RAPPORTI
TECNICI INGV

The new ITalian ACcelerometric Archive ITACA: database, web-services, and tools to access and analyze earthquakes waveforms

ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

429

ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

RAPPORTI
TECNICI INGV

The new ITalian ACcelerometric Archive ITACA: database, web-services, and tools to access and analyze earthquakes waveforms

Maria D'Amico, Chiara Felicetta, Emiliano Russo, Sara Sgobba, Giovanni Lanzano, Francesca Pacor, Lucia Luzi

INGV | Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Milano

429

INDICEX

Abstract

The growing interest in the open access of ground-motion data, web-services, and tools motivate every year a new release of the ITalian ACcelerometric Archive (ITACA), which represents the most complete collection of accelerometric waveforms recorded in Italy.
In the last years, ITACA has undergone an important renewal process resulting in a new layout of the website (http://itaca.mi.ingv.it) and substantial changes in the database content and services. In this work we present the main feature of the last release of ITACA v3.1 (last accessed on July 2020), providing Users with information to browse the website and to access and analyze data.

Keywords   Strong-motion; Earthquakes waveforms; Italy

Introduction

The ITalian ACcelerometric Archive (ITACA, https://itaca.mi.ingv.it) is the national database of accelerograms recorded since 1972 by the National Accelerometric Network (RAN, Rete Accelerometrica Nazionale), operated by the DPC (Dipartimento di Protezione Civile), the National Seismic Network (RSN, Rete Sismica Nazionale), operated by the INGV (Istituto Nazionale di Geofisica e Vulcanologia), and other regional and international networks operated by various providers. The ITACA project started in the 2004 in the framework of the Agreements between the Italian Department of Civil Protection (DPC) and the National Institute of Geophysics and Volcanology (INGV), funding two dedicated seismological projects: S6-Database of Italian accelerometric data for the period 1972-2004 and S4-The Italian accelerometric databank.

The goals of these projects were the collection, qualification and dissemination of the accelerometric waveforms acquired since the '70s by the National Agency for Electricity (ENEL), the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), and the DPC. The development of ITACA has continued over the time (Figure 1) driven by institutional activities in the framework of the Agreement between the INGV and the DPC (Annex A 2012 - 2021). In 2010, ITACA became the Italian infrastructure for accelerometric data distribution co-funded by the European Plate Observing System (EPOS, https://www.epos-eu.org/), a long-term project aimed to the creation of European research infrastructures for the solid Earth. Today, ITACA is the national node of the Engineering Strong Motion database (ESM, https://esm-db.eu/) which is designed to archive strong-motion data of magnitude M≥4.0 recorded in the European-Mediterranean and the Middle-East regions and distributed under the umbrella of Observatories & Research Facilities for European Seismology (ORFEUS, https://www.orfeus-eu.org/).

The different versions of ITACA released over the years (Figure 1) testify the great effort invested to its development and, in particular, to: 1) populate the database; 2) standardize waveforms and metadata formats according to the Federation of Digital Seismograph Networks (FDSN, https://www.fdsn.org/networks/); 3) process the signals by means of standard procedures; 4) create queries for exploring events, stations, and waveforms metadata; 5) access and download accelerometric waveforms and related metadata; 6) select suites of spectrum-compatible ground-motion waveforms; 7) provide seismological products useful for the calibration of ground-motion models.

The stakeholders of ITACA are researchers and students in the fields of applied seismology and earthquake engineering, professional engineers or geologists, and policy makers.
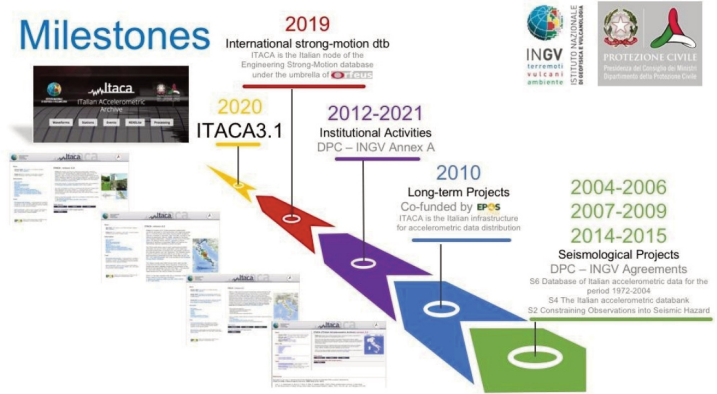
The number of users has grown over time, every day ITACA is visited by nearly 250 unique users. A considerable number of scientific papers is linked to the ITACA services or products [Luzi et al., 2009; Paolucci et al., 2010; Pacor et al., 2011; Bindi et al., 2011; Puglia et al., 2018, among others]. The growing interest to the open access to ground-motion data, web-services, and tools motivate every year a new ITACA release. Since ITACA version 3.0 (and thus including the current 3.1 released on April 2020, http://itaca.mi.ingv.it), the ITACA project features a renewed web interface (Figure 2) and substantial changes in the database content and services compared to the ITACA 2.x releases.

The architecture of the ITACA 3.1 database has been modified respect to the earlier versions (v2.x) in order to archive additional information in 100 tables, and is now managed by the open-source software PostgreSQL (https://www.postgresql.org/) which supersedes MySQL. The software that operates the database is entirely based on open-source software: the ITACA project uses custom Python (https://www.python.org/) software for data archival, management, and user authentication and AngularJS (https://angularjs.org/) for the web site design. Many routines are based on ObsPy (https://www.obspy.org), an open-source project with the aim to facilitate rapid application development for seismology providing a common Python framework. The new database architecture is constituted by three main branches (seismic events, recording stations, and accelerometric waveforms) and enables the association of multiple information related to the same object; for instance, a recording station can be linked to several geological maps and/or ambient noise measurements. In case of multiple sources of information, the database operators always recommend the most reliable one.

In this report, an overview of the main features of the ITACA 3.1 web site will be presented.

Figure 2 Home page of ITACA 3.1.

1. Statistics

ITACA 3.1 (last accessed July 2020) contains 45,170 accelerograms related to 1,946 seismic events with magnitude greater than 3.0, occurred in the time frame 1972-2019 and recorded by about 1,450 recording stations. More than 30,500 waveforms have been judged of good quality by specialists. A remarkable contribution is given by the main sequences occurred in the last decades in Italy (Figure 3): 2009 L'Aquila (April 6 - Mw6.1), 2012 Emilia (May 20 - Mw6.1, May 29 - Mw6.0), and 2016 Central Italy (August 24 - Mw6.0, October 26 - Mw5.9, October 30 - Mw6.5). The bulk of data relates to crustal events (depth<30 km) with epicentral distance 10 - 100 km in the magnitude interval 3.5 – 4.5. Given the predominance of middle and low magnitudes, the majority of focal mechanisms are undefined, while for the largest events (Mw>6) the normal faulting prevails (Figure 3).

The stations are classified according to Eurocode 8 (EC08 - CEN, 2204) on the basis of the time-average shear-wave velocity in the uppermost 30 m (Vs30) or of the surface geological information. The distribution of stations according to the EC8 soil classification is represented in Figure 3f. About 314 stations are classified as EC8-A (21.6%), 688 in EC8-B (47.5%), and 391 stations are in EC8-C (27%) class. A lower percentage of recording sites (3.6%) are classified in EC8-D or EC8-E soil classes. For very few stations (0.3%) the soil classification is still undefined. The 21% of the EC8 classes are assigned according to Vs30 estimates derived from measured 1D velocity profile and 71% on the base of the surface geological information. When velocity profiles or geological maps are missing (8%), the Vs30 value is estimated using empirical correlations with topographic slope [Wald and Allen, 2007].

The ITACA dataset is mainly characterized by maximum horizontal (RotD100) peak ground acceleration lower than 1 cm/s^2 (72%), while the 0.9% of waveforms overcomes 100 cm/s^2 that is particularly significant for engineering applications.
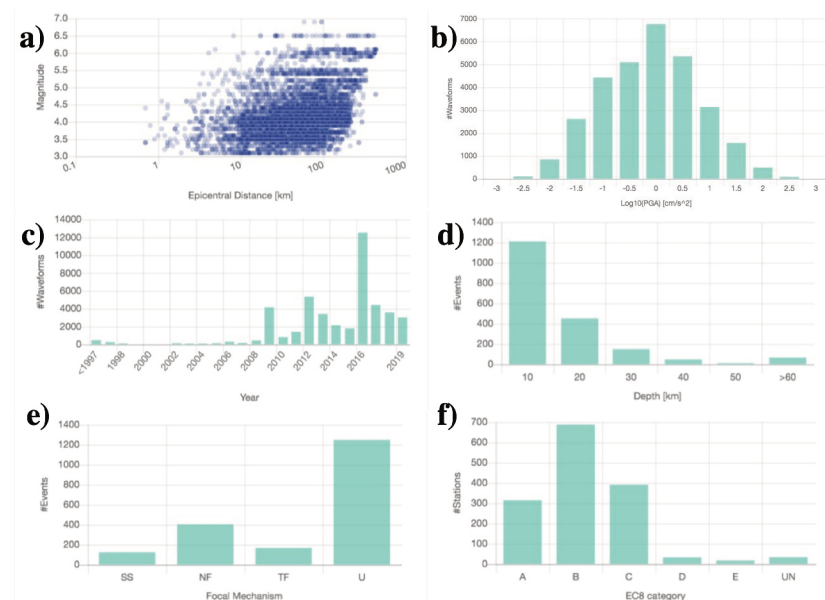
Figure 3 Statistics of ITACA 3.1 for waveforms, stations, and events: a) magnitude-distance distribution; b) histogram of the RotD100 peak ground accelerations; c) number of waveforms recorded per year; d) number of earthquakes as a function of the Epicentral distance; e) number of earthquakes as a function of the style of faulting; f) soil classification according to the Eurocode 8.

The strong-motion waveforms have been recorded in Italy and surrounding regions by the National Accelerometric Network (RAN), operated by the DPC, the National Seismic Network (RSN), operated by the INGV, and by regional or international networks operated by various providers (Figure 4).

Figure 4 Main data providers in ITACA 3.1: 1st - IT, Italian Strong-Motion Network (RAN); 2nd - IV, Italian National Seismic Network (INSN); 3rd - RA, Réseau Accélérométric Permanent (RAP); 4th - 3A, Centro di microzonazione sismica Network, 2016 Central Italy seismic sequence.

2. Website

The website of ITACA 3.1 is available in two languages, Italian and English, and provides various tools to explore the database content (i.e. glossary, pop-up). As in the earlier ITACA versions, the website contains three main sections related to waveforms, ev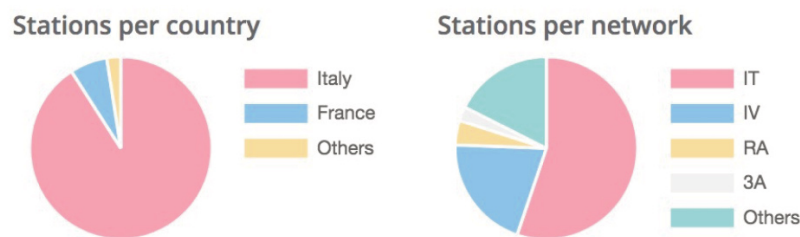ents, and recording stations. The ITACA 3.1 website navigation is facilitated by "tabs" which enable the User to switch between station, events, and waveforms sections and to query the database by means of several key fields. The queries outcomes are provided as tables or maps. OpenStreet is the map layer for the visualization of events and recording stations.

2.1 Waveforms

The *Waveform Search* page (Figure 5) manages the information related to the accelerometric records available in the ITACA 3.1 database. The waveform query can be made through a *quick* or *advanced search*. The *quick search* allows selecting waveforms records for some relevant key-fields: magnitude, epicentral distance, PGA (Peak Ground Acceleration), PGV (Peak Ground Velocity), ITACA ID, date time, station code, station name, method to estimate Vs30/EC8, EC8 code. The advanced search includes all the key-fields useful to query ITACA records specifically related to *Waveforms*, *Events*, and/or *Stations metadata* (Table 1).

Figure 5 Waveforms search in ITACA 3.1: 1) link to the authentication system; 2) tabs to query accelerometric records to the ITACA database by means of quick or advanced search; 3) buttons to execute or refresh the queries; 4) button to download a parametric table in TXT format with intensity measures and metadata of the required waveforms; 5) embedded table with relevant metadata and parameters of the selected records.

	Key-fields	Description	Example
Waveform	Epicentral distance (km)	Distance on the ground surface between the site and the earthquake epicenter	From: 10 To: 50
	Rjb (km)	Minimum distance between the site and the fault projection on the ground surface	From: 10 To: 50
	PGA (cm/s^2)	Maximum ground acceleration recorded during the seismic shaking by the three components of the seismic station related to processed waveforms	From: 100 To: 300
	Unprocessed PGA (cm/s^2)	Maximum ground acceleration recorded during the seismic shaking by the three components of the seismic station related to unprocessed waveforms	From: 10 To: 50
	PGV (cm/s)	Maximum ground velocity recorded during the seismic shaking by the three components of the seismic station related to unprocessed waveforms	From: 5 To: 20
	PGD (cm)	Maximum ground displacement recorded during the seismic shaking by the three components of the seismic station related to unprocessed waveforms	From: 5 To: 10
	Housner Intensity (cm)	Integral of the 5%-damped pseudo-velocity response spectrum over the period interval [0.1, 2.5]	From: 20 To: 50
	T90 (s)	Time interval between the points at which 5% and 95% of the total energy has been recorded	From: 0 To: 100
	Arias Intensity (cm/s)	Parameter proportional to the integral of the square of the acceleration history over the duration of the record	From: 20 To: 50
	Late trigger	Flag (Yes or No) to specify if the recording instrument has been triggered early enough or not to properly describe the first arrival of the seismic waveform	No
	Pulse-like	Feature to define if the waveform is characterized by double-sided pulse in the velocity trace	Pulse-like behavior

Event	ITACA ID	Alphanumeric code to identify the seismic event	Contains: EMSC
	Event name	Name associated to the seismic event	Contains: Italy
	Date Time	Origin time of the seismic event	From: 2016-08-24 To: 2019-01-01
	Style of faulting	Geometry of fault rupture during an earthquake	Normal faulting
	Latitude	Latitude in decimal degrees (WGS84) of the earthquake epicenter	From: 42.0 To: 43.5
	Longitude	Longitude in decimal degrees (WGS84) of the earthquake epicenter	From: 13 To: 14
	Hypocentral depth (km)	Distance between the hypocenter and the epicenter of the earthquake	From: 0 To: 30
	Magnitude	Measures the intensity of the seismic event (ML, Mw)	From: 5.5 To: 6.5
	Country	To select only the seismic events localized within the country border	Italy
	Region	To select only the seismic events localized within the region border	Abruzzo
	Province	To select only the seismic events localized within the province border	L_Aquila
	Municipality	To select only the seismic events localized within the municipality border	L_Aquila
Station	Network	The code associated to the recording Seismic Network according to the International Federation of Seismograph Network (FDSN)	IT - Italian Strong Motion Network (RAN)
	Station code	3 to 5 characters that identify the recording instrument	CLO
	Station name	Name associated to the recording station	Castelluccio di Norcia
	Latitude	Latitude of the station in decimal degrees (WGS84)	From: 42.0 To: 43.0
	Longitude	Longitude of the station in decimal degrees (WGS84)	From: 12.0 To: 15.0
	Country	To select only the recording stations within the country border	Italy
	Region	To select only the recording stations within the region border	Umbria
	Province	To select only the recording stations within the province border	Perugia
	Municipality	To select only the recording stations within the municipality border	Norcia
	Method to estimate Vs30/EC8	To select recording stations with Vs30/EC8 estimate by means of: a) geophysical survey; b) surface geology; c) topography	Surface geology
	EC8 code	To select recording stations according to Eurocode8	C
	Vs30 (m/s)	Average measure of the shear waves propagation velocity in the soil, within the first 30 meters of depth from the ground level	From: 180 To: 360

Table 1 List of the *Waveform Search Key-fields*.

2.1.1 Authentication system

The User may have access to data and services distributed through the ITACA 3.1 web site, like *Anonymous* or *Registered User* by an authentication system (Figure 6). In the former case, just the metadata exploration is allowed; in the latter case, the authentication is mandatory to navigate and/or download: i) waveforms together with related spectra and metadata; ii) parametric table (flat-file) user defined with metadata and intensity measures of selected records. The authentication can be accomplished compiling the ITACA registration form and approving the *Terms of Use*, as well as by using a Google or ORCID account (Figure 6).
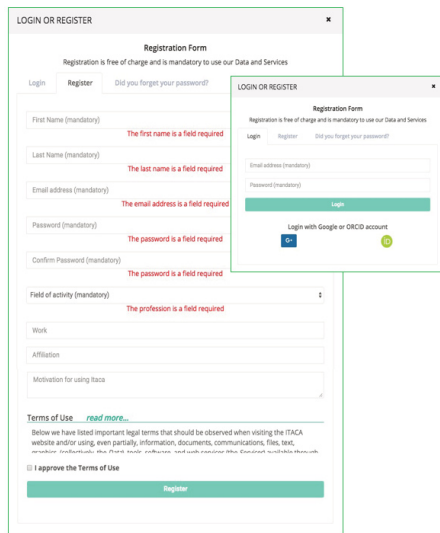

Figure 6 Registration form in ITACA 3.1. The User may have access to data and services through: i) ITACA registration form; ii) Google; iii) ORCID.

2.1.2 User Defined Flat-file

A parametric table (flat-file) with station and event metadata and intensity measures (peak ground acceleration, peak ground velocity, peak ground displacement, energy content, duration parameters, and spectral acceleration ordinates assuming 5% damping in the period range 0.01-10s) of accelerometric records contained in ITACA 3.1 can be downloaded from the waveforms page. The User can customize the ITACA flat-file by using the key-fields of the waveforms search. An optional software (filter_flatfile.sh) enable the User to filter the flat-file according to fields (columns) and metadata (e.g. event, station, network) and it is designed to run over different operating systems (e.g. Windows through Cygwinm MacOSX, Linux). The flat-file is distributed as TXT file (semicolon separated) and the complete list of the flat-file fields is available in the related *User manual* (see also Table 2). The column headers "y_yyy", in particular, identifies the period corresponding to the spectral ordinate (where the "_" indicates the decimal separator, e.g. 0_100 corresponds to T=0.100s). The waveform components are flagged as U, V or W, where U stands for the 1st horizontal component, V for the 2nd horizontal component and W for the vertical component.

Fields	Description	Example
event_id	The internal id of the event in ITACA	EMSC-20161030_0000029
event_time	Origin time of the event (format YYYY-MM-DDTHH:MM:SS)	2016-10-30T06:40:18

ev_nation_code	ISO code of the country where the epicenter of the event is located (see Table A1)	IT
ev_latitude	Latitude of the epicenter of the event in decimal degrees (WGS84)	42.8322
ev_longitude	Longitude of the epicenter of the event in decimal degrees (WGS84)	13.1107
ev_depth_km	Depth of the hypocenter of the event (km)	9.2
fm_type_code	Style of faulting (NF normal fault; TF reverse fault; SS strike-slip fault; O oblique fault)	NF
ML	Local magnitude	6.1
Mw	Moment magnitude	6.5
network_code	The code associated to the recording network according to the International Federation of Seismograph Network (http://www.fdsn.org)	IT
station_code	3 to 5 characters to identify the station	CLO
st_nation_code	ISO code of the country where the station is located (see Table A1)	IT
location_code	two characters' code that identifies the sensor location (see http://www.fdsn.org/seed_manual/SEEDManual_V2.4.pdf, pag 37) corresponding to the "Sensor ID" on the ITACA web interface	00
sensor_depth_m	Sensor depth (m), where negative values indicate meters above the ground surface	0
proximity	Description of the proximity to structures (e.g., free-field, close to structure or inside structure)	Inside structure
housing	Description of the station housing (e.g., fiberglass box, building, cave or vault)	Building
installation	Description of the installation conditions (e.g., pillar, building basement or directly on the ground)	Building basement
st_latitude	Latitude of the station in decimal degrees (WGS84)	42.829399
st_longitude	Longitude of the station in decimal degrees (WGS84)	13.206
st_elevation	Elevation of the station (m) in meters above the sea level	13.206
preferred_ec8_code	The most reliable EC8 site category (CEN, 2003) according to expert judgment	C
preferred_vs30_m_s	The most reliable estimate of the average shear wave velocity VS,30 (CEN, 2003) according to expert judgment	—
method_ec8_vs30	The method used to estimate "preferred_ec8_code" and/or "preferred_vs30_m_s	Surface geology
epi_dist	Epicentral distance (km)	7.8
JB_dist	Joyner-Boore distance (km)	0
rup_dist	Distance from the rupture plane (km)	1.92

instrument_type_code	The first letter specifies the general sampling rate and the response band of the instrument; The second letter specifies the family to which the sensor belongs (http://www.fdsn.org/seed_manual/SEED Manual_V2.4.pdf)	HG
instrument_type	'Analog' or 'Digital' instrument type	Digital
X_channel_code	Channel code of the X component according to the SEED convention http://www.fdsn.org/seed_manual/SEEDM anual_V2.4.pdf, pag. 134)	E
acceleration_units	The unit of acceleration is always cm/s^2	cm/s^2
processing_status	Indicates if the waveform has been processed or not, or if the waveform has been flagged as a bad quality record. In the latter case, no ground motion intensity measures are provided, with the exception of "X_un_PGA" (uncorrected PGA)	processed
processing_type	Indicates if the waveform is manually or automatically processed and the processing scheme adopted. At the moment, the standard processing is modified after Paolucci et al. [2011] and described in Puglia et al. [2018]	manual [Paolucci et al., 2011]
late_triggered_flag_01	Flag for late triggered waveforms (triggered means the record starts at the S waves arrivals)	—
X_hp	High pass filter frequency of the X component	0.1
X_lp	Low pass filter frequency of the X component	30
pulse_like_code	Flag for the waveforms that present a pulse in the velocity time histories according to the procedure by Shahi and Baker [2014]	1
X_un_pga	Unprocessed peak ground acceleration of the X component	- 419.554382
X_pga	Peak ground acceleration of the X component (cm/s^2)	- 418.616203
X_pgv	Peak ground velocity of the X component (cm/s)	571.423851
X_pgd	Peak ground displacement of the X component (cm)	-14.128563
X_T90	Duration (s) of the time interval between the points of 5% and 95% of the total energy [Trifunac and Brady, 1975] of the X component	9.64
X_housner	Housner intensity (cm) of the X component (cm)	234.313
X_ia	Arias intensity (cm/s) of the X component (cm)	245.755991166
X_Ty_yyy	5% damping acceleration spectral ordinates for the y_yyy period and the X component	419.15
Original_data_mediator	The service that provides the original raw data	http://www.mot1.it/ran download/

Table 2 Search fields of the User Defined Flat-file of ITACA 3.1.

2.1.3 Waveforms and Spectra Plots

From the list of the outcomes at the bottom of the *Waveform search* page (Figure 7a), the *User* can look into Waveform detail page, where the main waveform metadata area summarized (Figure 7b). The waveform and the spectra plots are visible in the related tabs only after logging to the website. Figure 7c,d show the waveforms and the spectra plots of the 2016, Mw6.5 Norcia earthquake recorded by the IT.CLO, respectively.

Figure 7a Example of waveform selection in case of the 2016, Mw6.5 Norcia earthquake recorded by the IT.CLO station: 1) link to the *Waveform detail* page of the selected record; 2) link to the *Event detail* page; 3) link to the *Station detail* page.

Figure 7b Example of *Waveform detail* with overall information about the 2016, Mw6.5 Norcia earthquake recorded by the IT.CLO station. The User can view the waveforms plots of acceleration (unprocessed and processed), velocity, displacement, and spectra (Fourier, 5% damped Pseudo Acceleration and Displacement response spectra).
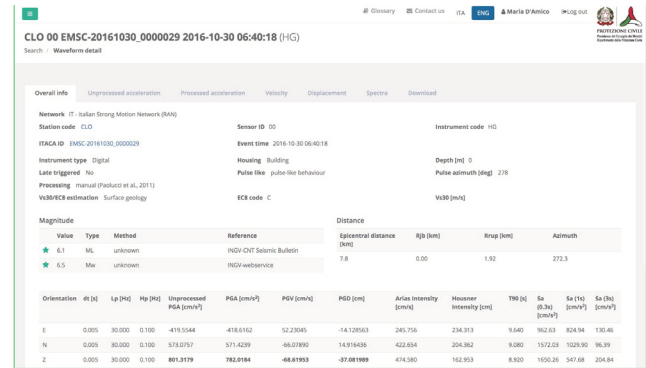

Figure 7c Example of *Waveform plot* of the 2016, Mw6.5 Norcia earthquake recorded by the IT.CLO station: a) unprocessed acceleration; b) acceleration; c) velocity; d) displacement.
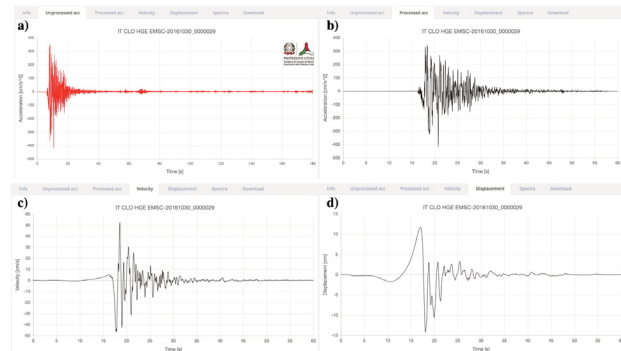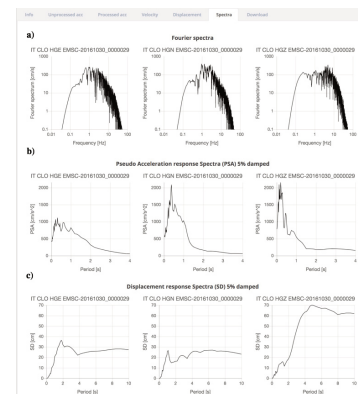

Figure 7d Example of *Spectra* related to the 2016, Mw6.5 Norcia earthquake recorded by the IT.CLO station: a) Fourier spectra; b) Pseudo Acceleration response Spectra (PSA) 5% damped; c) Displacement response Spectra (SD) 5%.

2.1.4 Download and Data Format

Upon *User* registration, waveforms data can be accessed from the *Download* tab of the *Waveform detail* page in ASCII (American Standard Code for Information Interchange) or ASDF (Adaptable Seismic Data Format, see http://seismic-data.org/) format (Figure 8). Waveforms data are unprocessed acceleration time series, processed acceleration, velocity, and displacement time series, or acceleration and displacement response spectra (5% damping) calculated at 105 periods (0.01–10 s).

The structures of ASCII (Figure 9a) and ASDF tags (Figure 9b) are defined by a sequence of codes (Table 3). According to the Standard for the Exchange of Earthquake Data (SEED, http://www.fdsn.org/pdf/SEEDManual_V2.4.pdf) the first four codes (*net_code*, *station_code*, *location_code*, *channel_code*) identify the recording instrument, while the other ones relate to the seismic event (*event_id*), the data processing (*processing_type*), the ground-motion parameter type (*file_type*), and the data format (*data_format*).
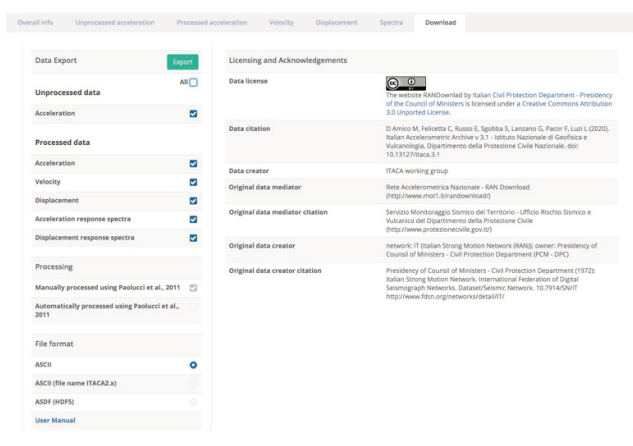

Figure 8 Example of data download from the *Waveform detail* page of ITACA 3.1. After logging in the website, the User can download unprocessed acceleration waveforms, acceleration waveforms processed by the procedure described in Paolucci et al. [2011] together with related velocity, displacement, and 5% damped acceleration and displacement response spectra. The ITACA 3.1 data are available in ASCII or ASDF (HDF5) formats.

Figure 9a Example of ASCII filename in ITACA 3.1 for the unprocessed (CV) acceleration (ACC) waveform recorded by the sensor 00 of the SGIUB station belonging to the IT network during the seismic event IT-2002-0075. The instrument is a High Broad Band (H) accelerometer (N) North-South (N) oriented.

Figure 9b Example of ASDF tag in ITACA 3.1 for the unprocessed (*cv*) acceleration (*acc*) waveform recorded by a High Broad Band (*h*) accelerometer (*n*) North-South (*n*) oriented during the seismic event *it_2002_0075* linked to the sensor *00*.

Placeholders	Description	ASCII example	ASDF example
net_code	code associated to the recording network according to the International Federation of Seismograph Network (http://www.fdsn.org)	*IT*	*it*
station_code	3 to 5 characters to identify the recording station	*SGIUB*	*—-*
location_code	0 to 2 characters to identify a specific recording instrument at a station; Note that double-zeros are always replaced by empty string in ASCII filenames	*00*	*00*
channel_code	3 digits to indicate the band code, the instrument code (N, L, G for accelerometer), and the orientation code (e.g. Z, N, E for the specific orientation Vertical, North-Shout, East-West or Z, 2, 3 for generic orthogonal orientation)	*HNN*	*hnn*
event_id	alphanumeric code to identify the seismic event	*IT-2002-0075*	*it_2002_0075*
file_type	2 to 3 characters to indicate the ground motion parameter of time series (ACC: acceleration; VEL: velocity; DIS: displacement) or 5% damped response spectra (SA: acceleration; SD: displacement)	*ACC*	*acc*
processing_type	2 characters to indicate the type of data processing (CV: acceleration ConVerted in physical units; MP or AP: acceleration Manually or Automatically processed using the workflow proposed in Paolucci et al. [2011]). In the earlier versions of ITACA (v2.x) the *processing_type* of unprocessed and processed acceleration was X or C, respectively	*CV*	*cv*

Table 3 Description of ASCII and ASDF tag for waveforms and spectra downloadable from the Waveform detail page of ITACA 3.1.

The ASCII format contains a metadata header of 64 rows that describes the event, the recording site, and the instrument metadata. Processing parameters are also included to ensure the proper use of data. Each waveform or spectrum distributed by ITACA is associated to an individual *Data license*, if indicated by the data provider, and opportunely acknowledged.

Line	Field	Example
1	EVENT_NAME	*CENTRAL_ITALY*
2	EVENT_ID	*EMSC-20161030_0000029*
3	EVENT_DATE_YYYYMMDD	*20161030*
4	EVENT_TIME_HHMMSS	*064018*
5	EVENT_LATITUDE_DEGREE	*42.8322*
6	EVENT_LONGITUDE_DEGREE	*13.1107*
7	EVENT_DEPTH_KM	*9.2*
8	HYPOCENTER_REFERENCE	*INGV-CNT_Seismic_Bulletin*
9	MAGNITUDE_W	*6.5*
10	MAGNITUDE_W_REFERENCE	*INGV-webservice*
11	MAGNITUDE_L	*—*

12	MAGNITUDE_L_REFERENCE	—
13	FOCAL_MECHANISM	*Normal faulting*
14	NETWORK	*IT*
15	STATION_CODE	*CLO*
16	STATION_NAME	*Colfiorito*
17	STATION_LATITUDE_DEGREE	*42.829399*
18	STATION_LONGITUDE_DEGREE	*13.206*
19	STATION_ELEVATION_M	*1456*
20	LOCATION	*00*
21	SENSOR_DEPTH_M	*0*
22	VS30_M/S	—
23	SITE_CLASSIFICATION_EC8	*C (inferred from geology)*
24	MORPHOLOGIC_CLASSIFICATION	—
25	EPICENTRAL_DISTANCE_KM	*7.8*
26	EARTHQUAKE_BACKAZIMUTH_DEGREE	*272.3*
27	DATE_TIME_FIRST_SAMPLE_YYYYM-MDD_HHMMSS	*20161030_064003.768*
28	DATE_TIME_FIRST_SAMPLE_PRECISION	*milliseconds*
29	SAMPLING_INTERVAL_S	*0.005*
30	NDATA	*12000*
31	DURATION_S	*60*
32	STREAM	*HGE*
33	UNITS	*cm/s^2*
34	INSTRUMENT	*sensor = Unknown [Unknown]	digitizer = Unknown [Unknown]*
35	INSTRUMENT_ANALOG/DIGITAL	*D*
36	INSTRUMENTAL_FREQUENCY_HZ	—
37	INSTRUMENTAL_DAMPING	—
38	FULL_SCALE_G	—
39	N_BIT_DIGITAL_CONVERTER	—
40	PGA_CM/S^2	*-418.616211*
41	TIME_PGA_S	*20.775*
42	BASELINE_CORRECTION	*BASELINE REMOVED*
43	FILTER_TYPE	*BUTTERWORTH*
44	FILTER_ORDER	*2*
45	LOW_CUT_FREQUENCY_HZ	*0.1*
46	HIGH_CUT_FREQUENCY_HZ	*30*
47	LATE/NORMAL_TRIGGERED	*NT*
48	DATABASE_VERSION	*0.5*
49	HEADER_FORMAT	*DYNA 1.2*
50	DATA_TYPE	*ACCELERATION*
51	PROCESSING	*manual [Paolucci et al., 2011]*
52	DATA_TIMESTAMP_YYYYMMDD_HH-MMSS	*20200407_094011.326*

53	DATA_LICENSE	CC-BY3_0-IT (http://creativecommons.org/licenses/by/3.0/deed.en)
54	DATA_CITATION	D'Amico, M. C., Felicetta, C., Russo, E., Sgobba, S., Lanzano, G., Pacor, F., Luzi, L., [2020, April 2]. ITalian ACcelerometric Archive (ITACA), version 3.1 (Version 3.1). Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/ITACA.3.1
55	DATA_CREATOR	ITACA working group
56	ORIGINAL_DATA_MEDIATOR_CITATION	Ufficio Rischio Sismico e Vulcanico del Dipartimento della Protezione Civile (http://www.protezionecivile.gov.it/)
57	ORIGINAL_DATA_MEDIATOR	Rete Accelerometrica Nazionale - RAN Download (http://www.mot1.it/randownload/)
58	ORIGINAL_DATA_CREATOR_CITATION	Presidency of Counsil of Ministers - Civil Protection Department (1972): Italian Strong Motion Network. International Federation of Digital Seismograph Networks. Dataset/Seismic Network. 10.7914/SN/IT http://www.fdsn.org/networks/detail/IT/
59	ORIGINAL_DATA_CREATOR	Network IT (Italian Strong Motion Network (RAN)); owner: Presidency of Counsil of Ministers - Civil Protection Department (PCM - DPC)
60-64	USER1-5	—

Table 4 List of the 64 fields which compose the header of the ITACA ASCII file.

ASDF format [Krischer et al., 2016] is a modern file format specifically tailored for researchers and analysts in seismology that allows to create a comprehensive data set, including all the available metadata related to waveforms and spectra (https://seismic-data.org/). The intent is to facilitate the interoperability with several programming languages such as C/Fortran or Python. The ASDF format is conceived so that the ITACA data are organized in a hierarchical structure inside a HDF5 [Hierarchical Data Format version 5; The HDF Group 1997–2015] container. The ITACA HDF5 container is organized in three sections related to events, waveforms, and auxiliary data (Figure 10). The ASDF volume can be read by dedicated softwares (e.g. HDFView, https://support.hdfgroup.org/products/java/hdfview/; ASDF SEXTANT, https://github.com/SeismicData/asdf_sextant) or programming languages such as Python or Matlab. An example of ASDF data set reading with the Python Application Programming Interface *pyasdf* (https://seismicdata.github.io/pyasdf/) is described in the following.

Once acquired an ASDF file from the download window of ITACA 3.1 (Figure 8), the data set can be parsed to an ObsPy object using the ASDFDataSet class of the *pyasdf* module (Figure 11a). Information about the seismic event (e.g. date time, epicentral coordinates, magnitude) are stored in a QuakeML [Schorlemmer et al., 2004; 2011] inside the HDF5 container (Figure 11b).
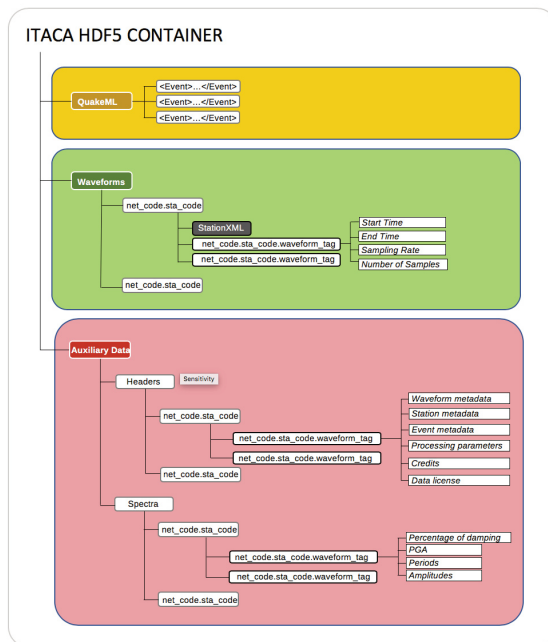
Figure 10 Structure of the ITACA ASDF file in the HDF5 container (modified after Krischer et al. [2016]): yellow) information about an arbitrary number of seismic events stored in a single QuakeML; green) seismic waveforms are stored in a FDSN StationXML; red) response spectra and metadata related to stations, events, waveforms, processing, credits, and license are hierarchically stored in the auxiliary data section.

```
[In [1]: import pyasdf

[In [2]: ds = pyasdf.ASDFDataSet('query.h5', mode='r')

[In [3]: print(ds)
ASDF file [format version: 1.0.2]: 'query.h5' (1.0 MB)
        Contains 1 event(s)
        Contains waveform data from 1 station(s).
        Contains 2 type(s) of auxiliary data: Headers, Spectra
```

Figure 11a Example of reading an ITACA 3.1 ASDF data set containing information about seismic events, waveforms recorded by one or more stations, and auxiliary data (e.g. response spectra, processing parameters, credits).

```
In [4]: type(ds.events)
Out[4]: obspy.core.event.catalog.Catalog

In [5]: print(ds.events)
1 Event(s) in Catalog:
2016-10-30T06:40:18.000000Z | +42.832,  +13.111 | 6.5 mw

[In [6]: print(ds.events[0])
Event:  2016-10-30T06:40:18.000000Z | +42.832,  +13.111 | 6.5 mw

                     resource_id: ResourceIdentifier(id="smi:esm.mi.ingv.it/fdsnws/event/1/query?event_id=EMSC-20161030_0000029")
                      event_type: 'earthquake'
            event_type_certainty: 'known'
                   creation_info: CreationInfo(agency_id='INGV', author='Engineering Strong-Motion database (ESM)', author_uri=Res
              preferred_origin_id: ResourceIdentifier(id="smi:esm.mi.ingv.it/fdsnws/event/1/query?origin_id=EMSC-20161030_0000029-C
           preferred_magnitude_id: ResourceIdentifier(id="smi:esm.mi.ingv.it/fdsnws/event/1/query?magnitude_id=EMSC-20161030_000002
      preferred_focal_mechanism_id: ResourceIdentifier(id="smi:esm.mi.ingv.it/fdsnws/event/1/query?focal_mechanism_id=EMSC-20161030_
                          ---------
              event_descriptions: 2 Elements
                          origins: 1 Elements
                       magnitudes: 1 Elements
```

Figure 11b Example of reading event metadata from an ITACA 3.1 ASDF data set.

Accelerometric waveforms are stored per single station together with a FDSN (International Federation of Digital Seismograph Network) StationXML (https://www.fdsn.org/xml/station/). Each waveform is a piece of a continuous time-series (hereinafter *trace*) regularly sampled and recorded by a seismic instrument within the time interval start time – end time (Figure 10). Every station can contain an arbitrary number of traces related to multiple channels, ground-

motion parameters, and data processing (Figure 11c). Each trace is named according to the following scheme (hereinafter *waveform tag*):

'location_code'_'channel_code'_'event_id'_'file_type'_'processing_type'

```
In [7]: ds.waveforms.IT_CLO
Out[7]:
Contents of the data set for station IT.CLO:
    - Has a StationXML file
    - 12 Waveform Tag(s):
        00_hge_emsc_20161030_0000029_acc_cv
        00_hge_emsc_20161030_0000029_acc_mp
        00_hge_emsc_20161030_0000029_dis_mp
        00_hge_emsc_20161030_0000029_vel_mp
        00_hgn_emsc_20161030_0000029_acc_cv
        00_hgn_emsc_20161030_0000029_acc_mp
        00_hgn_emsc_20161030_0000029_dis_mp
        00_hgn_emsc_20161030_0000029_vel_mp
        00_hgz_emsc_20161030_0000029_acc_cv
        00_hgz_emsc_20161030_0000029_acc_mp
        00_hgz_emsc_20161030_0000029_dis_mp
        00_hgz_emsc_20161030_0000029_vel_mp

In [8]: print(ds.waveforms.IT_CLO['00_hge_emsc_20161030_0000029_acc_cv'])
1 Trace(s) in Stream:
IT.CLO.00.HGE | 2016-10-30T06:40:13.769000Z - 2016-10-30T06:43:13.764000Z | 200.0 Hz, 36000 samples

In [9]: print(ds.waveforms.IT_CLO['00_hge_emsc_20161030_0000029_acc_cv'][0])
IT.CLO.00.HGE | 2016-10-30T06:40:13.769000Z - 2016-10-30T06:43:13.764000Z | 200.0 Hz, 36000 samples

In [10]: print(ds.waveforms.IT_CLO['00_hge_emsc_20161030_0000029_acc_cv'][0].data)
[-3.83038211 -3.82375908 -3.82659698 ...,  2.51704001  3.18688297
  3.62445688]
```

Figure 11c Example of reading waveforms and related metadata from an ITACA 3.1 ASDF data set.

All the meta information related to waveforms are hierarchically stored in the *Auxiliary Data* section (*Headers*) together with 5% damped acceleration and displacement response spectra (*Spectra*). Both *Headers* and *Spectra* are identified within the two respective data groups by the correspondent *waveform tags* (Figure 11d).

```
In [39]: print ds.auxiliary_data.Spectra.IT_CLO['00_hge_emsc_20161030_0000029_dis_mp']
Auxiliary Data of Type 'Spectra'
    Path: 'IT_CLO/00_hge_emsc_20161030_0000029_dis_mp'
    Data shape: '(2, 105)', dtype: 'float32'
    Parameters:
        damping: 0.05
        pgd_cm: -14.128563

In [40]: print ds.auxiliary_data
Data set contains the following auxiliary data types:
    Headers (1 item(s))
    Spectra (1 item(s))

In [41]: print ds.auxiliary_data.Headers.IT_CLO
12 auxiliary data item(s) of type 'Headers/IT_CLO' available:
    00_hge_emsc_20161030_0000029_acc_cv
    00_hge_emsc_20161030_0000029_acc_mp
    00_hge_emsc_20161030_0000029_dis_mp
    00_hge_emsc_20161030_0000029_vel_mp
    00_hgn_emsc_20161030_0000029_acc_cv
    00_hgn_emsc_20161030_0000029_acc_mp
    00_hgn_emsc_20161030_0000029_dis_mp
    00_hgn_emsc_20161030_0000029_vel_mp
    00_hgz_emsc_20161030_0000029_acc_cv
    00_hgz_emsc_20161030_0000029_acc_mp
    00_hgz_emsc_20161030_0000029_dis_mp
    00_hgz_emsc_20161030_0000029_vel_mp

In [42]: print ds.auxiliary_data.Spectra.IT_CLO
6 auxiliary data item(s) of type 'Spectra/IT_CLO' available:
    00_hge_emsc_20161030_0000029_acc_mp
    00_hge_emsc_20161030_0000029_dis_mp
    00_hgn_emsc_20161030_0000029_acc_mp
    00_hgn_emsc_20161030_0000029_dis_mp
    00_hgz_emsc_20161030_0000029_acc_mp
    00_hgz_emsc_20161030_0000029_dis_mp

In [43]: print ds.auxiliary_data.Spectra.IT_CLO['00_hge_emsc_20161030_0000029_dis_mp']
Auxiliary Data of Type 'Spectra'
    Path: 'IT_CLO/00_hge_emsc_20161030_0000029_dis_mp'
    Data shape: '(2, 105)', dtype: 'float32'
    Parameters:
        damping: 0.05
        pgd_cm: -14.128563
```

Figure 11d Example of reading Auxiliary Data from an ITACA 3.1 ASDF data set. The Auxiliary Data consists on meta information about events, stations, waveforms, data processing, credits, and data license, as well as 5% acceleration and displacement response spectra.

2.2 Events

The *Event Search* page manages information about the seismic events related to the waveforms available from ITACA (Figure 12a). The tab *quick search* allows selecting a list of events corresponding to the alphanumeric code that identify the earthquake in the database (*ITACA ID*), the name assigned to the event (*Event name*), the occurrence time period (*Date Time*), and the number of records. The advanced search can be made by using other three tabs that includes all the key-fields useful to query ITACA events related to *Geophysical or Geographic metadata*, and *Events per station* (Table 5). The query outcomes are also provided as epicenters distribution over OpenStreetMap or OpenTopoMap layer and can be downloadable as KML format (Figure 12a).
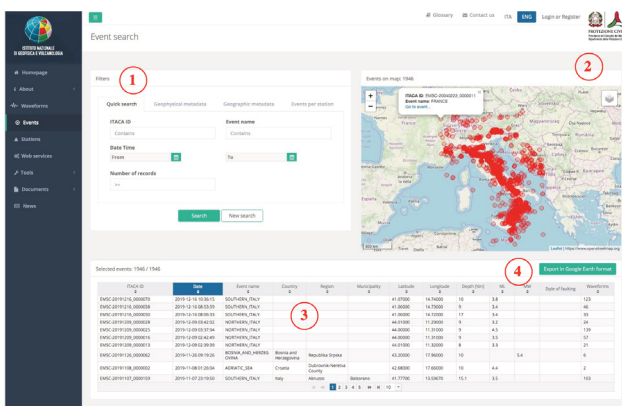
Figure 12a Event search in ITACA 3.1: 1) tabs to query earthquakes metadata to the ITACA database by means of quick or advanced search; 2) map of the epicenters represented on an OpenStreetMap inset; 3) embedded table with relevant metadata and parameters of the selected records; 4) button to export the epicenters distribution in KML format.

	Key-fields	Description	Example
Quick search	ITACA ID	Alphanumeric code to identify the seismic event in the ITACA database	EMSC-20161030_0000029
	Event name	Generic Event name as provided by the ITACA operators	CENTRAL_ITALY
	Date Time	Origin time of the seismic event	From: 2016-08-24 To: 2019-01-01
	Number of records	Minimum number or records	> = 100
Geophysical metadata	Hypocentral depth [km]	The distance between the hypocenter and the epicenter of the earthquake	From: 0 To: 30
	Magnitude	Measures the intensity of the seismic event (ML, Mw)	From: 5.5 To: 6.5
	Style of faulting	Fault mechanism during an earthquake	Normal faulting

	Latitude	Latitude of the epicenter	From: 42.0 To: 43.0
	Longitude	Longitude of the epicenter	From: 12.0 To: 15.0
	Country	To select only the recording stations within the country border	Italy
	Region	To select only the recording stations within the region border	Umbria
	Province	To select only the recording stations within the province border	Perugia
	Municipality	To select only the recording stations within the municipality border	Norcia
Events per station	Network	The code associated to the recording Seismic Network according to the International Federation of Seismograph Network (FDSN)	IT – Italian Strong Motion Network (RAN)
	Station code	3 to 5 characters that identify the recording instrument and its physical location	CLO

Table 5 List of the *Event Search Key-fields.*

By clicking on a specific record of the events list or on an epicenter (Figure 12a), the Users can access the *Event detail* page where all the available meta information are organized in six thematic sections. The section *Origin* shows all the available information about localization and magnitudes associated to the selected event, together with the distribution of the recording stations around the seismic source. Each station can be explored by clicking on the map. A collection of event data coming from web services when available of different scientific institutions (e.g. EMSC, https://www.seismicportal.eu/fdsn-wsevent.html; INGV, http://cnt.rm.ingv.it/ webservices_and_software) is also provided in QuakeML format through the specific link. The section *Moment tensor* shows the solutions of the seismic moment tensor, including the beach ball representation. The section *Source* provides information about source geometry, nucleation point, and focal mechanism. The tab *Record* allows to download all the accelerometric waveforms of the selected event and to display PGA and PGV recorded by each station, together with the epicenter and the seismogenic fault of the earthquake (Figure 12b). The section *Shakemap* (Figure 12c) shows the shaking maps in terms of macroseismic intensity (MCS), PGA, PGV, and Pseudo Spectral Acceleration (PSA) for T=0.3, 1, and 3s. The Shakemaps are reviewed and published by the INGV (http://shakemap.rm.ingv.it/shake/). Auxiliary information can be retrieved from the tab *Links*.

2.3 Stations

Information about recording stations can be accessed through the *Station* page according to several key fields related to network and station code, localization, and site characterization (Figure 13a). The tab *quick search* allows selecting a list of stations corresponding to the seismic network (*Network*), the name assigned to the station (*Station name*), and the minimum number of records. The advanced search can be made by using the other two tabs related to *Geographic metadata* and *Site metadata*. The query outcomes are provided in a list of selected stations with basic information about localization and site parameters, such as Vs30 and EC8 soil classification.
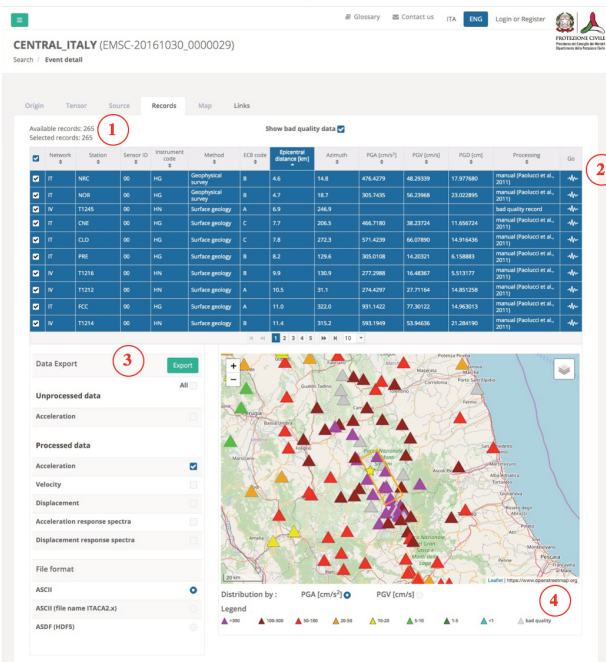
Figure 12b Example of *Event detail* about the 2016, Mw6.5 Norcia earthquake recorded by 265 stations in ITACA 3.1: 1) list of accelerometric records sorted by the epicentral distance; 2) link to the *Waveform detail* page of the selected record; 3) data download of the records selected in the table above; 4) distribution of PGA or PGV around the seismogenic source represented on an OpenStreetMap inset.
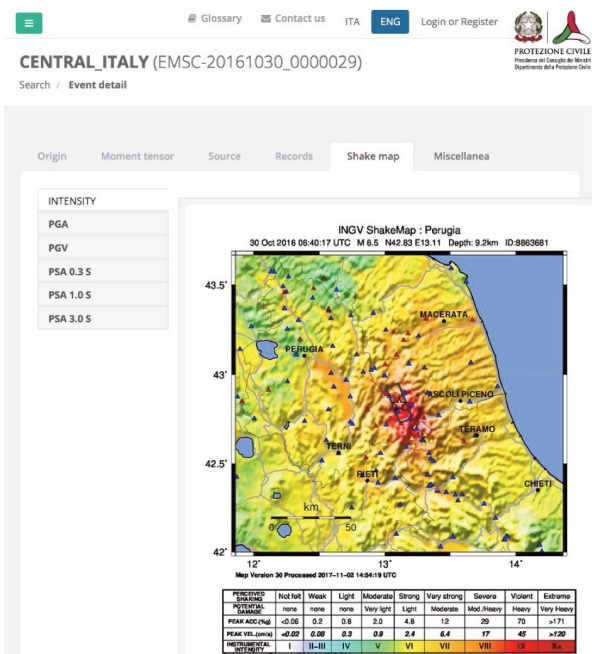

Figure 12c Example of *Shakemap* in terms of macroseismic intensity (MCS) for the 2016, Mw6.5 Norcia earthquake. The shakemaps are also available for PGA, PGV, and PSA at 0.3, 1, and 3 s.

The selected stations are also displayed over OpenStreetMap or OpenTopoMap layers and can be downloadable as KML format (Figure 13a). The *Station detail* pages of each station can be accessed by clicking on the correspondent row in the table or on the map (Figure 13a).

The Figure 13b shows the *Station detail* page of the example for the station AMT, belonging to the network IT. The station is displayed on an OpenStreetMap, and basic information about localization, housing, type of installation, and site characterization parameters related to the seismic sensors are provided in the tab *Info*. The site parameters displayed in the sensors table, such as Vs30, EC8 soil classification, are assigned on the base of expert judgement among all those available in the database (Figure 13b). The Vs30 in the ITACA database is estimated for a single station on the basis of S-wave velocity profile inferred by geophysical methods, and correlation with topographic slope [Wald and Allen, 2007]. For each estimate the EC08

soil category is assigned, also on the base of surface geology. When the S-wave velocity profile is available, the soil category of a recording site is assigned according to the current Italian seismic code [NTC 2018]. To retrieve all the station meta information stored in ITACA, the *User* can browse the Station detail page through a number of thematic tabs about *Topography, Geothematic maps, Geotechnical and Geophysical data, Seismological analysis,* accelerometric *Records, Photo gallery,* and supplementary *Documentation.*

Figure 13a Station search in ITACA 3.1: 1) tabs to query recording stations to the ITACA database by means of quick or advanced search; 2) map of the selected stations represented on an OpenStreetMap inset; 3) table with relevant metadata and parameters of the selected records; 4) button to export the stations map in KML format. As an example, we have selected only the stations belonging to the network IT. The Station detail pages can be accessed by clicking on each row of the table or on blue triangles on the map.
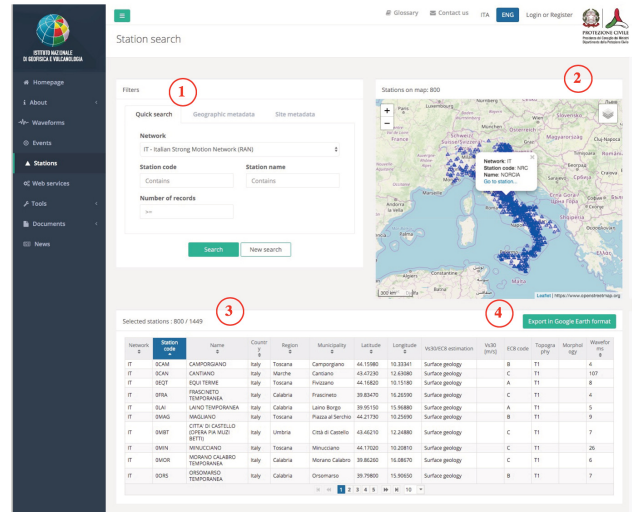
Figure 13b Example of Station detail of the station Amatrice (IT.AMT) in ITACA 3.1: 1) tabs to access station metadata about topography, geological maps, geotechnical and geophysical data, seismological analysis, accelerometric waveforms, photo gallery, and supplementary documents; 2) basic information about the station; 3) table with metadata and site characterization parameters related to the recording instrument identified by the Sensor ID; 4) selected station (blue triangle) and surrounding ones (grey triangles) represented on an OpenStreetMap inset.
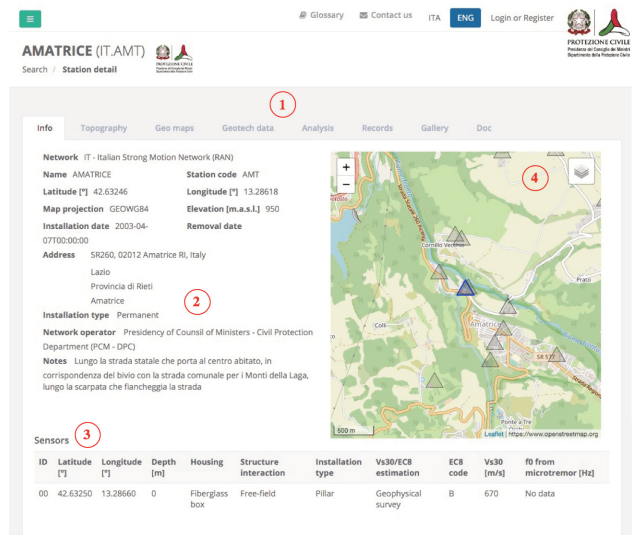
The section *Topography* shows the station on an OpenTopoMap layer together with the available information about morphology and topographic classification according to the NTC [2018] (Figure 13c). The Vs30 is estimated from the topographic slope [Wald and Allen, 2007]. The section *Geothematic map* collects all the site information inferred by geological maps. Since different maps can be linked to a specific recording instrument, only the most reliable is marked by a green star (Figure 13d). A summary map in the section *Geotechnical-Geophysical* shows the location of all the geotechnical and/or geophysical surveys around the station together with geographic coordinates, type of measurement, and references (Figure 14a). The outcomes of the surveys can be expanded or collapsed to reveals seismic waves velocity profiles (Figure 14b)

or microtremor analysis (Figure 14c). The *Seismological analysis* section contains Horizontal-to-Vertical spectral ratio of 5% damped acceleration response spectra (HVRS), Horizontal-to-Vertical spectral ratio of S-phase Fourier spectra (S-phase HVSR), and Horizontal-to-Vertical spectral ratio of coda Fourier spectra (coda HVSR). The seismological analysis can be accessed by clicking on the correspondent label (Figure 15a,b). All the accelerometric waveforms recorded by each station can be selected and downloaded in the *Records* section of the Station detail page. Photos of the housing of the recording stations, instrumentation, and surrounding landscape are shown in the tab *Photo gallery*. All the metadata related to each station are also available in the form of a report, accessed via the tab *Document*. Station reports are generated only for stations having site information, such as geological maps, geophysical measurements or site parameters inferred from seismological analysis.
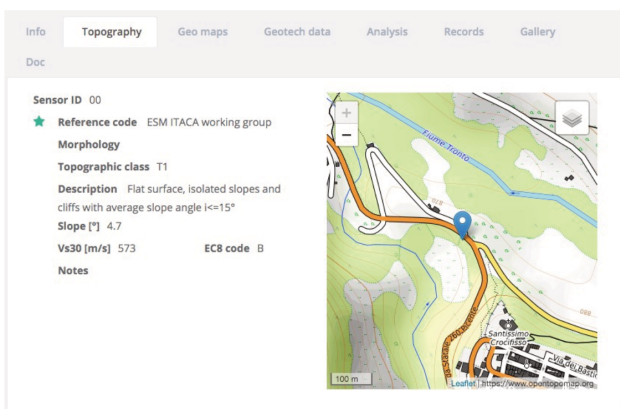
Figure 13c Example of the Topography tabs for the station Amatrice (IT.AMT) in ITACA 3.1.
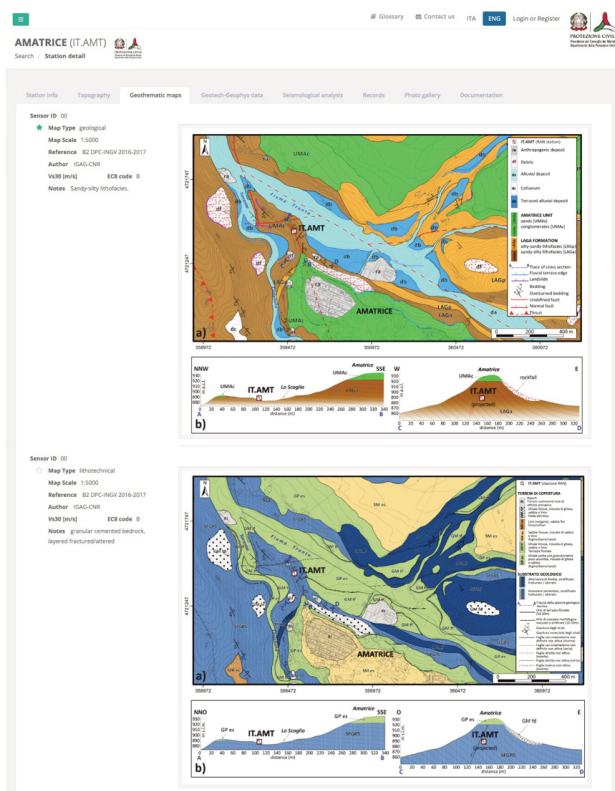
Figure 13d Example of the *Geothematic* tab for the station Amatrice (IT.AMT) in ITACA 3.1. Each geological map is associated to a recording instrument by the Sensor ID. According to an expert judgement, the most reliable geological map is marked by a green star.

Figure 14a Example of the *Geotechnical-Geophysical* tab for the station ROMA (SEDE INGV) (IV.ROM9) in ITACA 3.1. A summary map shows the location of geotechnical and geophysical surveys available in the database. The outcomes of the surveys can be expanded or collapsed to reveals seismic waves velocity profiles and/or microtremor analysis.
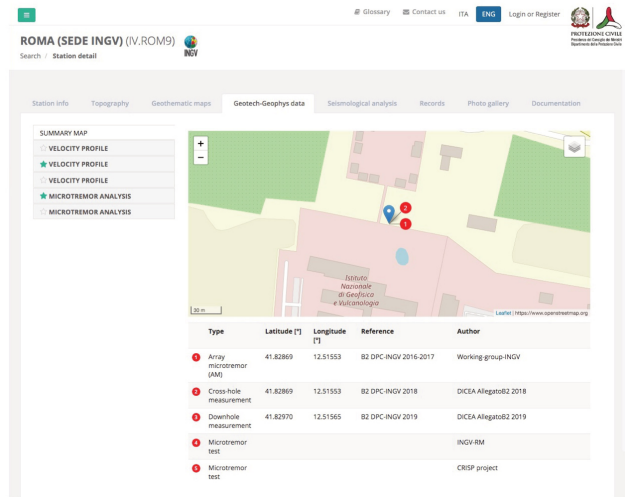

Figure 14b Example of P-wave and S-wave velocity profile for the station ROMA (SEDE INGV) (IV.ROM9) in ITACA 3.1. Metadata such as the Vs30 values, the EC8 and the NTC18 soil classification, the depth beyond which the S-wave velocity exceeds 800 m/s (H800), and the average velocity of the S-waves up to H800 are reported in the header of the velocity profile.
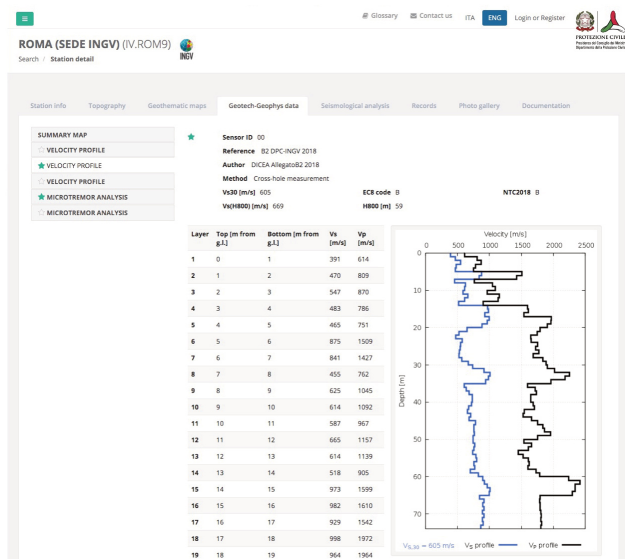
Figure 14c Example of microtremor analysis for the station ROMA (SEDE INGV) (IV.ROM9) in ITACA 3.1. are reported in the header of the velocity profile. The value of the fundamental resonance frequency (f0) and the H/V curve classification are reported in the header of the figure.
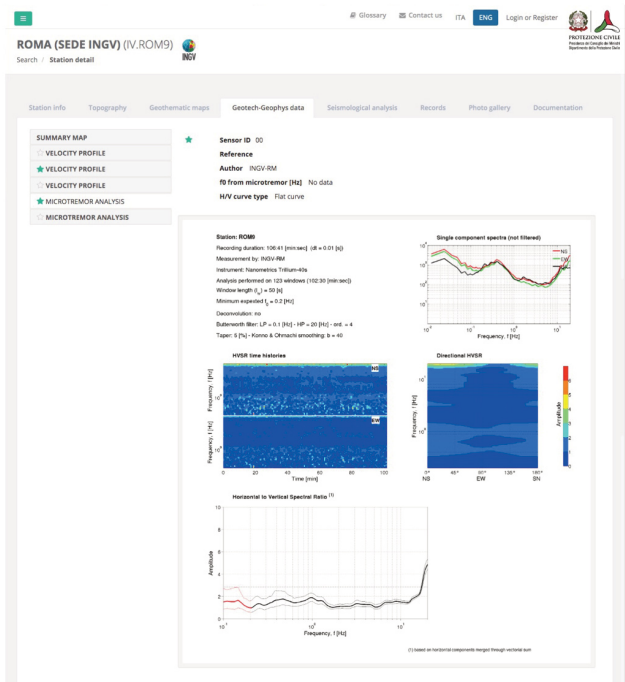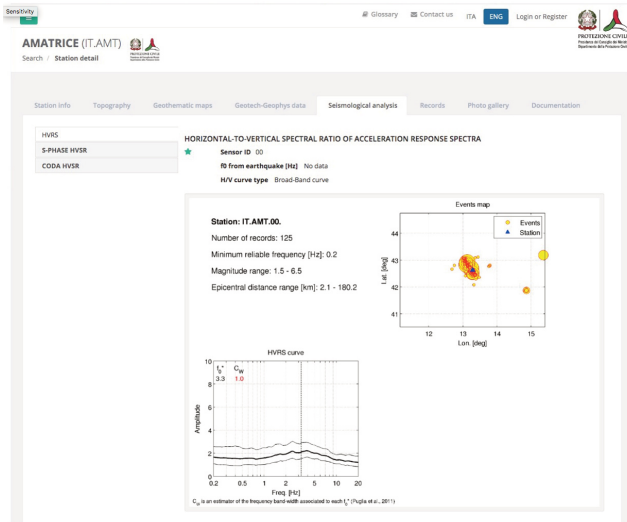
Figure 15a Example of *Seismological analysis* for the station AMATRICE (IT.AMT) in ITACA 3.1 in terms of Horizontal to Vertical Spectral Ratio of acceleration response spectra (HVRS).
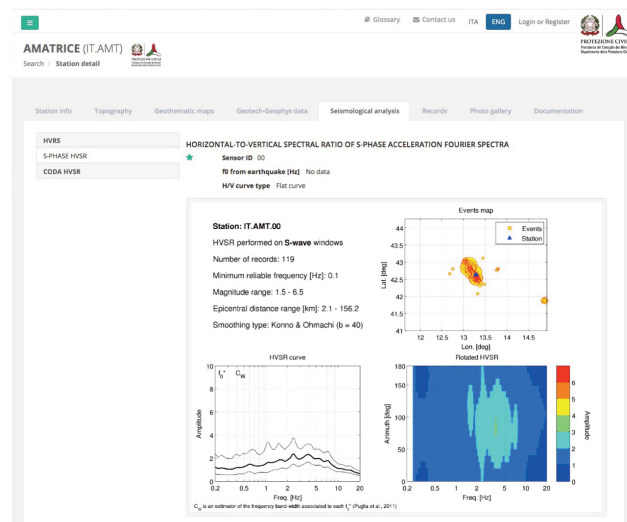

Figure 15b Example of *Seismological analysis* for the station AMATRICE (IT.AMT) in ITACA 3.1 in terms of Horizontal to Vertical Spectral Ratio of S-phase acceleration Fourier spectra (HVSR).
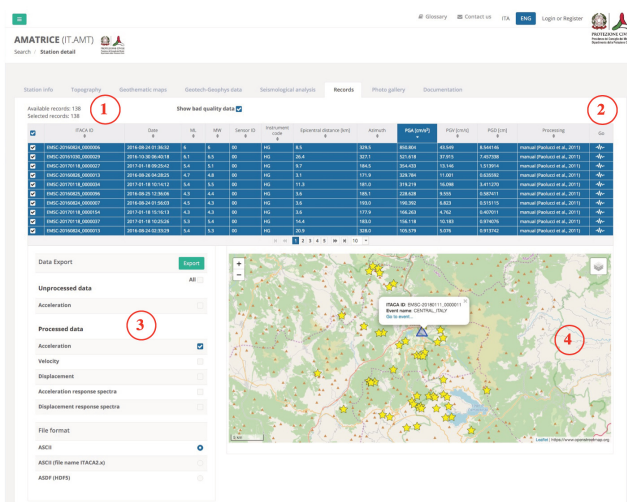

Figure 16 Example of accelerometric *Records* available in ITACA 3.1 for the station AMATRICE (IT.AMT). 1) list of accelerometric records sorted by the PGA; 2) link to the *Waveform detail* page of the selected record; 3) data download of the records selected in the table above; 4) distribution of the selected earthquakes (yellow stars) recorded by the IT.AMT station (blue triangle) represented on an OpenStreetMap inset.

2.4 Web services

ITACA provides machine-friendly access to data and metadata included in the database through the Web Services (WSs) reported in Table 6. The WSs available in ITACA 3.1 (http://itaca.mi.ingv.it/ItacaNet_31/#/data_and_services/web_services) are essentially dedicated to the download of waveforms and the selection of spectrum compatible records. For further details about the selection and the scaling of spectrum compatible waveforms see the section *Tools* of this technical report.

Web Service	Summary	Link
Event-Data	*Waveforms and response spectra download (ASCII or ASDF format)*	http://itaca.mi.ingv.it/itaca31ws/eventdata/1/
REXEL-target-spectrum	*Computation of design acceleration 5%-damping elastic spectra*	http://itaca.mi.ingv.it/itaca31ws/rexel-target-spectrum/1/
REXEL-spectra-selection	*Access to response spectra included in the ITACA 3.1 Database*	http://itaca.mi.ingv.it/itaca31ws/rexel-spectra-selection/1/
REXEL-run-matching	*Search of waveforms combinations in ITACA 3.1 that on average match a target spectrum within a specified tolerance*	http://itaca.mi.ingv.it/itaca31ws/rexel-run-matching/1/
REXEL-outputs	*After running REXEL-run-matching, returns available-combinations and a summary of the user-options, as well as waveforms and metadata-table for a selected combination*	http://itaca.mi.ingv.it/itaca31ws/rexel-outputs/1/
REXEL-plot	*Generates three types of plot in PNG format: target-spectra, spectra-selection and spectra-matching*	http://itaca.mi.ingv.it/itaca31ws/rexel-plot/1/

Table 6 Web Services to access data and metadata in ITACA 3.1.

In order to use the WSs, the User can require an authentication token to ITACA (http://itaca.mi.ingv.it/itaca31ws/generate-signed-message/1/) or EIDA (European Integrated Data Archive, https://geofon.gfz-potsdam.de/eas/). The data can be downloaded from ITACA using a HTTP-POST request contacting the WS address and specifying the request options by using the open source software CURL (https://curl.haxx.se/), together with the description of the option parameters supported by the ITACA Event-Data WS, are shown in Table 7.

Sample requests

- curl -X POST -F "message=@path/to/token.txt" "http://itaca.mi.ingv.it/itaca31ws/eventdata/1/query?eventid=usp000jknf&catalog=USGS&processing-type=CV,MP,AP" -o output.h5
- curl -X POST -F "message=@path/to/token.txt" "http://itaca.mi.ingv.it/itaca31ws/eventdata/1/query?eventid=IT-2012-*&station=NVL,MDN" -o output.h5
- curl -X POST -F "message=@path/to/token.txt" "http://itaca.mi.ingv.it/itaca31ws/eventdata/1/query?eventid=IT-2012-*&station=NVL,MDN&format=ascii" -o output.zip
- curl -X POST -F "message=@path/to/token.txt" "http://itaca.mi.ingv.it/itaca31ws/eventdata/1/query?eventid=IT-2012-0008,IT-2012-0009&station=M??&processing-type=CV,MP&data-type=ACC,DIS,SA,SD&add-xml=True&add-auxiliary-data=True" -o output.h5

Option	Description	Default
eventid	*Select a specific event by ID; IDs are catalog specific, you can specify a catalog using the '—catalog' option. Multiple IDs are comma-separated. It must be a valid ASCII. Allowed wildcard: '*' (matches zero or more characters) or '?' (matches a single character)*	*
catalog	*{ITACA,EMSC,USGS,ISC,INGV} Select a catalog for the specified '—eventid' option; available catalogs are: 'ITACA', 'EMSC', 'USGS', 'ISC' and 'INGV'*	ITACA
network	*Select one or more network codes. Can be SEED network codes or data center defined codes. Multiple codes are comma-separated. It must be a valid ASCII. Allowed wildcard: '*' (matches zero or more characters) or '?' (matches a single character)*	*
station	*Select one or more SEED station codes. Multiple codes are comma-separated. It must be a valid ASCII. Allowed wildcard: '*' (matches zero or more characters) or '?' (matches a single character)*	*
location	*Select one or more SEED location identifiers. Multiple identifiers are comma separated. As a special case '—' (two dashes) will be translated to a string of two space characters to match blank location IDs. It must be a valid ASCII. Allowed wildcard: '*' (matches zero or more characters) or '?' (matches a single character)*	*
channel	*Select one or more SEED channel codes. Multiple codes are comma-separated. It must be a valid ASCII. Allowed wildcard: '*' (matches zero or more characters) or '?' (matches a single character)*	*
format	*{hdf5,ascii} Specify format for result. Choose between an ASDF volume ('hdf5') or a zip file containing DYNA 1.2 ASCII waveforms ('ascii')*	hdf5

processing-type	{CV,MP,AP} Specify processing type for data selection; available types are: 'CV' (data converted to cm/s^2), 'MP' (manually processed using Paolucci et al. [2011]) and 'AP' (automatically processed using Paolucci et al. [2011])	MP
data-type	{ACC,VEL,DIS,SA,SD} Specify data type for data selection; available types are: 'ACC' (acceleration, in cm/s^2), 'VEL' (velocity, in cm/s), 'DIS' (displacement, in cm), 'SA' (acceleration response spectra, in cm/s^2) and 'SD' (displacement response spectra, in cm)	ACC
add-xml	{True,False} Add related QuakeML and FDSN StationXML. The value attributed to this option is ignored when 'format=ascii' is choosen	False
add-auxiliary-data	{True,False} Add related auxiliary data (e.g. processing parameters, site characterization metadata, etc.). The value attributed to this option is ignored when 'format=ascii' is chosen	False

Table 7 Usage examples and request parameters supported by the ITACA Event-Data web service.

2.5 Tools

2.5.1 Processing

Strong motion records can by employed in a variety of application fields such as applied seismology or earthquake engineering. Every specific application needs an adequate reliability level of data that should be free from signal distortion or noise both at high and low frequencies. This would ensure, as much as possible, a proper use of strong-motion records.

Accelerometric waveforms uploaded in the ITACA database are automatically processed using the procedure proposed by Paolucci et al. [2011] and summarized in the following:

- linear detrending of the uncorrected acceleration signal (subtraction of a first-order polynomial);
- application of a cosine taper, at the beginning and end of the signal, with percentage fixed to 5% of the signal length and with the possibility of being modified by the user;
- visual inspection of the Fourier spectrum to select the band-pass frequency range (band-pass frequency may be different for the three components);
- application of a second-order acausal time-domain Butterworth filter to the acceleration time series, with zero-pads added at the beginning and end of the signal before the acausal filter is applied;
- removal of zero-pads from the acceleration trace;
- (begin/end) taper of the acceleration signal, with percentage fixed to 5%;
- computation of the velocity signal and linear detrend;
- (begin/end) taper of the velocity signal, with percentage fixed to 5%;
- computation of displacement signal and linear detrend;
- (begin/end) taper of the displacement signal, with percentage fixed to 5%; and
- recursive differentiation to obtain the velocity and the acceleration time series, respectively.

A cosine taper with percentage of 5% and a low-pass filter with a fixed frequency of 30 Hz are applied to the three components of the ground motions (two horizontal and vertical). The high-pass frequencies depend on magnitude ranges defined on the base of statistical analysis on a representative set of processed waveforms [Puglia et al., 2018].

Automatically processed waveforms are revised from the ITACA operators in order to: i) validate the results of the automatic processing; ii) manually process accelerometric time series fine-tuning some parameters defined by the automatic processing such as time of the first and/or the last sample, percentage of signal tapering, or band-pass corner frequencies; iii) classify noisy records or waveforms containing spurious spikes as *bad quality* data.

Once that the waveforms are manually checked, corrected acceleration, velocity, and displacement time series are stored in the database, together with acceleration and displacement 5% damped response spectra. *Bad quality* data are maintained in the database only in the unprocessed version.

To accomplish the growing demand of reliable earthquake waveforms the new ITACA web site provides a renewed tool for the strong-motion data processing. After login to the ITACA web site (see 2.1.1 paragraph *Authentication system*), the new processing interface is available either by the home page or the side bar at http://itaca.mi.ingv.it/ItacaNet_31/processing/select. The *User* can execute a pre-selection of waveforms on the base of the earthquake identifier (*Event ID*), the time window, the magnitude and the epicentral distance intervals, the network and the station codes. In Figure 17 we show an example of data selection for waveforms recorded by the network IT characterized by magnitude greater than 6.0 and epicentral distance lower than 30 km. The result is a table of 67 records matching the search criteria that summarizes some relevant metadata such as network code, station code, location code, recording instrument, event identifier, origin time of the event, epicentral distance, local magnitude, moment magnitude, and maximum horizontal PGA before and after the data processing. The column *Processing* indicates whether the three components of the ground-motion have been manually processed by specialists (green flag), automatically processed (yellow flag), or manually checked by the database operators and judged as bad quality data (red flag). The *Search Result* table can be sorted by column in ascending or descending order and further filtered. The link *go* in the last column enables the *User* to access time series and spectra.
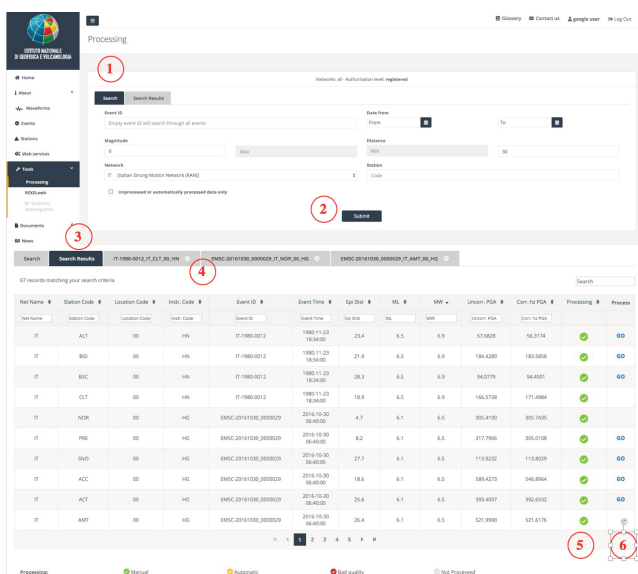

Figure 17 Data selection in the ITACA processing tool: 1) pre-selection on the base of earthquake identifier, data time of the seismic events, magnitude, epicentral distance, recording network, station; 2) button to submit the pre-selection query to the database; 3) list of the selected records with relevant metadata of recording stations, seismic event, and acceleration peak parameters (uncorrected and corrected maximum horizontal PGA); 4) tabs to process the selected records; 5) flags to check whether the accelerometric record has been manually processed by specialist, automatically processed by the standard ITACA procedure, judged as bad quality datum as visually inspection of database operators, or unprocessed; 6) link to the processing tabs.

Figure 18 shows the example of the three components of the ground-motion (two horizontal and vertical) recorded by the station NOR (NORCIA) during the Mw6.5 Norcia earthquake (2016, October 30th). Unprocessed (red) and processed (black) time series are plotted together with the first arrival of P- and S-phase coherently with the time window in which the raw traces have been cut out. The User can switch between meta information dealing with seismic events, stations, and waveforms to check the data to correctly interpret time series and spectra.

Figure 18 Example of data processing: a) acceleration time series; b) velocity time series; c) displacement time series; d) waveforms metadata; e) acceleration response spectra at 5% damping; f) displacement response spectra at 5% damping.
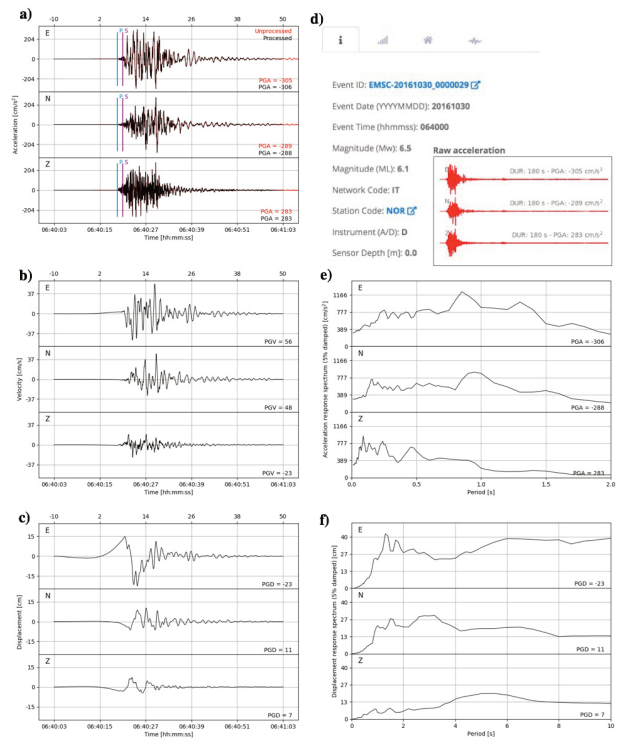
Figure 19 shows the three components of the Fourier amplitude spectra and the band-pass frequencies of the applied filter. Several parameters can be modified to control the data processing, such as value of the band-pass frequencies, filter order, time window to cut the raw waveforms, percentage of cosine taper at the beginning and at the end of the signal, type of trigger (normal or late depending on the pre-event window of the raw trace), sampling interval, and scaling factor. To simultaneously process the three components of the ground-motion, the *User* should press the button *Apply*. After a visually inspection of the processing results in terms of acceleration, velocity, displacement time series or 5% damped acceleration and displacement response spectra, the *User* can download the processed records both in ASCII or HDF5 format.

2.5.2 REXELweb

REXELweb is a tool for the automatic selection and scaling of spectrum-compatible accelerograms for dynamic analysis of structures. It allows to define target spectra according to user-definition or design provisions Eurocode 8 and Italian building code [NTC08], as well as to uniform hazard spectra (UHS) based on a European hazard model (ESHM13-SHARE, http://www.efehr.org/en/home/).
REXELweb is provided in ITACA (http://itaca.mi.ingv.it/ItacaNet_31/#/data_and_services/tools/rexel) as executable application developed in the MATLAB® environment. The MATLAB Component Runtime must be installed together with the executable file *ITACA_REXELweb*.
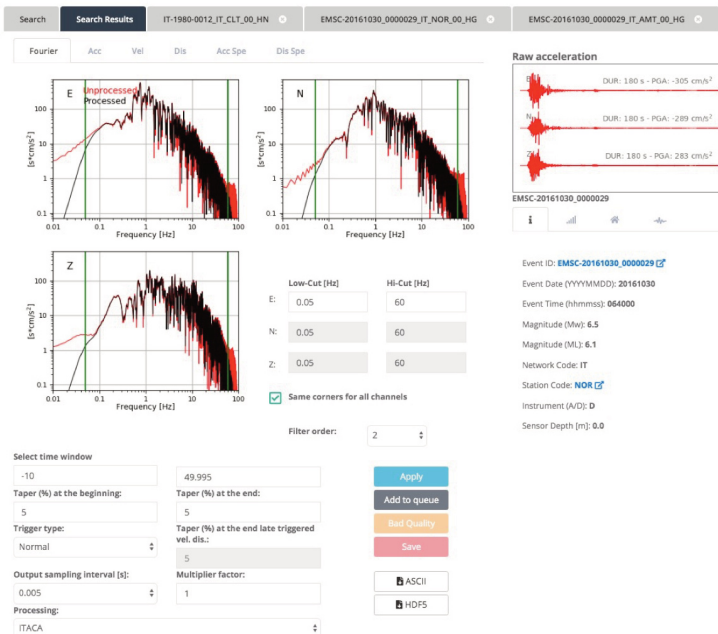
The MATLAB Component Runtime is a free software to run a MATLAB application without a MATLAB license.

Before running the *REXELweb* application, the following requirements must be fulfilled:
- Windows Operating System;
- registration of an account to the ITACA database: http://itaca.mi.ingv.it/;
- curl software (version 7.44.0 or above) must be installed on the system (in order to check for proper installation, you can type on the command/terminal window "curl –help" and the curl help command shall appear).

REXELweb is an utility to run the following web-services (WSs):
- **REXEL-target-spectrum** (http://itaca.mi.ingv.it/itaca31ws/rexel-target-spectrum/1/) allows computation of design acceleration 5%-damping elastic spectra.
- **REXEL-spectra-selection** (http://itaca.mi.ingv.it/itaca31ws/rexel-spectra-selection/1/) provides machine-friendly access to response spectra included in the ITACA Database.
- **REXEL-run-matching** (http://itaca.mi.ingv.it/itaca31ws/rexel-run-matching/1/) searches for sets of waveforms combinations that on average match the target spectrum within a specified tolerance.
- **REXEL-outputs** (http://itaca.mi.ingv.it/itaca31ws/rexel-outputs/1/) generates graphical representation from *REXELweb* outputs file.

REXELweb application can be downloaded after login to the ITACA web site (see 2.1.1 paragraph *Authentication system*). The application starts with a double click on the icon of the ITACA_REXELweb application. At the first step, the application requires to select the input options file *options.txt* compiled with the *User* preferred options (Figure 20).

After the selection of the input file, the application returns the target spectra (horizontal and vertical) (Figure 21). A dialog box *Select One* allows the user to choose *YES* and continue, or to quit the application choosing *NO* by the drop-down menu (Figure 21). In both cases, the results are saved in the directory selected by the User in the file *options.txt*.

If the User chooses *YES*, the application returns a plot of the selected spectra, both as scaled and unscaled (Figure 22). An interactive window *Select One* allows the user to choose *YES* and continue, or to quit the application choosing *NO* by the drop-down menu (Figure 22).

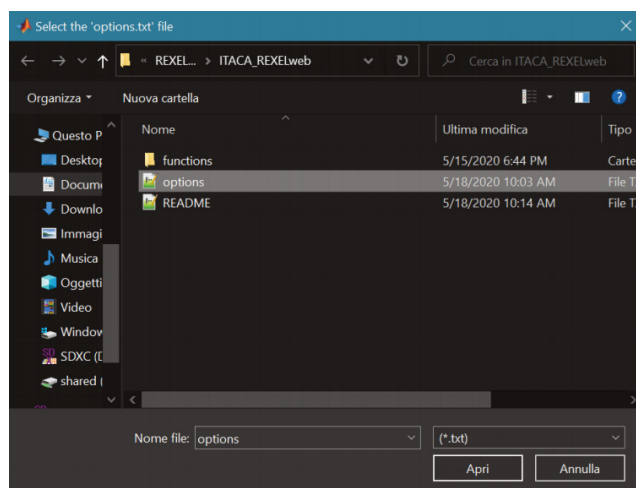**Figure 20** Window to select the user-input options.txt file.
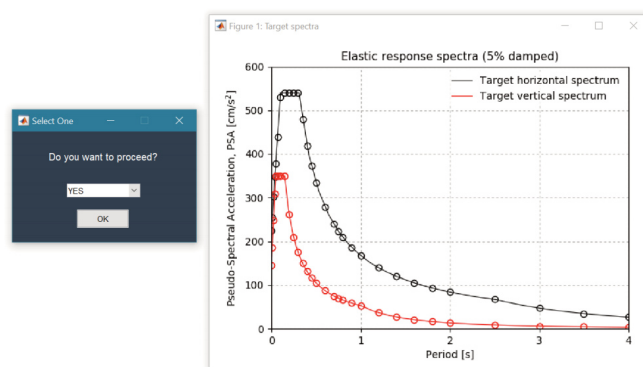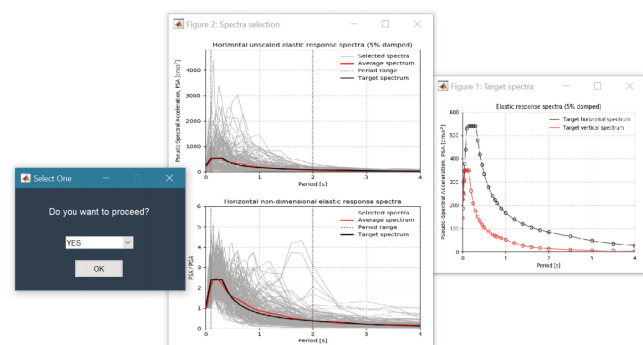

Figure 21 Plot of the target spectra.

Figure 22 Plot of target spectra (right) and selected spectra (left).

If the *User* chooses *YES*, the first combination of spectra is displayed (Figure 23) with a table list of the file names and corresponding metadata. The mean score [Iervolino et al., 2010], which is variable in the range (0-1), indicating the distance between the target and the spectrum of an individual record, is also reported (Figure 24). A dialog box informs the user about the currently displayed combination and the others available (Figure 25a). By clicking on *OK* the *User* can continue the selection of the waveform and choose whether to display another combination among those available (Figure 25b). *QUIT* button allows the *User* to close the application.

After choosing the combination, the *User* can decide whether to exclude one or more waveforms, stations, and/or events. Selecting *NO*, no data is excluded. In any case, if the *User* chooses *YES*, an interactive window allows selecting the records to be excluded by clicking on them (Figure 26 for the waveforms; Figure 27 for the stations, and Figure 28 for the events).
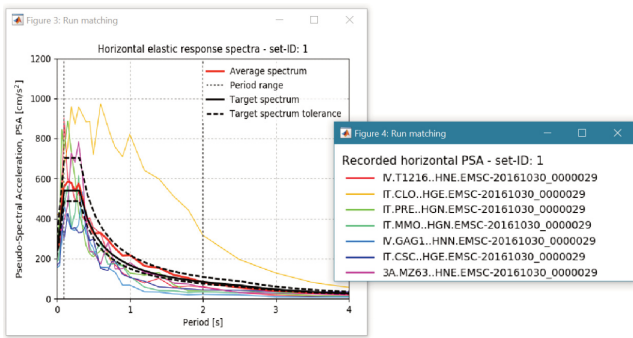
Figure 23 Plot of the first combination of the 7 selected spectra with the legend.

Figure 24 Record names and metadata corresponding to the First combination of the 7 selected spectra reported in Figure 23.
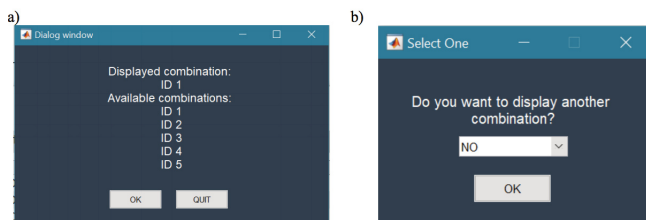

Figure 25 a) Displayed combination and the others available; b) dialog window to display or not another combination among those available.
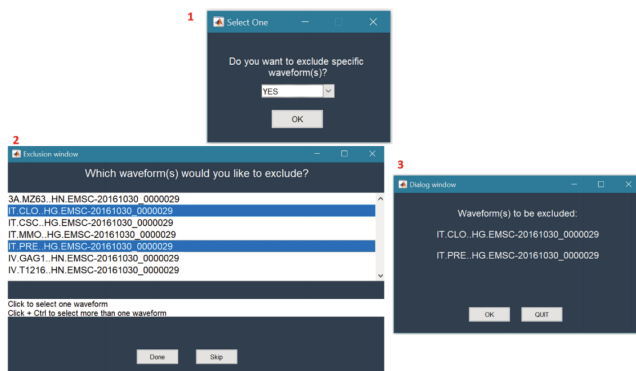

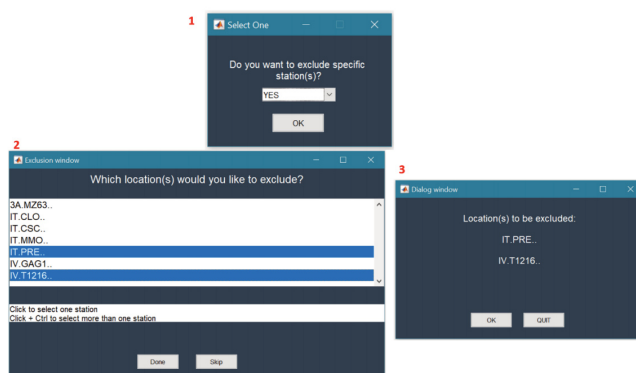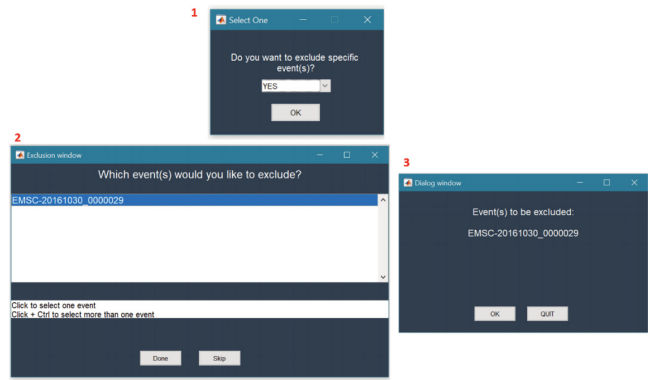**Figure 26** Dialog windows to exclude one or more waveforms.

Figure 27 Dialog windows to exclude one or more stations.

Figure 28 Dialog windows to exclude one or more events.

DONE button saves the data to be excluded and a summary window appears. When some data are excluded, the selection of the possible combinations restarts from the *Spectra Selection*. By clicking on *SKIP*, no data is excluded.

At the end of the selection, the *User* can choose to save or not the waveforms related to the selected combination (Figure 29). If *YES*, 2 different formats are available by the drop-down menu: ASCII or HDF5 (Figure 30). A summary window (Figure 31) resumes the user-options adopted for selection. For a complete list of all the output files see the rexel-output web-service page.

Figure 29 Dialog window to choose if save or not the waveforms related to a specific combination.
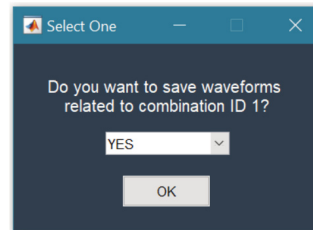

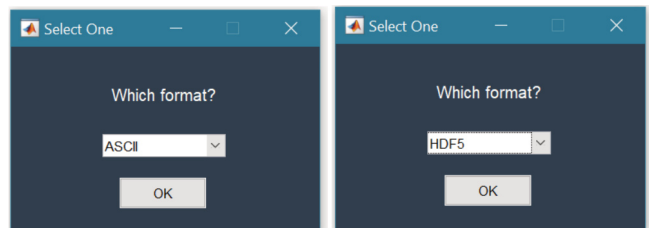**Figure 30** Dialog window to choose the data format.

Figure 31 Summary window.

3. Final considerations

ITACA constitutes a suitable web portal to open access authoritative strong-motion data and tools to manage accelerometric records. The standardization of waveforms and metadata, the data interoperability with open-source software (i.e. ObsPy), and the data availability in formats commonly used in the computational seismology and earthquake engineering make ITACA a flexible instrument to easily explore and download data and metadata of accelerograms recorded in Italy and surrounding areas.

References

D'Amico, M. C., Felicetta, C., Russo, E., Sgobba, S., Lanzano, G., Pacor, F., Luzi, L., (2020). *ITalian ACcelerometric Archive (ITACA), version 3.1 (Version 3.1)*. Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/ITACA.3.1.

Bindi D., Pacor F., Luzi L., Puglia R., Massa M., Ameri G. and Paolucci R., (2011). *Ground motion prediction equations derived from the Italian strong motion database*. Bulletin of Earthquake Engineering, 9(6), 1899-1920, doi:10.1007/s10518-011-9313-z.

CEN, (2004). EuroCode 8: Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings. Bruxelles: European Committee for Standardization.

Luzi L., Sabetta F., Mele F., Hailemikael S., Bindi D., Pacor F., Massa M., Lovati S., Gorini A. and S. P. W. Group, (2009). *The Italian strong motion data base: design, data input and web distribution*. Bollettino di Geofisica Teorica ed Applicata, 50(2), 179-192.

NTC, (2008). *Norme Tecniche per le Costruzioni* - D.M. 14/01/2008, Gazzetta Ufficiale n. 29-4 febbraio 2008. Suppl. Ordinario n. 30. Capitolo 3 Azioni sulle costruzioni (in Italian).

Pacor F., Paolucci R., Ameri G., Massa M. and Puglia R., (2011). *Italian strong motion records in ITACA: Overview and record processing*. Bulletin of Earthquake Engineering, 9(6), 1741-1759. doi:10.1007/s10518-011-9295-x.

Paolucci R., Pacor F., Puglia R., Ameri G., Cauzzi C., Massa M., (2010). *Record Processing in ITACA, the New Italian Strong-Motion Database*. In: Geotechnical, Geological, and Earthquake Engineering. Springer Netherlands, Dordrecht, pp 99-113.

Paolucci R., Pacor F., Puglia R., Ameri G., Cauzzi C., Massa M., (2011). *Record processing in ITACA, the new Italian strong-motion database*. In: Akkar S., Gülkan P., van Eck T. (eds), Chapter 8 of the book earthquake data in engineering seismology - predictive models, data management and networks. ISBN: 978-94-007-0151-9 (printed version) 978-94-007-0152-6 (E-book version). Geotechnical, Geological, and Earthquake Engineering, vol. 14, Springer, Netherlands.

Puglia R., Russo E., Luzi L., D'Amico M., Felicetta C., Pacor F., Lanzano G., (2018). *Strong-motion processing service: a tool to access and analyse earthquakes strong-motion waveforms*. Bulletin of Earthquake Engineering, https://doi.org/10.1007/s10518-017-0299-z.

Shahi S.K. and Baker J.W., (2014) *An efficient algorithm to identify strong velocity pulses in multi-component ground motions*. Bulletin of the Seismological Society of America, 104(5), 2456-2466.

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

Krischer L., Smith J., Lei W., Lefebvre M., Ruan Y., Sales de Andrade E., Podhorszki N., Bozdağ E., Tromp J., (2016). *An Adaptable Seismic Data Format, Geophysical Journal International*, Volume 207(2), 1003–1011, https://doi.org/10.1093/gji/ggw319.

The HDF Group, (1997 - 2015). *Hierarchical Data Format, version 5*. Available at: https://www.hdfgroup.org/HDF5/, last accessed 7 October 2015.

Schorlemmer D., Wyss A., Maraini S., Wiemer S. and Baer M., (2004). *Orfeus Newsletter 6(2): QuakeML - An XML schema for seismology*. Available at: http://www.orfeus-eu.org/organization/Organization/Newsletter/vol6no2/quakeml.shtml, last accessed 7 October 2015.

Schorlemmer D., Euchner F., Kastli P., Saul J. and Group Q.W., (2011). *QuakeML: status of the XML-based seismological data exchange format*, Ann. Geophys., 54(1), 59–65.

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

QUADERNI di GEOFISICA

I QUADERNI DI GEOFISICA (QUAD. GEOFIS.) accolgono lavori, sia in italiano che in inglese, che diano particolare risalto alla pubblicazione di dati, misure, osservazioni e loro elaborazioni anche preliminari che necessitano di rapida diffusione nella comunità scientifica nazionale ed internazionale. Per questo scopo la pubblicazione on-line è particolarmente utile e fornisce accesso immediato a tutti i possibili utenti. Un Editorial Board multidisciplinare ed un accurato processo di peer-review garantiscono i requisiti di qualità per la pubblicazione dei contributi. I QUADERNI DI GEOFISICA sono presenti in "Emerging Sources Citation Index" di Clarivate Analytics, e in "Open Access Journals" di Scopus.

QUADERNI DI GEOFISICA (QUAD. GEOFIS.) welcome contributions, in Italian and/or in English, with special emphasis on preliminary elaborations of data, measures, and observations that need rapid and widespread diffusion in the scientific community. The on-line publication is particularly useful for this purpose, and a multidisciplinary Editorial Board with an accurate peer-review process provides the quality standard for the publication of the manuscripts. QUADERNI DI GEOFISICA are present in "Emerging Sources Citation Index" of Clarivate Analytics, and in "Open Access Journals" of Scopus.

RAPPORTI TECNICI INGV

I RAPPORTI TECNICI INGV (RAPP. TEC. INGV) pubblicano contributi, sia in italiano che in inglese, di tipo tecnologico come manuali, software, applicazioni ed innovazioni di strumentazioni, tecniche di raccolta dati di rilevante interesse tecnico-scientifico. I RAPPORTI TECNICI INGV sono pubblicati esclusivamente on-line per garantire agli autori rapidità di diffusione e agli utenti accesso immediato ai dati pubblicati. Un Editorial Board multidisciplinare ed un accurato processo di peer-review garantiscono i requisiti di qualità per la pubblicazione dei contributi.

RAPPORTI TECNICI INGV (RAPP. TEC. INGV) publish technological contributions (in Italian and/or in English) such as manuals, software, applications and implementations of instruments, and techniques of data collection. RAPPORTI TECNICI INGV are published online to guarantee celerity of diffusion and a prompt access to published data. A multidisciplinary Editorial Board and an accurate peer-review process provide the quality standard for the publication of the contributions.

MISCELLANEA INGV

MISCELLANEA INGV (MISC. INGV) favorisce la pubblicazione di contributi scientifici riguardanti le attività svolte dall'INGV. In particolare, MISCELLANEA INGV raccoglie reports di progetti scientifici, proceedings di convegni, manuali, monografie di rilevante interesse, raccolte di articoli, ecc. La pubblicazione è esclusivamente on-line, completamente gratuita e garantisce tempi rapidi e grande diffusione sul web. L'Editorial Board INGV, grazie al suo carattere multidisciplinare, assicura i requisiti di qualità per la pubblicazione dei contributi sottomessi.

MISCELLANEA INGV (MISC. INGV) favours the publication of scientific contributions regarding the main activities carried out at INGV. In particular, MISCELLANEA INGV gathers reports of scientific projects, proceedings of meetings, manuals, relevant monographs, collections of articles etc. The journal is published online to guarantee celerity of diffusion on the internet. A multidisciplinary Editorial Board and an accurate peer-review process provide the quality standard for the publication of the contributions.