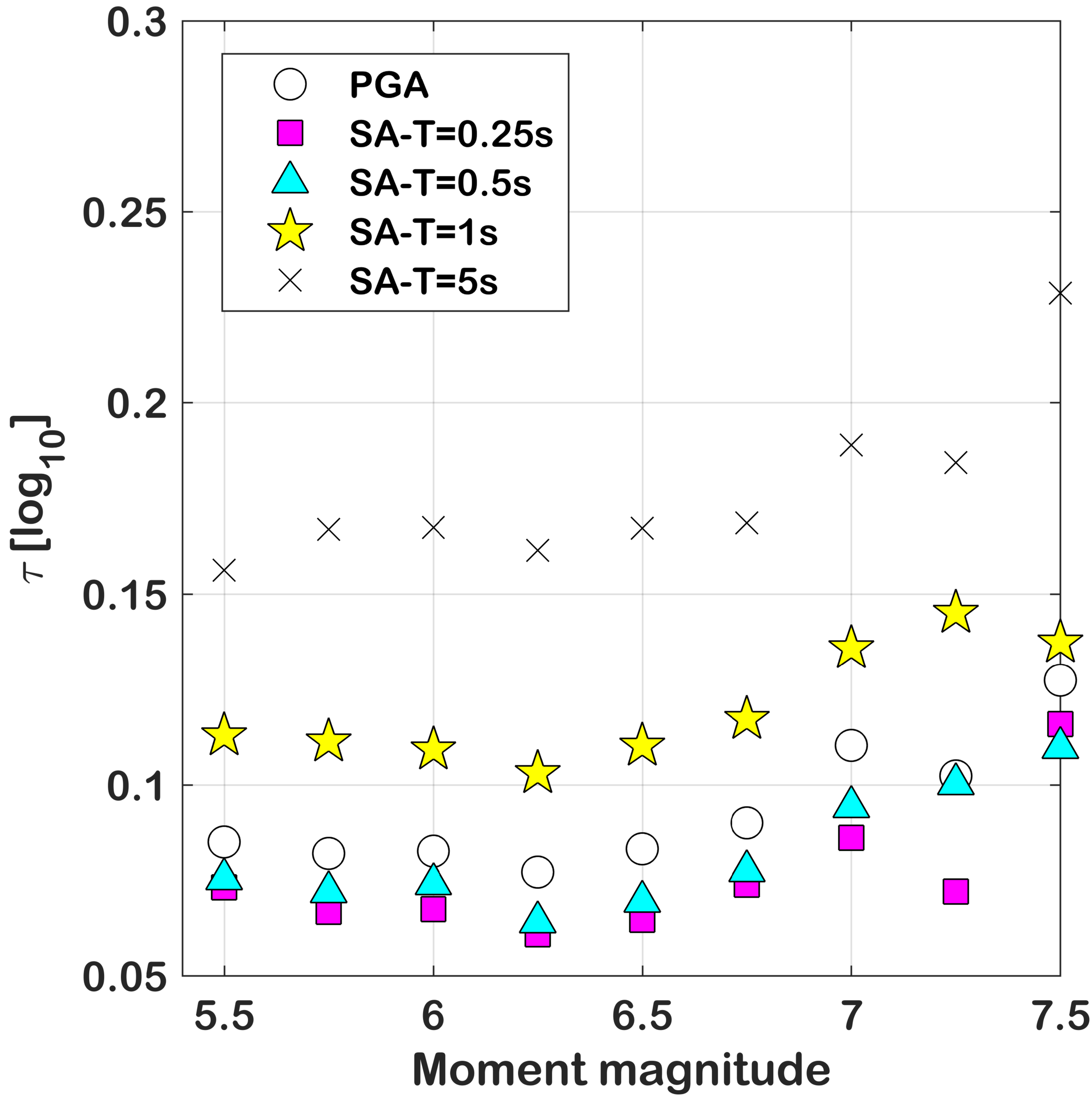


Figure 13b

Click here to access/download;Figure;Figure13b.pdf

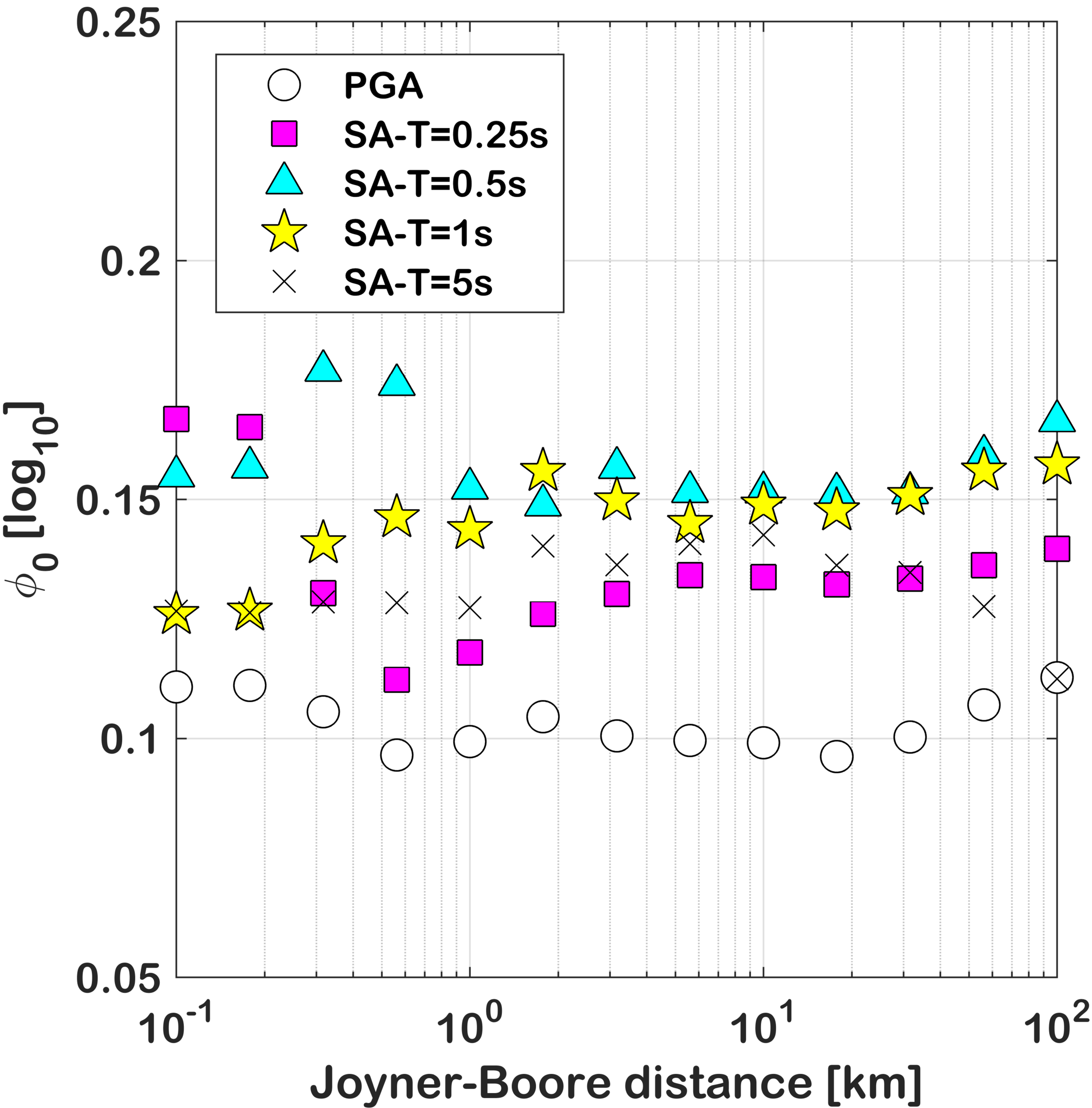


Figure 13c

Click here to access/download;Figure;Figure13c.pdf

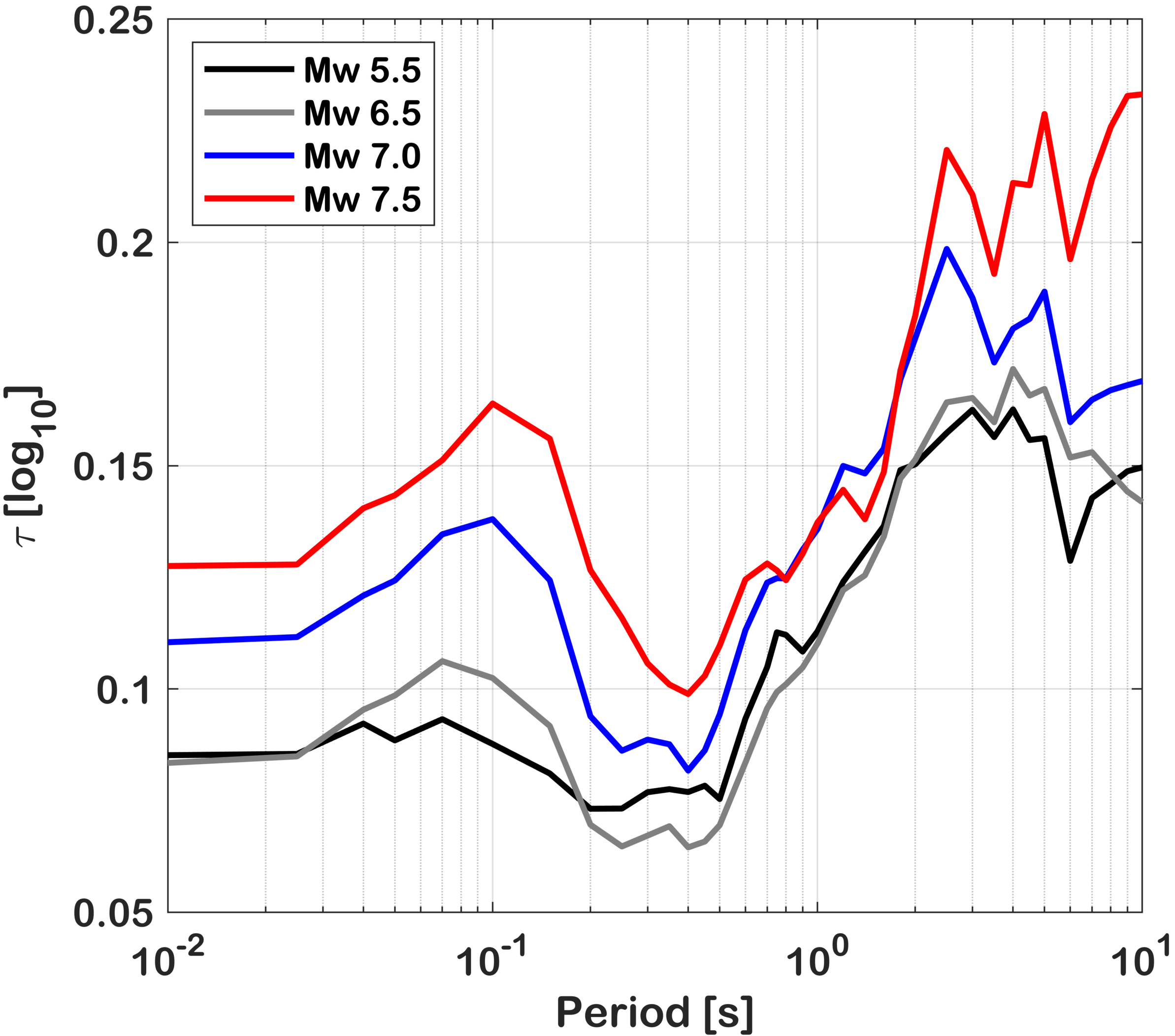


Figure 13d

Click here to access/download;Figure;Figure13d.pdf

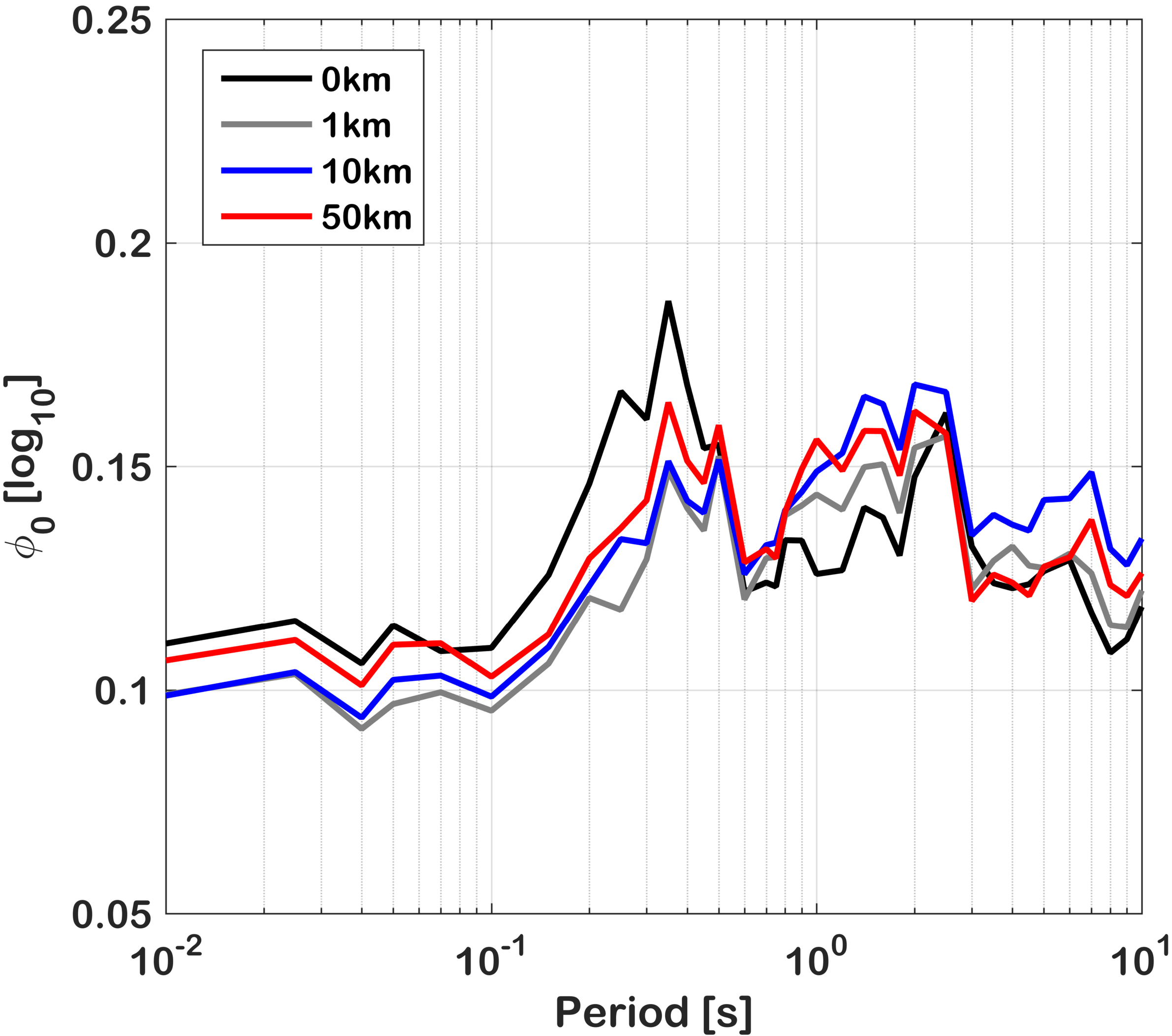


Figure 1A-a

Click here to access/download;Figure;Figure1Aa.pdf

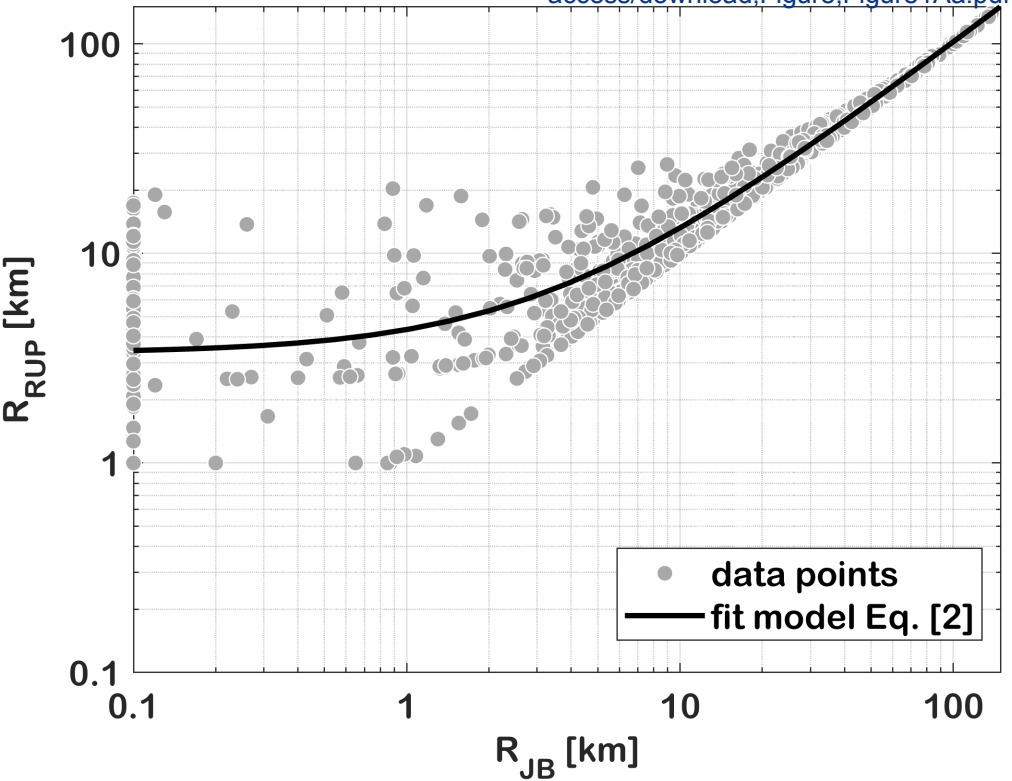
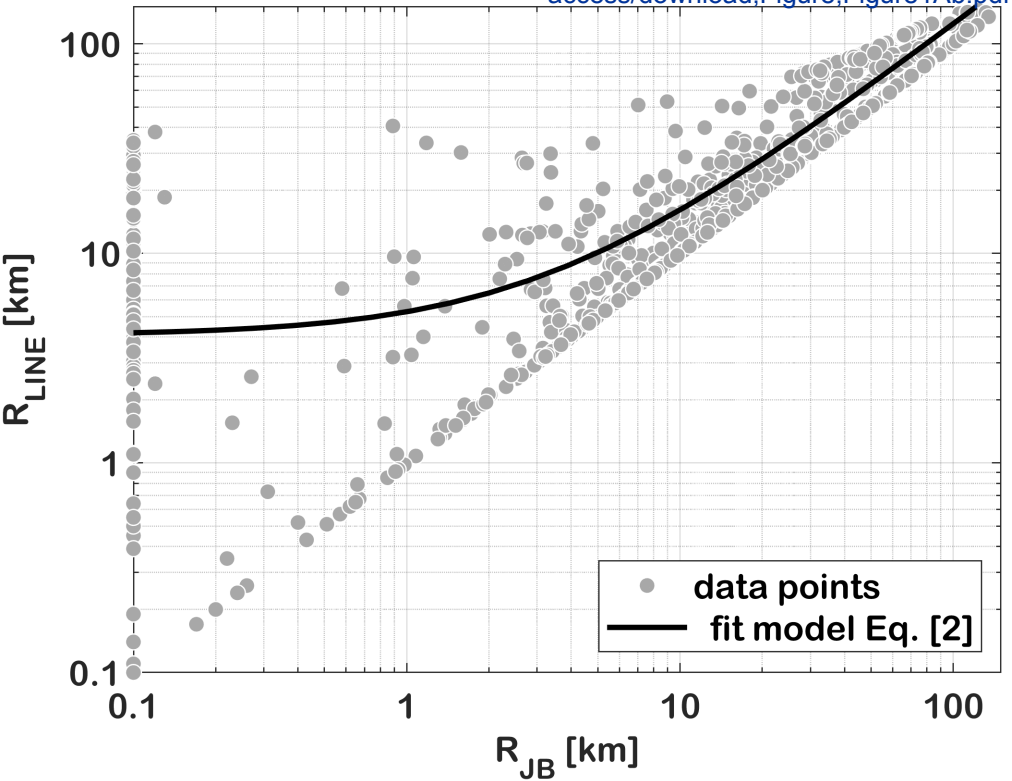


Figure 1A-b

Click here to access/download;Figure;Figure1Ab.pdf



Supplementary materials for: "NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground-motion models in near-source"

by Sara Sgobba, Chiara Felicetta, Giovanni Lanzano, Fadel Ramadan, Maria D'Amico, Francesca Pacor

supplement-S1: Flat-file (.CSV) of the permanent displacements recovered from NESS2-eBASCO dataset and coseismic displacements from GPS data

supplement-S2: Further details on the comparison between permanent displacements recovered from NESS2-eBASCO dataset and coseismic displacements from other validation data

supplement-S3: Total residuals vs ground motion parameters at SA(T=1s)

supplement-S4: A table containing the periods T[s], the model coefficients a_R, b_{1R}, c_R, f_{1R}, f_{2R}, h_R, M_{ref} and the statistical significant p-values, at each period T; i.e. p-Values(a_R), p-Values(b_{1R}), p-Values (c_R), p-Values(f_{1R}), p-Values(f_{2R})

supplement-S5: A table containing the periods T[s] and the standard deviations of between-event (tau), site-to-site (psi), event-and-site corrected (phi_0), total (sigma) in log-10 units

Supplementary material #S2

for: "NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground-motion models in near-source"

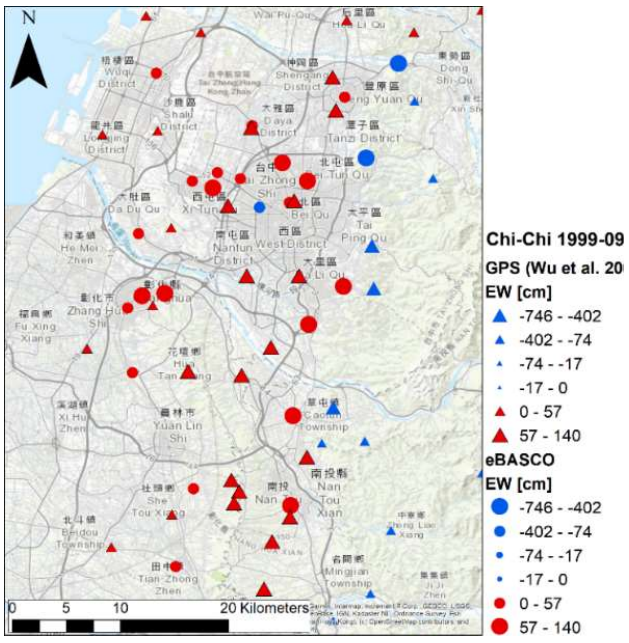
by Sara Sgobba, Chiara Felicetta, Giovanni Lanzano, Fadel Ramadan, Maria D'Amico, Francesca Pacor

Further details on the comparison between permanent displacements recovered from NESS2-eBASCO dataset and coseismic displacements from other validation data

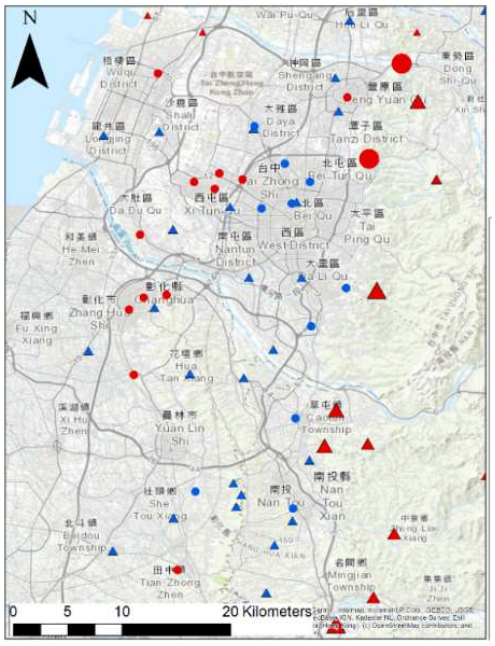
Figure S2-1 reports further comparisons between the permanent displacements (PDS) estimated with Extended BASeline Correction (eBASCO) algorithm and the coseismic values from GPS stations made available by Wu et al. (2001) for the Chi-Chi earthquake M_w7.58 1999-09-20 (Taiwan) horizontal EW (Figure S1a) and vertical components (Figure S1b).

In Figure S2-1 c,d fling-step benchmarks are represented by the ground-motion simulations of the Norcia event 2016-10-30 M_w6.5 (Italy). The simulated ground-motion is based on the 3D physics-based numerical model obtained by Özcebe et al. (2019).

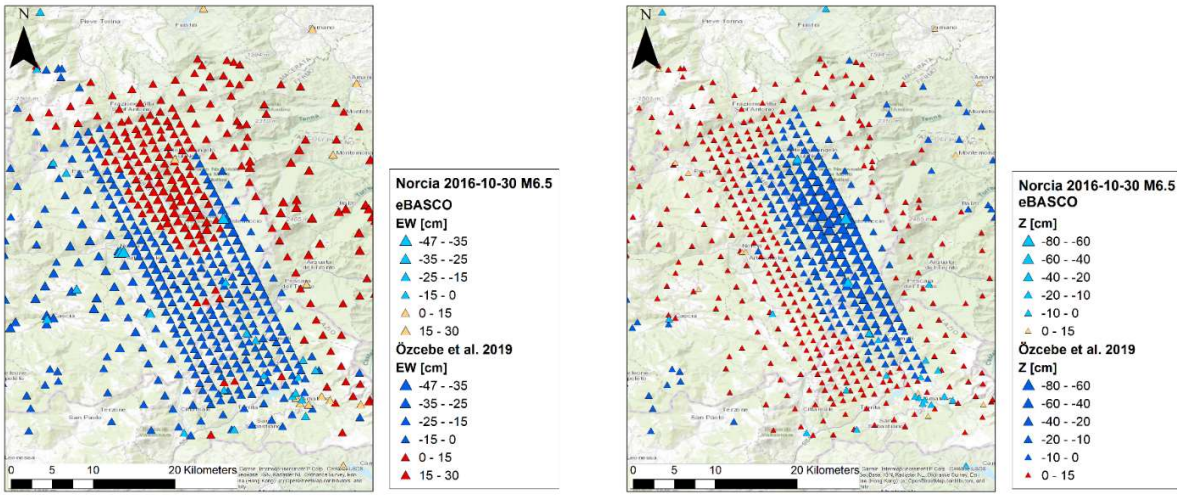
In Table S2-1 are compared the eBASCO-PDS estimates with the values based on 3D numerical simulations obtained by Evangelista et al. (2017) for the event of L'Aquila 2009-04-06 M_w6.1 (Italy).



(a)



(b)



(c) **(d)**

Figure S2-1. Static displacements of the Chi-Chi earthquake 1999-09-20 (Wu et al., 2001) recovered from NESS2 Extended BASeline Correction (eBASCO) compared with nearby GPS data, for the components EW (a) and Z (b); static displacements of Norcia earthquake 2016-10-26 from 3D numerical simulations (Özcebe et al., 2019) for the components EW (c) and Z (d).

Table S2-1. Permanent Displacements (PDS) Recovered from NESS2-eBASCO correction compared with PDS from 3D numerical simulation for L'Aquila earthquake 2009-04-06 (Evangelista et al., 2017) for the horizontal (EW – NS) and vertical (V) components.

				eBASCO-based PDS			Simulations-based PDS (Evangelista et al. 2017)		
Network	**Station**	**Latitude**	**Longitude**	**EW**	**NS**	**V**	**EW**	**NS**	**V**
		[°]	**[°]**	**[cm]**	**[cm]**	**[cm]**	**[cm]**	**[cm]**	**[cm]**
IT	AQA	42,38	13,34	2,62	-3,16	-3,09	2,3	-1,8	-2,8
IT	AQG	42,37	13,34	1,72	-4,39	-2,42	1,95	-1,8	-2,36
IT	AQK	42,34	13,40	0,77	-4,14	-13,81	3,35	1,67	-11,3
IT	AQV	42,38	13,34	1,50	-4,13	-3,20	2,68	-1,9	-3,25
IT	GSA	42,42	13,52	5,48	6,71	1,30	4,44	4,31	2,5

References

Wu, G., Minoru Takeo, Satoshi Ide; Source Process of the Chi-Chi Earthquake: A Joint Inversion of Strong Motion Data and Global Positioning System Data with a Multifault Model. Bulletin of the Seismological Society of America 2001; 91 (5): 1128–1143. doi: https://doi.org/10.1785/0120000713.

Özcebe, A. G., C. Smerzini, R. Paolucci, H. Pourshayegan, R. Rodríguez Plata, C.G. Lai, E. 675 Zuccolo, F. Bozzoni, and M. Villani (2019). On the comparison of 3D, 2D, and 1D numerical approaches to predict seismic site amplification: the case of Norcia basin during the M6.5 2016 October 30 earthquake, Proceedings of the 7th International Conference on Earthquake Geotechnical Engineering, Rome, 17-20 June 2019.

Evangelista, L., S. Del Gaudio, C. Smerzini, A. D'Onofrio, G. Festa, I. Iervolino, L. Landolfi, R. Paolucci, A. Santo, and F. Silvestri (2017). Physics-based seismic input for engineering applications: a case study in the Aterno River valley, Central Italy, Bulletin of Earthquake Engineering, 15, no. 7, 2645–2671.

Supplementary material #S3

for: "NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground-motion models in near-source"

by Sara Sgobba, Chiara Felicetta, Giovanni Lanzano, Fadel Ramadan, Maria D'Amico, Francesca Pacor

Total residuals vs ground motion parameters at SA(T=1s)

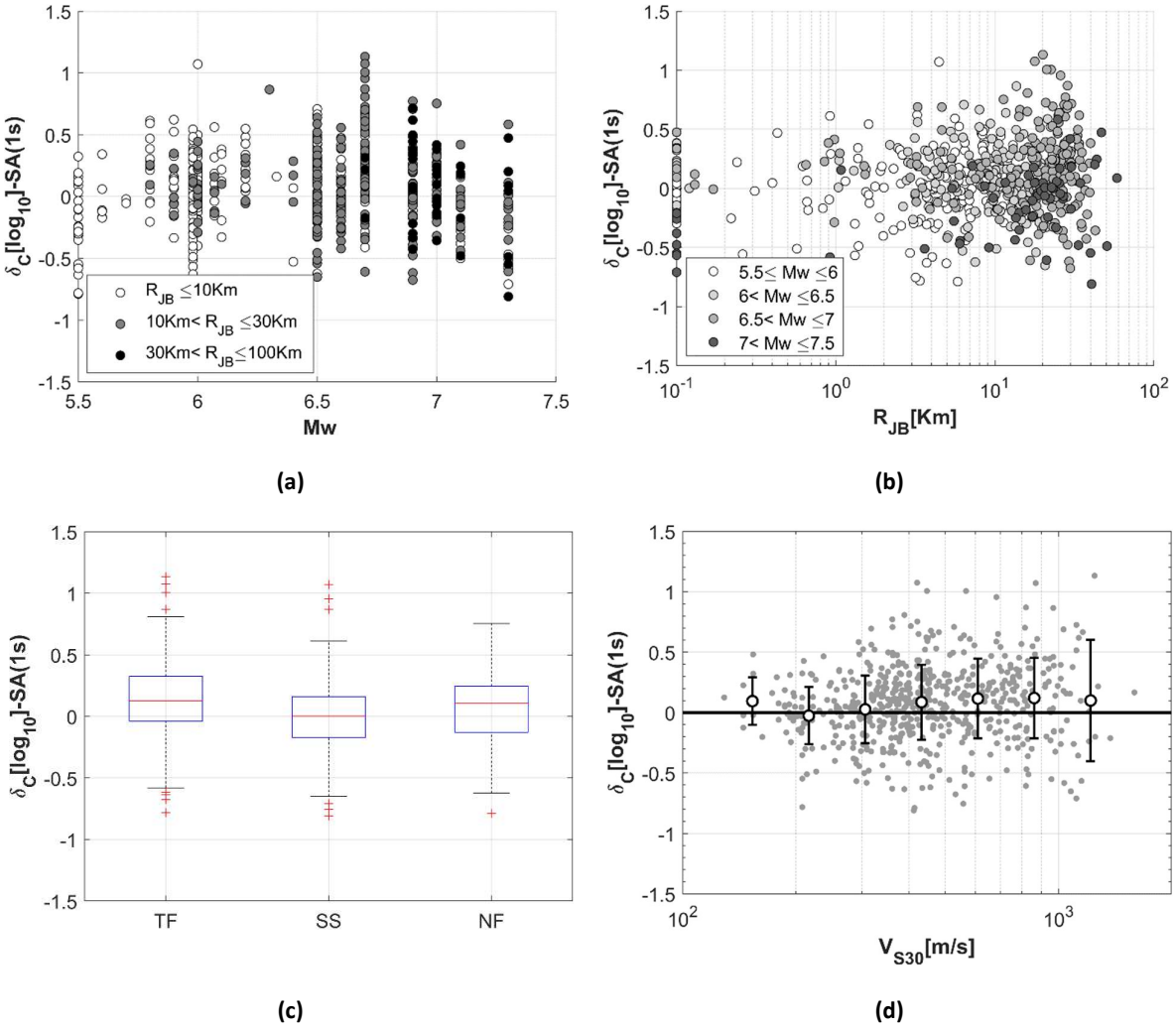


Figure S3-1. Total residuals δc, computed as logarithm difference between NESS2 observations and ITA18 predictions for SA(T=1s), as a function a) moment magnitude; b) Joyner-Boore distance; c) style of faulting; d) average shear-wave velocity in the uppermost 30m.

Supplemental Material (All Other Files, i.e. Movie, Zip, tar, csv)

supplement-S1.csv

Supplemental Material (All Other Files, i.e. Movie, Zip, tar, csv)

supplement-S4.csv

Supplemental Material (All Other Files, i.e. Movie, Zip, tar, csv)

supplement-S5.csv