Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2019-113-RC2, 2019

Earth System Science Data Discussions

Interactive comment on "The INSIEME seismic network: a research infrastructure for studying induced seismicity in the High AgriValley (southern Italy)" *by* Tony Alfredo Stabile et al.

Anonymous Referee #2

Received and published: 13 December 2019

General comments

Within this manuscript the authors describe a seismic network deployed in a region prone to induced seismicity, tailor made to better understand this process. Before describing the actual seismological infrastructure they provide an extensive introduction to the anthropogenic seismicity, to conclude with a short discussion about both scientific findings and a summary of the peculiarities of the collected dataset. Although being well organized the paper is still unbalanced towards scientific results rather than emphasizing the potential of this dataset for other users. The collected dataset made openly available to the community using standard formats and services has great value

and potential for the community to better understand the generation of induced seismicity and test alternative methods. The dataset should be the core of this manuscript without too many distractions for the readers about the own scientific findings of the authors. Considering the journal target and the high quality of the dataset described here I would suggest (in the detailed comments) a number of changes aiming at reducing the parts about the scientific findings while enhancing the presentation of this peculiar dataset.

Detailed comments

Page 1, line 27 "... Data collected until the end of the INSIEME project (2019-03-23) are already released ..." Indeed at IRIS DMC there are data to 25.06.2019. Check and correct if needed. What about real-time data? See also later comments about the Data availability section.

Page 1, line 29 "...available from (https://doi.org/10.7914/SN/3F_2016; Stabile et al., 2016)." Replace with "...available from IRIS DMC." Details about how to retrieve and use data should be provided with all details in the Data availability section.

Page 1 to page 2, line 9 Remove/reformulate the introduction aiming at keeping only three short paragraphs (∼5 lines each) about: The project (Funder), the importance of high quality seismic networks to better understand induced seismicity and a short summary of the paper content (the actual lines 10-16 at page 2 can stay).

Figure 1 Change the color used for the INSIEME station to improve the visibility (currently with the dark blue is difficult to spot the triangles on the map).

Page 3, line 18 to page 4, line 5 This section, which is important to understand the context, could be included in the reshaped introduction.

Page 5, line 13 "They provide a flat response to ground velocity up to 100 Hz." Redundant, the higher limit is written already in the previous sentence.

Page 7, line 10-11 "Dataless of all the INSIEME seismic stations, which include com-

prehensive information of each station and the respective instrument response, are provided in the Supplement." This is not needed since metadata are provided in standard stationXML or text format. The authors can provide additional details (including URLs to station, dataselect and availability web services) in the data availability section.

Page 7, lines 15-20 The authors are providing here a long explanation about how to reach the remote with dynamic IPs. Do they try to use VPN? OpenVPN is supported by the hardware in use and they all connect to the same server. Add a sentence why they used this approach rather than creating a Virtual Private Network.

Page 7, line 23 "The router Teltonika RUT-500 is compliant with SeedLink, and therefore adopted as transmit tool" What's the meaning of this sentence? Probably that this hardware supports TCP/IP protocol?

Page 7, line 25 Probably the same can be achieved with the ping reboot functionality without the need to force daily reboots. Not sure though this version of router has this functionality.

Page 7, line 30 to page 8, line 7 Add a reference here to the data availability section where I suggest to add a figure with the data availability (%) for all stations for the entire period of operation of the network (e.g. using obspy-scan).

Page 8, line 27 Data quality Section and related figures 5 and 6 I suggest to add figures with Probability Density Functions for all stations/components and accordingly comment them in this section. The PDFs should be calculated over the entire period of operation of the network which according to the data available at IRIS DMC is 01.04.2016 - 25.06.2019. Current figures are only including 4 days of data which is not enough to have an idea about the actual data quality and argue about quality at different depths/locations. To facilitate the visualization of the figures median, 5th and 95th percentile should be plotted on each panel. To show the difference between the surface, shallow and deep installation (INS1 at 50 m) would be enough to have an ad-

ditional figure with the comparison of the median values for sensors at different depths and locations.

Page 12, line 3 "...INS5 seismic station, where a small peak..." replace small with relevant. Indeed in figure 8 the H/V ratio exceeds 3, making this station the most amplified in the frequency band 2-5 Hz. Would have been interesting though to compute also spectral ratios among the stations having fixed one station as reference (e.g. INS4 or one station nearby free of site effects belonging the other permanent networks).

Page 14, line 1 Data availability section: this should evolve in a comprehensive description of the dataset availability. Provide details about where and how to access the data (fdsn web services at IRIS DMC). Data are available at IRIS DMC from 01.04.2016 to 25.06.2019. Is just this the open dataset described in this paper or this includes also open real-time data or periodic releases after certain embargo dates? Please specify in this section. A figure with the availability for the given period should be added. Would also be ideal to clearly state here if there is a license applied to the data and accordingly ask also IRIS DMC to include this in the DOI metadata of the network (including additional metadata as Funder, Sponsors, ORCIDs of the creators etc.). Moreover within the paper a seismic catalogue is mentioned and would be ideal to provide a link to it from here, either to an fdsn-event service or simply add catalogue to the supplementary material.

Page 14, lines 15 to page 1, line 3 Check this part carefully as most of this is redundant from the previous sections. Try to reduce redundancy and emphasize the part starting at page 15, line 4 stressing the peculiarities of this dataset.

Page 14, line 26 "... detected 856 local natural and induced earthquakes ..." Can the earthquake catalogue, obtained from this network, be added to the supplementary material? Alternatively can the authors point to a repository where this catalogue is hosted?

Supplementary material

Being the dataset archived in a FDSN data centre providing standard data and metadata formats datalessSEED volumes can be omitted in my opinion. Within the manuscript the authors are referring to an own earthquake catalogue. This would be indeed a useful addition for the supplementary material. Or at least a link to an open standard fdsn web service where this can be obtained.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-113, 2019.