ANNALS OF GEOPHYSICS, 68, 5, DM571, 2025; doi:10.4401/ag-9328
OPEN ACCESS

Methods for data and metadata quality tests of large dense seismic networks – focus on AdriaArray

Petr Kolínský*[,1], Johannes Stampa[2,3], Luděk Vecsey[1], Felix Eckel[2], Tena Belinić Topić[4], Thomas Meier[2], the AdriaArray Seismology Group[5]

(1) Institute of Geophysics of the Czech Academy of Sciences, Prague, Czech Republic
(2) Institute for Geosciences, University of Kiel, Germany
(3) Department of Earth Sciences, Bullard Laboratories, University of Cambridge, Cambridge, United Kingdom
(4) Department of Geophysics, Faculty of Science, University of Zagreb, Croatia
(5) https://orfeus.readthedocs.io/en/latest/adria_array_main.html

Article history: received March 6, 2025; accepted October 28, 2025

Abstract

Data quality checks are essential for any broad-band seismological network, and in particular for data of temporary passive seismic experiments. These data quality checks concern (i) the availability and retrievability of the data from public data archives, (ii) the noise conditions at the stations, (iii) formal properties and the correctness of metadata, (iv) the mutual consistency between data and metadata, and finally (v) the quality of the data itself. Methods for these checks are introduced and applied to the AdriaArray Seismic Network. We present techniques for evaluating the quality of individual stations as well as techniques that allow us to detect outlying amplitudes and arrival times in case of a dense network. Results of the tests are summarized in the form of maps and in addition, details are given in an online repository. Our checks are continuously repeated and results are updated to secure high data quality. The aim of our study is to provide the user with useful information on the quality of AdriaArray data, as well as with suggestions for their own data quality assessment. In addition, the presented data quality checks form the basis for data curation by station and network operators. The suggested approaches can also be applied to other large dense seismic networks.

Keywords: Large Dense Seismic Network; Data Quality; Metadata Quality; AdriaArray; Passive Experiments

1. Introduction

Controlling the quality of seismological data has been an essential task since the early days of seismology. For example, noise in seismological recordings has already been analyzed by Lynch (1938). He measured the effects of car traffic, buildings, passing trains, man-made noise during working hours, effects of quarry blasts and of other sources. To distinguish the “disturbing” noise from the natural Earth ambient wavefield, attempts to define the “normal” noise conditions have been made over the years. Although nowadays considered as a classical standard, Peterson (1993) introduced his new high-noise model (NHNM) and new low-noise model (NLNM) as an update of the HNM and LNM (or OHNM and OLNM – "original" high/low noise models), which had been in use years before (see, e.g., Brune and Oliver, 1959). The probability power spectral density (PPSD), or probability density functions (PDFs), are now standard tools used to identify the noise conditions and to examine the overall station quality, following mainly the works by McNamara and Buland (2004) and McNamara and Boaz (2005). Usually, measured PPSDs are compared to Peterson's NHNM and NLNM.

McNamara and Buland (2004) characterized the background noise in the United States of America using tens of stations, calculating PDFs and discussing various sources of the observed spectra depending on time of the day, the season, location and type of installation. As a reference, they used Peterson's NHNM. Díaz et al. (2010) performed similar analyses and calculated PDFs for 55 stations within the IberArray broadband seismic network, assessing the performance of the temporary installations in comparison with nearby permanent stations and the NHNM. Particular sub-networks of the AlpArray experiment (Hetényi et al., 2018) were also described, with data quality tests based mainly on noise observations, provided by Fuchs et al. (2015, 2016) (Eastern Austria and Western Slovakia), Molinari et al. (2016) (Switzerland, Italy, Croatia, Bosnia and Herzegovina and Hungary), Govoni et al. (2017) (northern Italy), Vecsey et al. (2017) (MOBNET stations in the Czech Republic), Gráczer et al. (2018) (the Hungarian network including temporary stations), Heit et al. (2021) (SWATH-D network in the Eastern Alps) and Schlömer et al. (2022a) (DSEBRA stations in Germany, Austria and Hungary). All these papers introduced availability (completeness) and PPSD plots as basic measures of the data quality. From the examples given, we see that most of the stations do fit in between the Peterson's NHNM and NLNM, however, some stations, usually those deployed in large sedimentary basins, approach or even exceed the NHNM, especially at long periods of horizontal components. Schlömer et al. (2024), describing the PACASE experiment, dedicated a whole chapter to tests of the network data. Beside the very fundamental data availability (see the detailed plot for every PACASE station in their paper), they also introduced the temporal behavior of the PPSDs during the PACASE runtime, which clearly exhibited seasonal variations between summer and winter across the entire network.

Although the PPSD plots are a valuable measure of the data quality, to assess the usability of every single station, we need to introduce other parameters as well. Among the basic ones are the availability and retrievability tests, which allow us to assess the completeness of the data in the data center and the ability of the user to download the available data on request. Additionally, we introduce noise maps to reveal potential errors in sensor gain and we conduct metadata validation tests to formally verify the correctness of information contained in the metadata itself. We also provide tests using the deconvolved seismic data to assess the mutual consistency of data and metadata, and we test the correctness of the instrument response function given by poles and zeros (PaZ). We apply our tests to both permanent and temporary stations of the AdriaArray Seismic Network (Kolínský et al., 2025). Our aim is to introduce the methods and to show examples of the results in the form of station maps without labeling the stations by their names. Results of the tests for individual stations are summarized in an online repository and are subject to changes as the errors found here are continuously being corrected and the tests are repeated.

2. Seismic data quality control

The next seven subsections focus on various phenomena related to the quality control of services, data and metadata. Each of the subsections provides an overview of the state of the art and focuses then on how we handle the discussed issue in AdriaArray.

2.1 Large dense seismic networks

Large dense seismic networks deployed around the world in the last two decades enable studying the propagation of seismic wavefields, the structure of the Earth, and sources of seismic waves in unprecedented detail. Examples of such networks are USArray (Meltzer et al., 1999; Busby et al., 2018) in the United States, AusArray, first in Northern Territory and Queensland and later covering whole Australia (Gorbatov et al., 2020; 2024). In China, the example is the NECESSArray (Ranasinghe et al., 2015). In Europe, large seismic experiments of the last decade include AlpArray (Hetényi et al., 2018; Z3 – AlpArray Seismic Network, 2015), PACASE (Schlömer et al., 2024; ZJ – Hetényi et al., 2019) and AdriaArray (Kolínský et al., 2025). For more detailed overview of seismic arrays and large passive experiments, see Section 1.4 in Kolínský et al. (2025). Hundreds of broadband seismic stations spaced by only tens of kilometers produce large amounts of data, which are usually processed by automatic routines. The data is no longer supervised by seismologists at the level of every record as thousands of hours of data are handled at once. Ensuring the quality of data and accompanying metadata is nowadays a task by itself. Besides classical techniques, which investigate the properties of data and metadata at a single station, large dense seismic networks allow for a multi-station approach to review the quality of the data. Diagnostic multi-station tools include the detection of outlying stations or records among many others. Since properties of the wavefield vary smoothly for wavelengths longer than the station spacing in case of smoothly varying Earth structure, comparing the measurements of various quantities at neighboring stations can help identify anomalous behavior. These methods work under the assumption that most of the (meta-) data is correct, and therefore a small number of outliers can be detected. Thus, not only do large dense seismic networks contribute to seismic monitoring and research, which is their primary goal, but thanks to the design of large experiments, data quality can also be tested more precisely than before when records from only single stations were analyzed.

2.2 Tested data parameters

In addition to the site noise conditions mentioned in the introduction, amplitudes of seismic events are used to assess the correctness of the declared gains of sensors, as well as the information given in the metadata. Davis et al. (2005) measured free oscillations of the Earth, excited by the great Sumatra-Andaman earthquake from 26 December 2004, using the Global Seismic Network (GSN, Butler et al., 2004). The amplitudes of long-period oscillations should be uniform globally, which allowed the authors to test the instrument response given in metadata. Ekström et al. (2006) examined gains from a huge dataset of 600 large earthquakes recorded at 200 stations between 1990-2004. The observed frequency-dependent deviations might indicate errors in the instrument responses. The authors provided also a detailed list of stations with issues. Ringler et al. (2010) compared monthly mean PSDs with long-term average mean PSDs calculated for many years. This allowed for observing temporal variations of gains for a single sensor. They focused on long-period variations in sensor properties, speculating that the cause could be degradation of the sensor’s feedback electronics. They also analyzed broadband colocated sensors comparing their gains mutually, similar to the approach by Ringler et al. (2015). Vecsey et al. (2017) compared microseismic noise using a multi-station approach. This method is also suitable for temporary stations operating over shorter time periods, enabling the detection of gain inconsistencies with an accuracy of up to 1-2 dB. Zaccarelli et al. (2021) introduced a program for the evaluation of waveform amplitudes named SDAAS – Seismic Data Amplitude Anomaly Score. The program is designed to detect outliers using a machine-learning algorithm. It can detect recording artifacts (e.g., anomalous noise, peaks, gaps, spikes), sensor problems (e.g., digitizer noise) and metadata errors (e.g., wrong stage gain in StationXML). Anthony et al. (2022) calculated noise levels for the complete USArray comprising 1679 stations. That number of stations is comparable to the AdriaArray Seismic Network with more than 1500 broadband stations. However, in the case of AdriaArray, the stations recorded simultaneously, while in the case of the USArray, around 400 stations were recording at any given moment in a rolling deployment from west to east over the contiguous U.S. (Busby et al., 2018). Anthony et al. (2022) also discussed the temporal variability of noise, and their maps are similar to the noise maps presented in our current paper.

Another crucial parameter of every three-component seismic station is its orientation towards north. Larson and Ekström (2002) determined misorientations of GSN sensors by measuring the polarization of long-period surface waves. They included a table with station names and measured values of misorientations. A similar method was used by Ekström and Busby (2008) for USArray. Various methods and techniques for determining the orientation of sensors were introduced by Ringler et al. (2013), including the description of field measurements. Rayleigh wave polarization was also implemented by Rueda and Mezcua (2015) to estimate sensor misorientations of the Spanish network.

Timing issues – another key parameter affecting the precision of seismic measurement – were found for a single station at Spitsbergen by measuring arrival times from colocated events by Gibbons (2006). Misorientations and timing issues were discussed also by Vecsey et al. (2017) and Petersen et al. (2019), who in addition provided maps showing stations with anomalously high amplitudes and misorientations in the AlpArray and SWATH-D networks. Instrumental time shifts can be detected using ambient noise cross-correlations, as shown, e.g., by Stehly et al. (2007) for three pairs of stations in Southern California or by Sens-Schönfelder (2008) for a small network at the Merapi volcano in Indonesia.

Davis and Berger (2012) focused their efforts on metadata in general, by comparing records from temporary colocated broadband sensors deployed for several hours or days during maintenance visits at the permanent GSN stations. Weidle et al. (2013) introduced an automated quality check of timing, polarity and instrument response issues by the two-station dispersion curve measurements. They showed synthetic and real examples of how various technical problems and errors in metadata affect the velocity measurement. They mentioned particular issues with response functions, including PaZ given for a different sensor than actually deployed at the stations. They also pointed to several particular stations and addressed the issues found. Cauzzi et al. (2022) analyzed the spectra of colocated sensors in Europe, comparing acceleration and broadband records. Vecsey et al. (2017) demonstrated the use of daily metrics, such as daily mean (median), standard deviation, and max-min values, for monitoring the state of health of seismic stations. These metrics can be applied to detect issues such as mass position drift, amplitude saturation, and even clipping, as well as to identify component interchange. This method is simple yet effective and, thanks to the WFCatalog service (Trani et al., 2017) in EIDA (European Integrated Data Archive; Strollo et al., 2021), it can be rapidly applied to a large number of stations.

2.3 Development of tests

The motivation for the development of quality testing procedures usually emerges during scientific or routine work of a researcher who needs to ensure that the analysis is based on reliable data. Almost every seismologist first implements specific procedures ensuring that the quality of the recordings fits the needs of the particular seismic analysis. Over the years, some of these procedures have evolved into stand-alone quality tests. This is the case, for example, of AutoStatsQ developed by Petersen et al. (2019). AutoStatsQ is now a well-established software, used to identify misorientations, polarity, and gain problems routinely by researchers as well as it is implemented at several EIDA nodes. It works as a single-station method and therefore can be used for quality checks of individual stations and sparse networks. It can detect erroneous transfer functions, amplitude gain factors (using P waves), polarity and component misorientations (using Rayleigh waves). Its original purpose was to check the dataset to be used for the moment-tensor inversion.

Similarly, a research-oriented set of tests was introduced by Kolínský and Bokelmann (2019) to ensure that the analysis of surface waves propagating across the AlpArray network was based on reliable data. The test consisted of a) gap checks, b) visual inspection, c) group velocity comparison, d) phase wavefront tracking and e) array beamforming analysis. Out of these five stages, the phase wavefront tracking is routinely used for the detection of outliers in dense networks (AlpArray, PACASE, AdriaArray), which are usually caused by errors in metadata. The gap checks and wavefront tracking tests are described here in our paper. Apart from their original research-oriented purpose, these tests are now used to detect problematic stations without reference to any particular research.

2.4 Dedicated quality tests

Additional methods were developed specifically as quality controlling mechanisms. This applies to other procedures described in our paper, namely the data retrievability test, the noise level test, checks of sensor corner periods as well as tests of StationXML formal consistency. Tests described in this paper were used for controlling the quality of data and metadata of the AdriaArray Seismic Network from the beginning, starting during the phase of the station deployment in 2022. Many errors in metadata have been corrected already. Some stations were moved to sites with lower noise when their PPSD plots did not meet the usual requirements. We also faced issues, as methods based on the detection of outliers only work for dense networks, while in the early stages of station deployment, the network was still rather sparse. Also, when many stations face problems in a certain region, detection of outliers is potentially difficult as there is no reference frame of what are the “normal” values of amplitudes, time arrivals and other observables. Therefore, the tests were repeated over the years from 2022 to 2025 using a higher number of stations and longer time periods of the AdriaArray data. As the network became denser, and as issues with some stations have been corrected, spotting the additional problems became easier and faster.

2.5 Range of tested services

The AdriaArray Seismic Network inherited 135 temporary stations from the PACASE experiment (2019-2022). Some of these stations were even incorporated from the previous AlpArray network (2016-2019). Since May 2022, when the AdriaArray started, almost 300 new temporary stations were set up in just one year. At the same time, additional almost 100 permanent stations (Kolínský et al., 2025), deployed years and decades before AdriaArray, started to share the data in EIDA managed by ORFEUS (Observatories and Research Facilities for European Seismology; Cauzzi et al., 2024). The number of mobile pools and local permanent network operators (23 mobile pools, tens of local networks, 64 institutions) was higher in the case of the AdriaArray experiment than it was before in the AlpArray and PACASE projects. This added to the complexity and heterogeneity of approaches and technical challenges that we needed to tackle to reach a uniform database of seismic waveforms.

Data of AdriaArray is stored in EIDA, which is a European federated system of datacenters (Strollo et al., 2021). AdriaArray uses 9 EIDA nodes to store data from temporary stations, and altogether 12 EIDA nodes to store data from permanent stations. Data can be accessed directly querying the particular EIDA node, via a routing service which points the user to particular node without the need of knowing what data is stored at which node. Also, open data without embargo can be accessed via the EIDA Federator. It is a complex interplay of how network operators provide the data, how the metadata is stored and accessible, the way of how the access tokens are handled, particular software used by the end user to download the data, stability of internet connection, number of queries from particular user as well as other users to the EIDA nodes and other parameters. These all together, when working properly, allow the end user to download the data. This complex environment needs to be maintained and altering any piece of the puzzle could potentially make portions of data from a station or a whole network inaccessible.

Tests presented here not only focus on data and metadata quality. We test the whole chain of steps between the data harvest at the beginning and the user performing the research at the end. We use the availability services of EIDA and compare them with our retrievability tests. This gives network providers an overview of the EIDA performance and points to problematic points between the providers and EIDA. AdriaArray accommodates both near-real-time data streamed from stations directly to EIDA as well as whole epochs of data which has been gather manually and uploaded to EIDA later. This is needed to either backfill gaps of the formerly telemetered data, or to add portions of data from times when the particular network was not online, or the data was not shared with EIDA yet. Near-real-time and backfilled data requires different handling and given the high number of network-provider – EIDA-node relations, issues arise and need to be detected. Availability and retrievability as well as access to metadata are the basic properties of the FAIR principles (Findability, Accessibility, Interoperability, and Reuse; Wilkinson et al., 2016), which the seismological community follows.

2.6 Metadata quality

Note that by "data quality" we refer to both the quality of the records themselves and the correctness of the information contained in the metadata. The fundamental difference between the two is that the metadata can be corrected at any time, while the quality of the data – once recorded – cannot be improved after recording. Problematic stations may be detected by regular checks of the records allowing to improve the setting of the stations for the future. If an error is found in the metadata, it can be corrected and used to reprocess the older data. Therefore, it makes sense to take care of the quality of metadata even retrospectively.

According to experience from the previous experiments as well as the data quality tests of AdriaArray, the majority of the data quality issues we tackle are related to erroneous metadata. Metadata is an integral part of the data. A perfectly recorded seismic waveform has no meaning without proper metadata. As mentioned earlier, about 400 new stations have been added to EIDA in the last three years in the framework of the AdriaArray experiment alone, not counting the permanent stations deployed for other networks in Europe. It is important to note, however, that the main problem lies in the metadata of the already existing permanent stations in the AdriaArray region. The same problem with permanent stations was encountered earlier, during the course of AlpArray. The metadata of newly installed permanent and temporary stations are generally correct. On the contrary, the older existing permanent stations may have undergone some changes in instrumentation over the years that are not reflected in the metadata. The most common problem is that instrumentation is not properly documented in the StationXML. Thanks to the installation of many new temporary stations in the AdriaArray region, we are now able to detect issues of permanent stations, as the network is dense enough to reveal outliers with the multi-station methods described in our paper. Application of these techniques was not possible before the AdriaArray Seismic Network was operational. Our tests described in this paper include the AdriaArray Seismic Network, which encompasses 1092 permanent and 436 temporary stations (1528 stations in total).

2.7 Accessibility of the results

To allow the station operators and field crews to fix the issues, communication channels needed to be established. Although most of the communication was based on personal links among the members of the AdriaArray community, we also prepared many posters and talks at conferences, meetings and workshops to share the results and experience with the first data testing and processing attempts. These contributions are collected in a GitHub repository, see the link in Section 5 below, including the summarizing poster by Kolínský et al. (2024) – see also Section 4.2. The feedback both from the station operators as well as from the data users led also to establishing an online repository containing the results of our tests. In this paper, we introduce a summary sheet that is publicly available online on GitHub (Section 4.1), where the results of all tests are presented in a homogeneous way so that each particular seismic station can be easily checked for all the tested properties. The sheet lists more than 2500 stations in the AdriaArray region, including the strong-motion and short-period stations, which are not part of the AdriaArray Seismic Network. In this paper, we focus on the AdriaArray Seismic Network, and all the test examples presented in the form of maps are shown for the AdriaArray stations even though some of the tests were performed for the other stations as well. In addition to the summary sheet, the results of each individual test are also available online on GitHub (Section 4.2). The purpose of all our tests is twofold: On one hand, the users should be made aware of potential issues. On the other hand, the station and network operators should be notified so that they can fix the detected problems.

The iterative process of checking the AdriaArray data has not been finished yet, and we plan to continue with the retrievability, data and metadata quality checks in the next few years. The summary sheet will be updated accordingly.

3. Tests for the AdriaArray Seismic Network

The next five subsections focus each on a particular test. Their overview is given in Table 1. We list here the tested properties, type of stations investigated, method used (single-station or multi-station), source of information needed for testing (data, metadata, metrics from EIDA). The table further shows what is the tested period of time, if the result of the test depends on when the test is performed and what is the output and estimated time cost of running the test.

3.1 Data retrievability

In case of European seismic experiments, data is stored at the EIDA federated infrastructure (Strollo et al., 2021) by ORFEUS. There are various tools to examine what EIDA stores, for example the WFCatalog (Trani et al., 2017) or the newly established FDSN (International Federation of Digital Seismograph Networks) availability service (FDSNWS-availability, see the links in Section 5 below). These services tell the user what data is available, meaning present in EIDA. To assess what data the user can get in practice, we developed a retrievability test. “Retrievability” and “availability” have different meanings. “Retrievability” tells us what can actually be downloaded by users at a given time and place. If the retrievability test gives the same result as the availability service, we know we have all the data we could possibly have from EIDA. If retrievability provides a lower score than availability, we did not manage to download all the data from EIDA. This could be caused by various reasons, such as unstable internet connection or temporary unavailability of a particular EIDA node. We have also experienced varying results of retrievability tests when performing the test at the same time from different sites across Europe. These differences are usually due to internet connection issues specific to the individual sites where the tests are performed. The results of the retrievability score also differ when the same test for the same time period is run at different times. This is usually due to service outages of the EIDA nodes during the time of testing. In addition, the network

Table 1. List of tests and their properties discussed in this paper. The column labeled as "section" refers to the subsection of this paper. Here we point to the basic characteristics of the tests. In fact, all the tests using metadata depends on time as metadata could be potentially updated anytime. "Output" lists what quantity we use to communicate the results and "hours to perform" shows, how long does it take to run the test. Note, that StationXML formal checks concern many properties, which are listed separately in Table 2.

<table>
<tr><th>Test</th><th>Section</th><th>Tested properties</th><th>Station type</th><th>Single-/ multi</th><th>Source of information</th><th>Tested period</th><th>Depends on time</th><th>Output</th><th>Hours to perform</th></tr>
<tr><td rowspan="4">Retrievability</td><td rowspan="4">3.1</td><td>FDSN-availability</td><td rowspan="4">Any</td><td rowspan="4">Single</td><td rowspan="2">Taken from EIDA</td><td rowspan="4">Any/ Month</td><td rowspan="4">Yes</td><td>Percentage</td><td rowspan="4">4</td></tr>
<tr><td>WFCatalog availability</td><td>Percentage</td></tr>
<tr><td>Retrievability</td><td>Data</td><td>Percentage</td></tr>
<tr><td>Metadata deconvolution</td><td>Metadata</td><td>Yes/No</td></tr>
<tr><td rowspan="3">Noise levels</td><td rowspan="3">3.2</td><td>Amplitude @3.0 Hz</td><td rowspan="3">Any</td><td rowspan="3">Multi</td><td rowspan="3">Data
+
Metadata</td><td rowspan="3">Day</td><td rowspan="3">No</td><td>Histograms</td><td rowspan="3">8</td></tr>
<tr><td>Amplitude @5.0 s</td><td>Histograms</td></tr>
<tr><td>Amplitude @20.0 s</td><td>Histograms</td></tr>
<tr><td>StationXML</td><td>3.3</td><td>See Table 2 for details</td><td>Any</td><td>Single</td><td>Metadata</td><td>All/AdA</td><td>No</td><td>0-5 grades</td><td>8</td></tr>
<tr><td rowspan="2">Earthquake data</td><td rowspan="2">3.4</td><td>Technical checks</td><td rowspan="2">BB</td><td rowspan="2">Single</td><td>Data</td><td rowspan="2">40-90 mins</td><td rowspan="2">No</td><td>Keywords</td><td rowspan="2">4</td></tr>
<tr><td>Metadata deconvolution</td><td>Data + Meta</td><td>Yes/No</td></tr>
<tr><td rowspan="3">Wavefront tracking</td><td rowspan="3">3.5</td><td>Transfer function in PaZ</td><td rowspan="3">BB</td><td rowspan="3">Multi</td><td rowspan="3">Data + Metadata</td><td rowspan="3">40-90 mins</td><td rowspan="3">No</td><td>Yes/No</td><td rowspan="3">4</td></tr>
<tr><td>Time shifts</td><td>Yes/No</td></tr>
<tr><td>Polarity (if phase used)</td><td>Yes/No</td></tr>
</table>

operators backfill the data to the EIDA over time, and in this case, later tests give higher retrievability percentages. Sometimes, we download more data than what the availability service suggests. It happens when the availability was not yet updated in EIDA after backfilling.

The procedure consists of requesting five random 10-minute blocks of data for a given set of stations within a specific time period (e.g. March 2024). All requests are using the EIDA routing service. If any request fails for any reason, it counts as negative and is not repeated. Whenever possible, this test is done simultaneously from multiple different places around Europe. The result is presented in the form of a map showing the percentage of retrieved data using a red-green color scale. The percentage is calculated by the number of successfully downloaded time windows from all simultaneous test runs divided by the number of total requests from all test runs. Note again that the results do not describe the presence of waveform data at EIDA nodes but the ability to download the data at the time of the test from a user's perspective.

We show three examples, to demonstrate both the methodology as well as the actual results. While the test has been used also for the region of whole Europe, here we focus on the AdriaArray Seismic Network. The retrievability tests were repeated for 22 months of AdriaArray data starting with June 2022 and ending with March 2024. Every month was evaluated by a separate test. All these tests were made on April 10th and 11th, 2024. The map in Fig. 1 shows the results for January 2024, while Fig. 2 shows the same for March 2024. The maps show all stations included in the AdriaArray Seismic Network, both permanent and temporary. Besides the red-green colored stations, other stations are shown by white and gray triangles. Stations which were not operational at the tested time period are marked in gray. These stations were either not deployed yet, or had terminated its operation before the tested time period. In case of closed stations, their location is taken from their metadata in EIDA. In case of stations not deployed yet, their location is taken from the AdriaArray inventory (see Kolínský et al., 2025). Before every test, an updated inventory of stations available in EIDA is created using the FDSNWS-station webservice – it means, we search for the StationXML metadata. Stations for which the metadata was found are then queried during the test. However, there are also cases when some of the existing stations do not show up in the particular EIDA inventory before the test, even though we know these stations are in EIDA. To keep the map of the AdriaArray Seismic Network complete, we plot these stations as white triangles – information about them is again taken from the AdriaArray station inventories. White triangles with black dots were stations not registered in EIDA at the time of performing the test. These were planned to be connected to EIDA in the future.

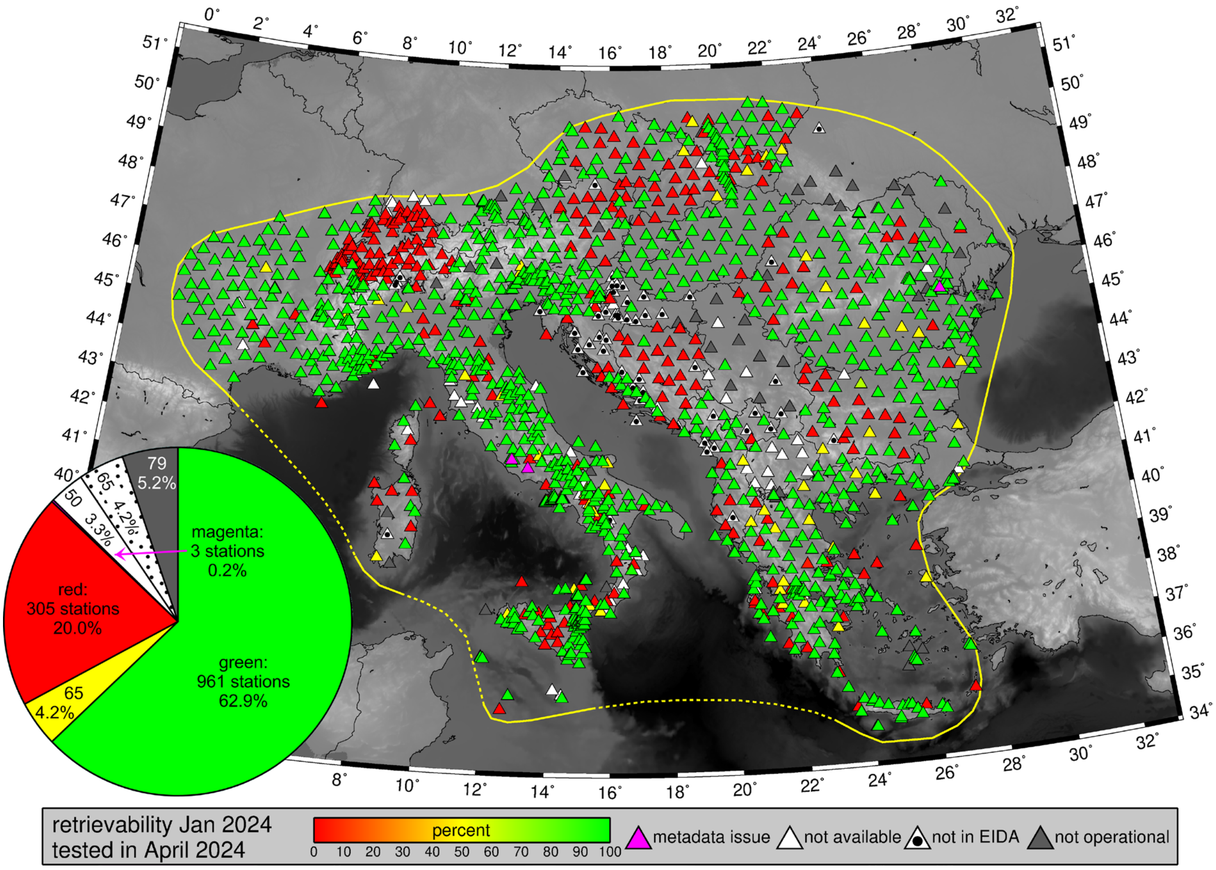

Figure 1. Data retrievability of January 2024 tested in April 2024. The pie chart on the left shows numbers and percentages for particular categories of stations. For simplicity, stations with retrievability between 20% and 90% are shown in yellow in the pie chart. Similar pie charts are also used in next figures with maps. They always cover 1528 stations in total.

Note, that in case of the operational – non-operational stations, their status relates obviously to the tested time period – January or March 2024 in our examples. However, in case of EIDA – non-EIDA stations, their status relates to the time when the test was performed – April 2024. There could be stations, which were operational during the tested time period, but not connected to EIDA at that moment. Registering the station to EIDA and backfilling the previously recorded data allows retrieving data also for those stations later. Magenta triangles in Figs. 1 and 2 show stations where a serious metadata problem did not allow to deconvolve the instrument response.

The red-green scale shows the percentage of data retrieved for each particular station. There are several reasons why the fully red triangles, meaning "no data retrieved", appear on the map. 1] The data is indeed not available, although the station operates, is online and streams its data to EIDA. This could be caused by malfunction of the station itself, by unreliable telemetry, or by temporal unavailability of the EIDA node. 2] The station is operational and online, registered in EIDA, but data is sent to the network operator only, and not yet to EIDA. This is the usual "testing" phase for many newly installed stations. 3] The station is deployed and working properly, but does not have telemetry and no data had been collected manually before the test was performed.

Comparing Figs. 1 and 2, we see clear differences. Localized patches of red color in the January 2024 (Fig. 1) test show that an entire EIDA node was not providing data during that specific test – we see red stations in Switzerland and in Bosnia and Herzegovina, which are all stored at the ETH EIDA node. This was a temporary issue which did not affect the next test for March 2024 (Fig. 2). Improved telemetry also plays a significant role in that the test in Fig. 2 shows more green stations – stations in eastern Austria were not online in January and started to stream the data later in March.

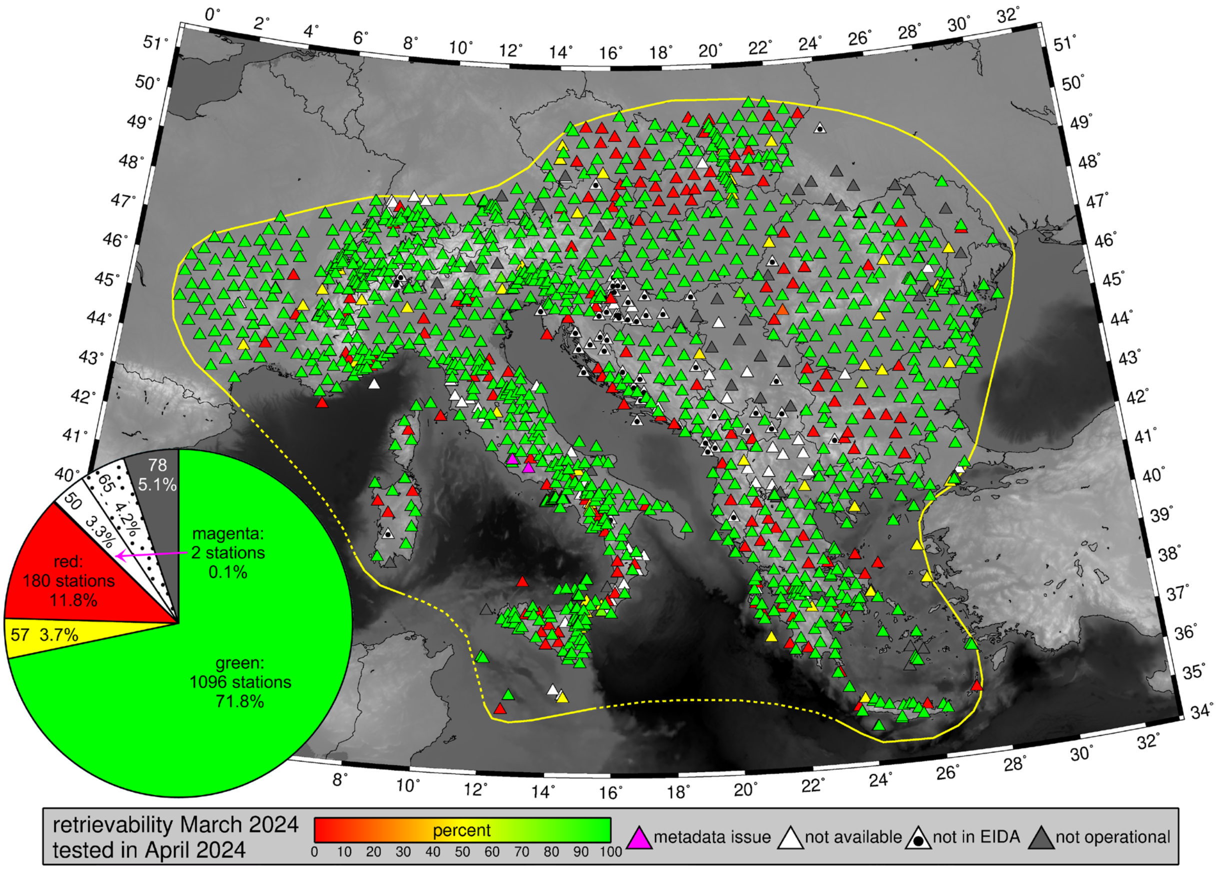

Figure 2. Data retrievability of March 2024 tested in April 2024.

The third example is given in Fig. 3. It again shows a test for March 2024, however, performed in February 2025. We see that more stations were available from EIDA – there are fewer white triangles than in Fig. 2. Also, overall, there are many more green stations compared to April 2024 as data has been backfilled to EIDA meanwhile.

During the test, not only the retrievability is tested, but also the availability from the WFCatalog is gathered, and the FDSN availability webservice is exploited. The summary sheet (discussed in Section 4.1 below) includes the percentage of all three quantities (retrievability, WFCatalog, FDSN availability). In practice, the scores differ for reasons mentioned above. Sometimes, retrievability gives an even higher percentage than the availability services, which could be caused by the fact that these services were not updated when data was backfilled retrospectively.

This retrievability test is based on sampling short time windows to give a quick statistical overview about data of large networks and possibly larger time periods. Other papers of this special issue show data availability

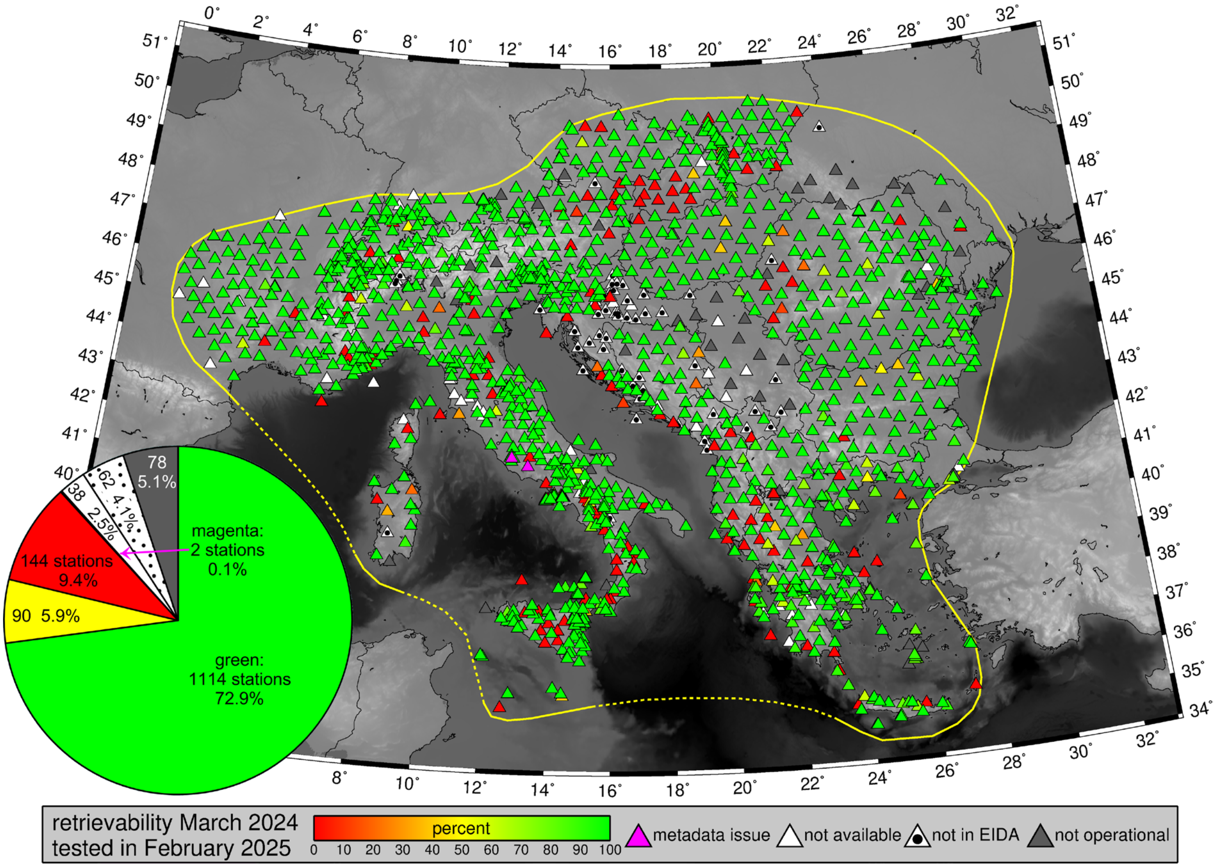

Figure 3. Data retrievability of March 2024 tested in February 2025.

based on downloading all data for a given time span, usually from the beginning of the AdriaArray initiative, May 2022, see Kampfová Exnerová et al. (2025), Vecsey et al. (2025), and Borleanu et al. (2025). These are not tests, but true figures of data available in EIDA – the completeness plots. Creating these availability/completeness plots is time consuming, and unfortunately still requires repeated querying of the EIDA nodes. It is usually done only for a particular network or regional subset of stations. The retrievability test presented in our paper is fast and allows for a quick overview of the whole continental-scale station coverage. It was performed not only for the AdriaArray region, but for the whole dataset encompassed in ORFEUS EIDA, see the GitHub link in Section 5 below. The retrievability test allows us to assess not only the performance of individual stations, but also potential issues with the availability of the particular network or individual EIDA node.

3.2 Noise levels

The noise level test shows amplitudes of noise for each tested station. Although the measurement is done for each station separately, the main strength of the test is that it allows comparing every station with its neighboring stations to identify outliers. This test is therefore one of those that make use of the dense distribution of stations (multi-station method). The noise level maps can show smooth variation of the noise among particular regions, but they mainly highlight stations that report amplitude levels which are orders of magnitude off the neighboring stations.

The test is based on data from a whole day without major teleseismic earthquakes. It includes broadband, strong-motion and short-period sensors. Using data from multiple days was considered but the noise levels were found to be stable already when only 24 hours of data were used in the processing. The raw waveform records are downloaded and the response of the instrument is deconvolved. This implies that possible errors in the measured amplitudes can also be caused by erroneous response functions in the metadata. The procedure is repeated for 3 frequency bands: 3.0 Hz (bandwidth: 1.5-6 Hz) and 5.0 s (2.5-10 s) for the vertical component and 20 s (10-40 s) for the horizontal component. From the absolute values of noise amplitudes, the 95th percentile is calculated and represented by a color scale on the map.

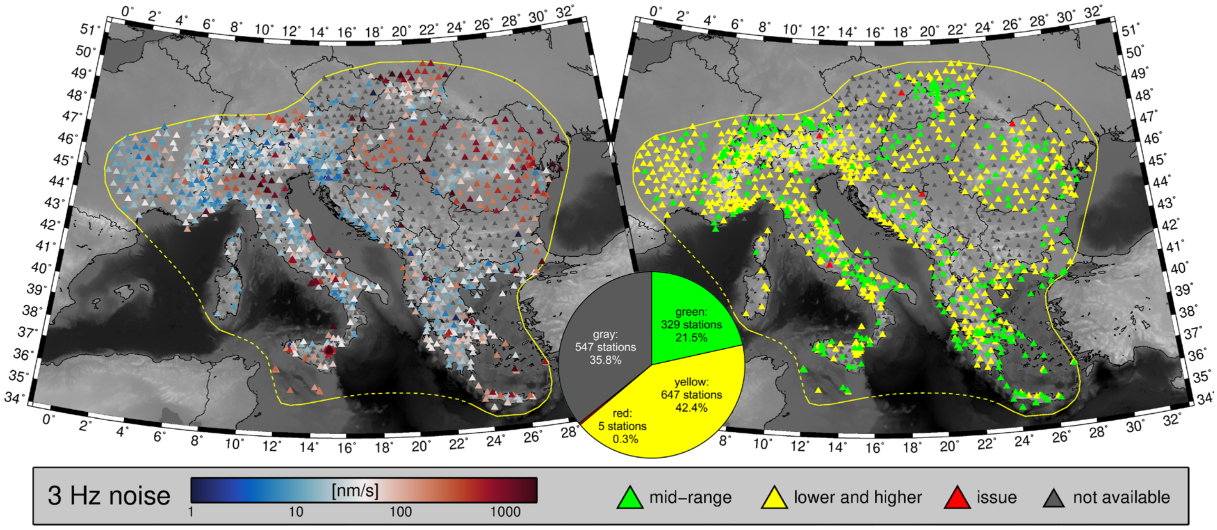

Figure 4. Left: noise levels for the AdriaArray stations extracted for the 3 Hz band. Right: Simplified evaluation of outlying stations. Not-available stations are shown by smaller dark gray triangles to distinguish them from the high-noise stations which also appear with darker color.

Left panels of Figs. 4, 5 and 6 show the noise levels for the AdriaArray region for the 3 Hz, 5 s and 20 s bands, respectively. The color axis is logarithmic because the noise levels span over multiple orders of magnitude. Note the different ranges of the color scales in the three figures.

The 3 Hz band test (Fig. 4) is performed on the vertical component. This frequency band is sensitive to local noise. It can highlight stations at noisy sites. There is a clear correlation with topography – the 3 Hz band shows lower noise in the mountain regions (Alps, Dinarides, Apennines) and higher noise in the sedimentary basins (Po Plain, Pannonian Basin, Molasse Basin, Moesian Platform). The 5 s band test (Fig. 5) is performed also on the vertical component. This frequency band is mainly affected by secondary microseisms. It showcases stations that are sensitive or insensitive to microseisms and we see a clear trend of higher noise levels in the north and lower levels in the south of the AdriaArray region, corresponding to the expected stronger microseism sources

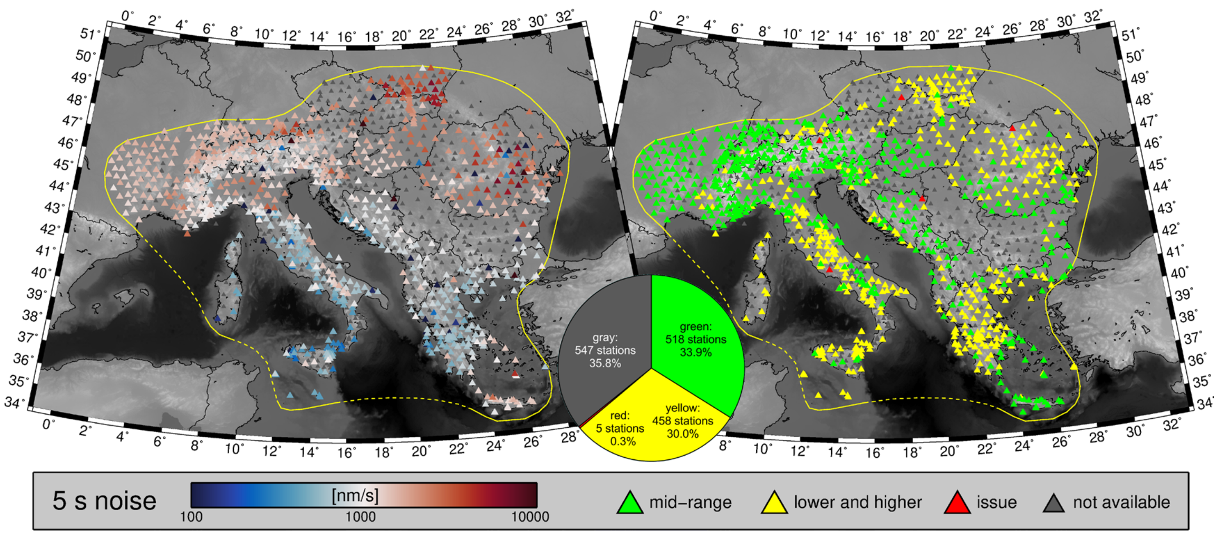

Figure 5. The same as in Fig. 4 for the 5 s band.

in the North Atlantic than in the Mediterranean Sea (Juretzek and Hadziioannou, 2017; Lu et al., 2022). The 20 s band test (Fig. 6) is performed on the horizontal component. Site effects play a minor role in this frequency band, and also the north-south trend is missing. Instead, there is a slight general decrease of noise levels from northeast to southwest. Seismic noise at this period is partly affected by long-period microseisms.

In addition to smoothly varying noise levels, outliers are observed throughout the AdriaArray region. To reveal these, the number of stations for narrow noise level bins is plotted as histograms in Fig. 7 for each frequency band. We see a quasi-Gaussian distribution (note again the logarithmic horizontal axis) with a small number of detached stations with unreasonably high or low noise levels. Those noise levels below 0.1 nm/s or above 10^6 nm/s are unrealistic and are labeled as "invalid" for all three frequency bands.

Maps on the right in Figs. 4, 5 and 6 then show a simplified evaluation of the noise levels based on the distribution in histograms in Fig. 7. Mid-range levels (between 25 nm/s and 200 nm/s for 3 Hz and 20 s band, and between 800 nm/s and 2000 nm/s for 5 s band) are indicated in green. Stations above and below these mid-range levels are in yellow – they cluster in regions of specific topography and represent mainly smoothly varying noise levels. Outlying stations, detached from the main histogram distribution below 0.1 nm/s and above 10^6 nm/s limits are plotted in red.

Performing the test repeatedly over the last years has already revealed outlying stations clustered in particular regions. Thanks to this finding, a systematic error in metadata handling caused by different versions of software used by different EIDA nodes was found. This issue was corrected and the test was performed again to confirm the accuracy of the metadata handling. The test was performed for all types of seismic stations, see the summary sheet (Section 4.1 below). In the figures, however, we show only the AdriaArray stations for simplicity.

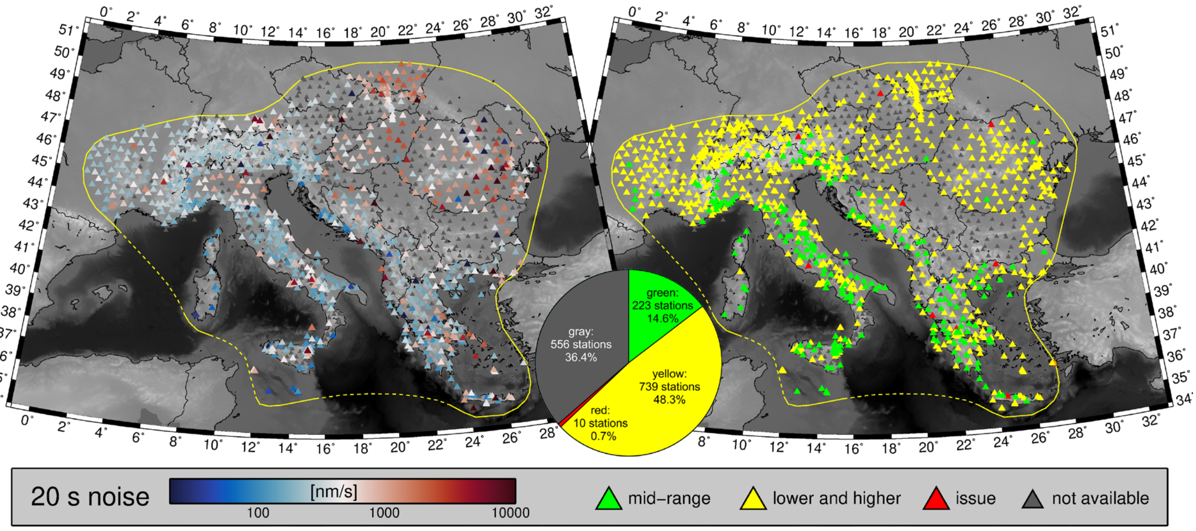

Figure 6. The same as in Figs. 4 and 5 for the 20 s band on horizontal components.

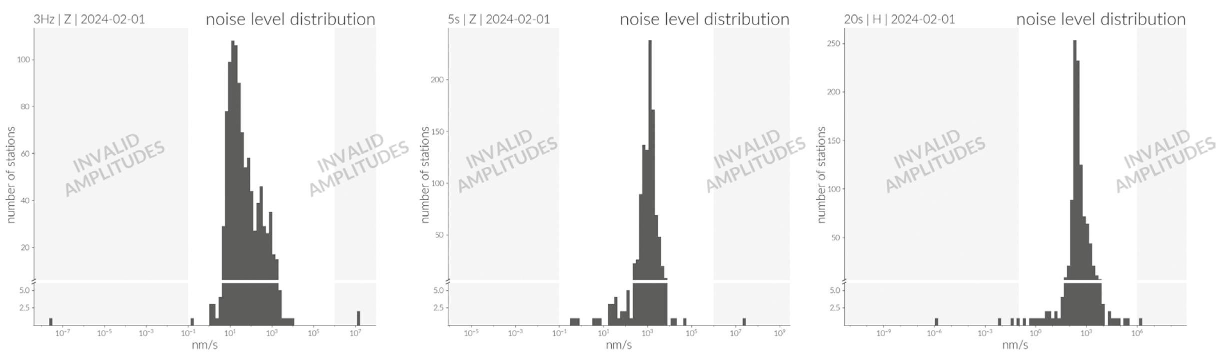

Figure 7. Noise levels given in histograms for the three frequency bands as in Figs. 4, 5 and 6. Outlying stations are clearly visible, detached from the main distribution of noise levels.

3.3 StationXML internal checks

We also developed and performed a suite of metadata tests based on checking various formal properties of StationXML files downloaded from EIDA. Tests include validating the FDSN StationXML format by using the FDSN validation schema (see the links in Section 5 below), a set of validation checks performed by the IRIS StationXML Validator (IRIS, 2023) and additional newly developed tests aimed at PaZ stages (portions of StationXML file listing poles and zeros of the sensor transfer function), calculating and checking the corner periods of sensors and comparing the channel codes with respect to the MiniSEED convention (FDSN et al., 2012). Every issue is assigned one of the six grades of severity. Grades [0] and [1] (see Table 2 for details) are only notifications without any significance for applying the StationXML to data. These are issues like the wrong format of certain values, missing sensor description, using future dates, which is not recommended, and lower case/upper case letter inconsistency in input/output unit names. Grades [2] and [3] represent warnings that most probably do not affect the data in a wrong way when the StationXML file is used for deconvolution but may cause problems when searching for certain properties of the stations, for example, their operational period. These issues include overlaps of epochs, inconsistent start dates, for example when the Station element has an earlier startDate than the Network element, and incorrect channel names with respect to the sampling rate. Grades [4] and [5] are significant errors which distort the data so that it cannot be used for further processing or which do not allow to deconvolve the data at all. These are, e.g., significantly different coordinates of the Station and Channel elements, missing entries as Azimuth and Dip, invalid units of the first and last stages, inconsistent sample rates between stages, missing PaZ and several others. Table 2 lists all the checked parameters sorted first by the method used to check them, then by the targeted element in the case of the IRIS Validator, and then by their severity.

For every StationXML, all these tests are performed and listed in the detailed sheets, see Section 4.2 below for the link to the respective GitHub page. For each StationXML, the worst assigned grade is then selected from all the tests for the summary sheet (see Section 4.1. below) and it is also represented by one of the three colors in a map, see Fig. 8. Green triangles represent stations, where there was no issue in the StationXML found, or the test only gave a notification (grades [0] and [1]). Orange triangles denote stations with warnings (grades [2] and [3]) and red are stations with serious issues found in their metadata (grades [4] and [5]). Gray are stations which have not been tested as their metadata is not in EIDA. This includes both stations which are operational but not registered with EIDA, as well as stations which were not deployed yet at the time when the test was performed.

The summary sheet, as well as Fig. 8, gives an overview on the most serious detected issues but does not give details of the test results and the number of issues found for each individual station or what severity was found for the other issues. As an example of a particular subset of tests, we present here the tests of corner periods, see Fig. 9. PaZ of the instrument transfer function are used to calculate the corner period for every channel. Note that the corner period itself is not a mandatory parameter and hence does not need to be present in the StationXML. Many station operators now give this value in the optional element with a comment on the sensor type, which is good practice. The calculated corner period is then compared with the given channel name. Warnings (orange triangles) are given for stations where a potentially wrong band code was found. This means that the calculated corner period does not match the channel naming convention as given in SEED. StationXML affected by this issue can still be used for deconvolving the data successfully, if the channel name in the metadata matches the channel name in the data, meaning if both are wrong in the same way. The red triangles mark the stations where there is a confusion between the names of short-period and broadband channels. A user downloading only broadband (or only short-period) stations from EIDA may completely miss a given station or, conversely, may retrieve data recorded by a sensor with a different frequency range than intended.

There is an additional test, which is given in the detailed sheets but not shown in the figure. As we do have an independent source of information about the corner periods in our AdriaArray inventories (see Kolínský et al., 2025) we also provide a cross-check of the corner periods calculated from the StationXML with those listed in the station inventories. This comparison with the independent list of sensors and their corner periods shows that the corner period differs slightly for around 10% of stations and significantly for 3% of stations. However, the discrepancy does not tell us whether the corner period in the independent inventory or the transfer function in the StationXML file is wrong. The values in the station inventories were collected from the network and station operators in 2019 and 2020 for the permanent stations and could change over time. Still, this cross-check allows us to identify suspicious stations and correct the value in either of the two independent sources.

Table 2. List of StationXML checks.

Method	Element	Grade	Test
FDSN StationXML validation		1	wrong value format in ClockDrift
		3	missing attribute
IRIS StatinXML Validator	Network	1	Network and/or Station endDates set to future
		3	Station epochs overlap
		3	Station startDate before Network startDate
	Station	3	Channel startDate before Station startDate
		4	Station and Channel Latitudes/Longitudes differ too much
		4	Station and Channel Elevations differ too much
	Channel	1	no Sensor Description in Channel
		1	mis-orientation ≥ 5, component code should be one of [123]
		3	Channel.startDate should exist and be before Channel.endDate
		4	missing Azimuth and/or Dip
		4	invalid InputUnits (first Stage) and/or OutputUnits (last Stage)
	Response	0	case inconsistency in Input/OutputUnits
		3	invalid Input/OutputUnits
		3	InstrumentSensitivity.Frequency must be < SampleRate/2
		3	missing at least one Decimation in Stages
		4	invalid InputUnits in the first Stage
		4	Input/OutputUnits do not follow in Stages
		4	missing Value in InstrumentSensitivity
		4	StageGain.Frequency is zero in undue Stage
		4	missing InstrumentSensitivity in Response
		4	output samplerate from Stages not equal to Channel.SampleRate
		4	samplerates do not follow in Stages
		5	missing Decimation and/or StageGain in Stage(s)
Originally developed tests		2	corner period confusion between real and tabled values
		3	real(Pole) ≥ 0 in the analog PolesZerosResponseStage
		3	incorrect channel band due to sampling rate
		4	channel band confusion in broadband and short-period
		5	response failure
		5	no Poles in PolesZerosResponseStage
		5	the last response stage is faulty
		5	no conjugated Poles in the PolesZerosResponseStage

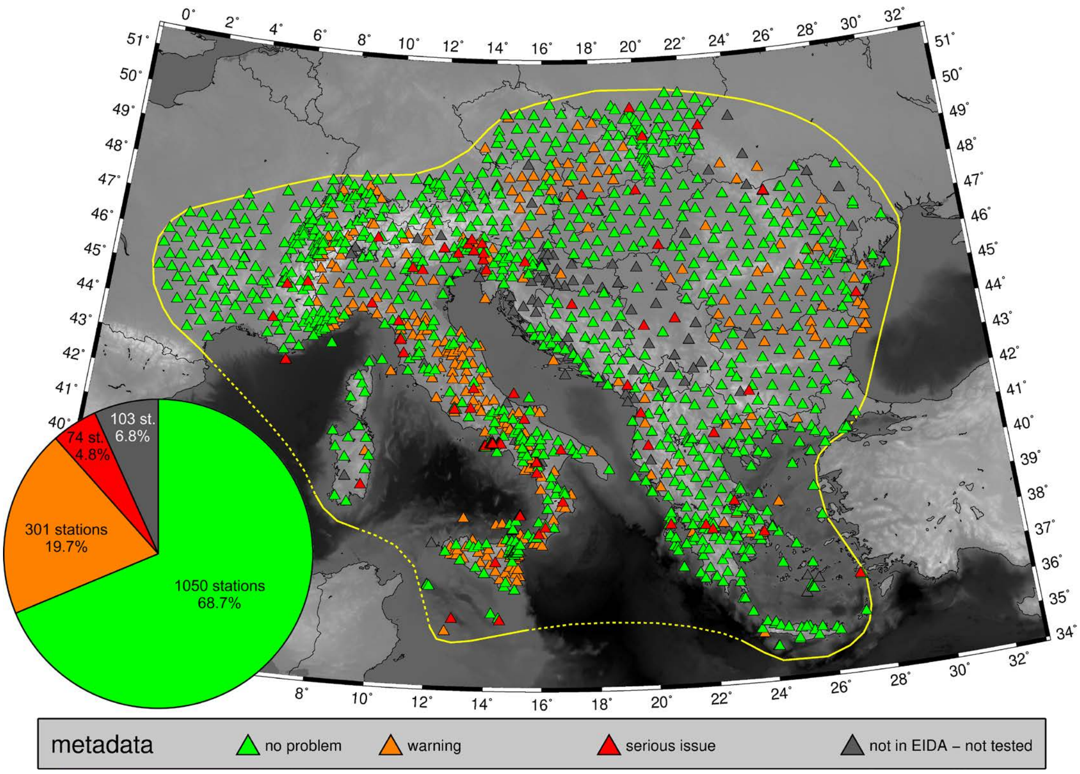

Figure 8. Overview of issues found by the metadata checks. Stations with serious problems affecting the deconvolution are in red. Gray are stations which do not have their metadata in EIDA. This includes both stations which are operational but not registered with EIDA, as well as stations which were not deployed yet at the time when the test was performed.

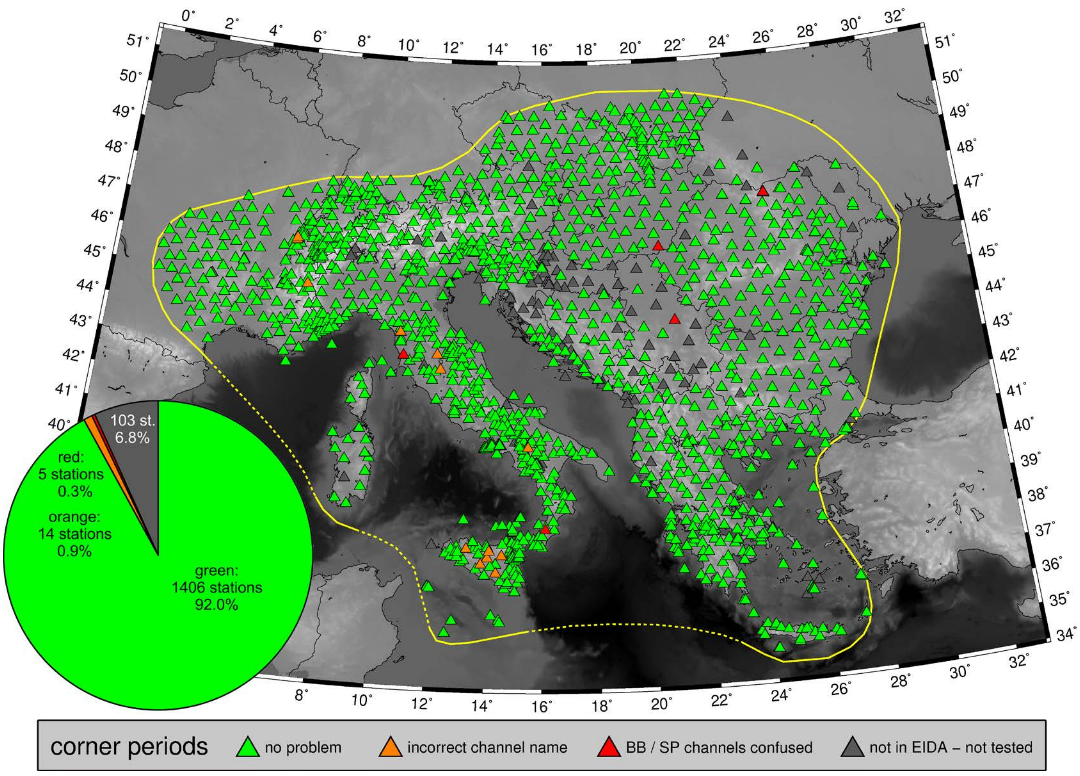

Figure 9. Overview of the corner period evaluation. Stations with severe inconsistency of channel naming are in red. Gray stations are the same as in Fig. 8.

As in the previous tests, both Figs. 8 and 9 show stations of the AdriaArray Seismic Network (1528 stations). However, the tests were performed for more than 2000 stations, including strong-motion and short-period ones. Note that the independent station inventories by Kolínský et al. (2025) also list the strong-motion and short-period stations.

These tests of StationXML were done for all the epochs included in the particular StationXML, which, for some permanent stations, means that we tested their metadata for years and even decades before the AdriaArray started. For the purposes of this paper and examples given in Figs. 8 and 9, we only selected the worst grades for epochs which do have an overlap with the operational period of AdriaArray, meaning for stations, which continued or started their operation after May 2022. Hence, tests of epochs closed before May 2022 are not shown here.

3.4 Earthquake data quality check

Similar to the method introduced by Kolínský and Bokelmann (2019) for AlpArray, we use teleseismic earthquake records of the AdriaArray stations to perform tests of the data quality. This test encompasses several stages of checking the data itself for technical imperfections, deconvolving the data using StationXML allowing for detecting anomalous behavior due to issues in the metadata and continuing with data processing and precise surface-wave group and phase velocity measurements, which allows to detect outlying arrival times across the whole dense network and to point out suspicious instrument transfer functions given by PaZ in StationXML. This Section focuses on the data quality and instrumental transfer function deconvolution detecting stations for which the StationXML is invalid in a formal sense, meaning it cannot be used for deconvolution. The next Section 3.5 then introduces the wavefront tracking for detecting issues with PaZ, evaluating the actual information given in the StationXML.

The evaluation of the quality of teleseismic records described in this section was developed for surface-wave analysis. It targets specific properties of the data, which are not necessarily required for other techniques. For example, we examine quite a long portion of data, usually 40-90 minutes, as surface waves are dispersive and are recorded over a long time window when the epicentral distance is several thousand kilometers. Moreover, to calculate the spectrum, we need a sufficient portion of data before and after the surface wavetrain itself. Our procedure, for example, includes a check for gaps in the data. In case a seismologist is interested in, e.g., receiver function analysis, the requirements for data quality would focus on significantly shorter time windows and gaps around that window would not matter. On the other hand, given a long-period nature of teleseismic surface waves, we allow for quite long gaps (up to 20 seconds). Other techniques would be probably more strict, allowing for significantly shorter gaps. The output of this procedure provides an overview of what is wrong with the data and metadata, however, the data can potentially still be used for specific tasks. The result is included in the online summary sheet of all tests in detail as well as simplified in the maps here in the paper.

As an example, we requested all the AdriaArray stations (1092 permanent + 436 temporary) for the M7.4 earthquake which occurred in Taiwan, 2024-04-02, origin time 23:58:11.0. We downloaded 3 hours of data, starting with the origin time. Then, we selected an 83-minute time window focused right on the earthquake record, on which we ran the quality check. The reason for that double windowing is that we want to have some margins around the selected time window and not knowing the exact characteristics of the record, we rather download a larger time window first. It might happen that although we do get a downloaded file for a particular station, there is no data in the shorter time window of 83 minutes.

The map in Fig. 10 shows the results. White, black and gray are colors denoting stations for which we did not have any data file. Stations not deployed on the examined date are denoted by gray triangles. Stations for which we did not get any downloaded file even though they were operational are denoted by white triangles. Those of them not connected to EIDA at the time of the download are marked with a black dot inside the white triangle. We knew beforehand that we could not get data from these “black-dotted” stations. However, it is considered here as a specific type of issue, as these stations are operational and the data could have been available potentially, if the station and network operators have shared it via EIDA. The purely white stations are clearly missing in our dataset even though we know they are connected to EIDA and operational. See a similarity with the previously introduced retrievability checks, where we also needed to assess the presence of the stations in EIDA at the time when the test was performed, and the operational status (deployed – not deployed) of the stations at the tested time period. For the stations with a data file, we deconvolved their instrument transfer function using the respective StationXML. Magenta triangles depict the stations for which the deconvolution was not possible due to an error in the metadata. These errors are

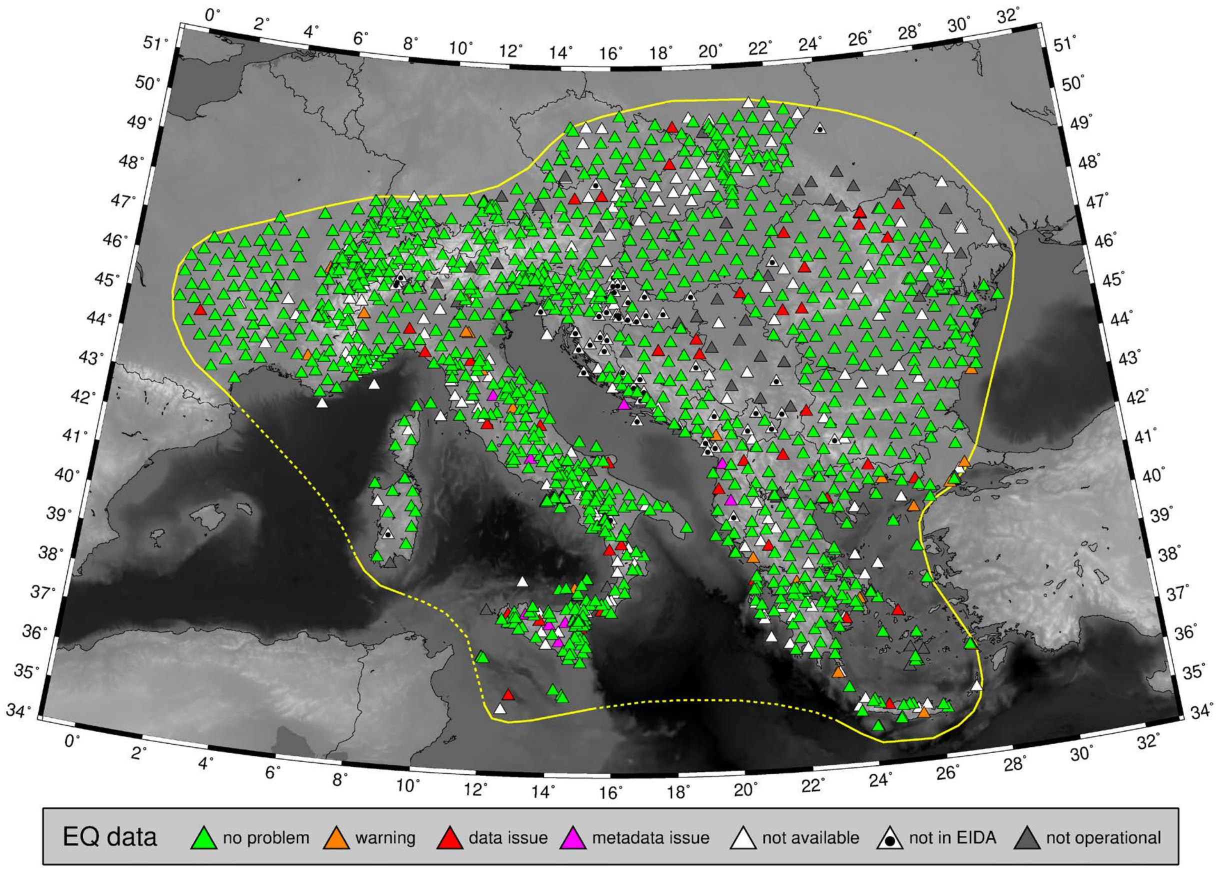

Figure 10. Data and metadata check for the Taiwan M7.4 earthquake from 2024-04-02. Magenta triangles are the stations for which we were not able to deconvolve the data. Red are stations where the data suffered by serious technical problems, mainly gaps, and could not be used for further analysis. The pie chart for this map is given in Fig. 14.

usually either different channel names in the data and in the metadata (HH* vs. BH*), wrong azimuth and dip for the components, so that they are not orthogonal, FIR filters yielding zero or infinity, missing elements (gain), missing stages (PaZ), or the epoch closed in metadata with a date before the earthquake.

For several of these issues, their correction is simple and obvious, and we manually corrected our local copy of the StationXML to have the transfer function deconvolved successfully. We did this to check if the StationXML may contain more problems – by correcting one of the errors, we sometimes discovered additional problems. Nevertheless, we still mark these stations as having an error in metadata, as the only one responsible for correcting the metadata properly and providing the updated version to EIDA is the network operator. After deconvolution, we tested the data itself. Red triangles are the stations, for which the data was deconvolved, but did not pass the quality check because either the gaps in the data are too long (any gap is longer than 18 s), the whole component is missing, the record is shorter then requested, there is no data in the selected 83 minutes, one (or more) of the components is zero, or there is a significant difference in maximum amplitude between any two components (by order of magnitude).

Orange are the stations where there was a minor data issue which was solved at the downloaded files so that the data could be used for further surface wave analysis. These issues are either that there were more than 3 traces in the stream and they were merged, there were small gaps (up to 18 s) needed to be covered with interpolation, traces were overlapping and the selection among them was made, or there were traces with mixed location codes and channels with the same location code were selected.

It means that whenever we need to perform any selection or merging of the data, we mark it as a warning, as such handling could potentially cause additional imperfections (interpolation, selection of overlapping traces). Green are the stations where there was no problem. To summarize the numbers for the example of the Taiwan earthquake, out of the 1528 AdriaArray stations, 15 temporary and 25 permanent stations were still not deployed at the time of the earthquake and 26 stations stopped their operation before the examined day. From the remaining 1462 operational stations, 65 were not connected to EIDA at the time of the download (December 2024) and hence we could not get their data. We downloaded data from 1208 stations, which means that there are still 189 stations (white triangles) where there was no data downloaded even according to all available information, the stations should have been operational and streaming data to EIDA. Then, 12 stations (0.8%) had invalid metadata (magenta) and 46 stations did not pass the data quality check (red). We perform the data check separately on vertical and horizontal components. 23 stations did not pass the vertical-component check, and additional 23 did not pass the horizontal-component check. We are then left with 1173 stations with useful vertical components. However, the tests done up to now still do not evaluate if maybe some of the records are noisy or contain other peculiarities which would prevent us from successful surface wave analysis. These will be identified and removed in the next step, see Section 3.5 below. If all three components are needed for the analysis, we can process data from 1150 stations.

3.5 Wavefront tracking

The wavefront tracking procedure is based on picking the group velocity propagation times, or picking the same phase of teleseismic surface waves with high SNR at all stations in a given network or region and then displaying the map of the propagating wave in terms of contours for constant time steps – usually 5 s. The wavefront is expected to be smooth. Before the measurement, we first inspect all the records visually and remove those where there is not a clear earthquake record. This step usually removes less than 1% of stations. Note that most of the technically insufficient records were already removed by the tests according to the previous Section 3.4. Performing the measurement only on high-quality data is important to ensure that the evaluation of the transfer function given in the metadata by wavefront tracking is not affected by the low quality of the data itself.

The procedure to obtain the wavefront-tracking maps is as follows: The fundamental mode of surface waves is extracted by classical frequency-time filtering (Kolínský and Bokelmann, 2019). During this step, we identified several stations (around 1%), where the fundamental mode selection was not possible due to the bad quality of the data itself. During the deconvolution, wrong PaZ or missing entries in the sensor response result in a significant phase (time) shift, which shows up as a distortion of the wavefront (phase velocity) and also of the envelope maxima (group velocity). This method is based on the assumption that most of the stations tested have correct metadata, and hence outliers can be identified. It is a typical multi-station method. The times (group or phase) of arriving surface waves are picked at each station individually, but the quality assessment is only done when the whole network is considered at once.

We first started with these wavefront-tracking tests in 2017 for the AlpArray dataset. Here we show an example of the wavefield propagating from the M6.9 earthquake in Japan, 2016-11-21, after separating the 102-s-period Rayleigh wave on the vertical component. Left panel in Fig. 11 shows the initial result distorted by incorrect PaZ

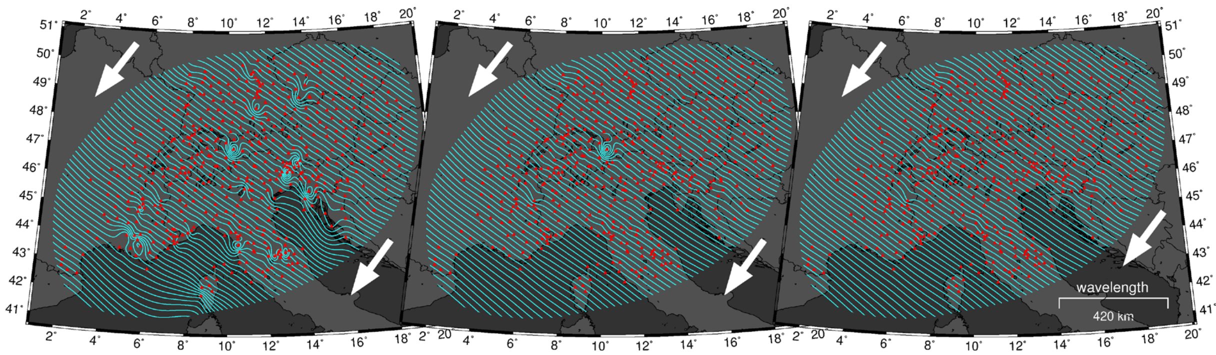

Figure 11. Wavefront tracking for 102 s Rayleigh wave propagating from the Japan M6.9 earthquake, 2016-11-21, recorded by the AlpArray Seismic Network. Phases are used in this example. Contours are plotted for every 5 s of wave propagation time (phase velocity). Left panel shows the original measurements at 526 stations (red triangles). In the middle plot, 25 suspicious stations are removed, and one is left. In the right panel, that one from the middle panel has been corrected. Hence, both the middle and the right panel show measurement at 501 stations. Arrows show the direction of wave propagation.

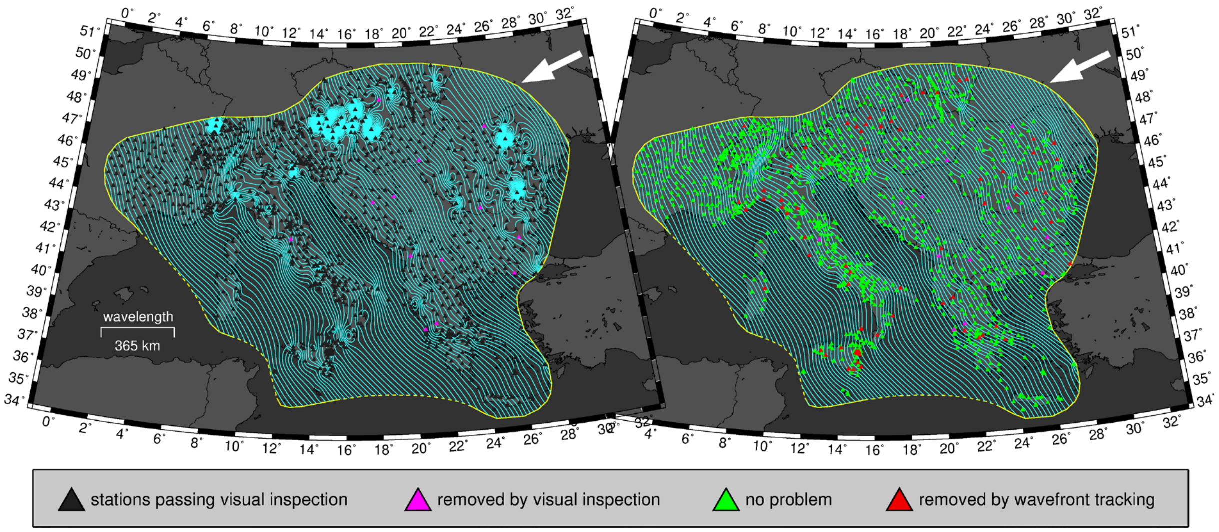

Figure 12. Wavefront tracking of 90 s Rayleigh wave propagating from the Taiwan M7.4 earthquake, 2024-04-02, recorded by the AdriaArray Seismic Network. Group propagation times are used in this case. Left panel shows the initial measurement using all stations with downloaded data which passed the visual inspection. Right panel shows the same wavefield after removing the suspicious stations identified in the left map. These removed stations are marked by red triangles in the right map. Stations removed by the visual inspection are plotted as magenta triangles in both maps. Arrows show the direction of wave propagation.

in 26 StationXML responses (out of 526 stations which passed the quality checks from Section 3.4). We see “bubbles” of heavily distorted propagation time contours, of size much smaller than the wavelength (given in the right map of Fig. 11). By removing the suspicious stations one by one and by replotting the wavefronts, we were able to confirm that indeed each of these particular stations distorted the smooth pattern of wave propagation. The corresponding StationXMLs were then manually examined and in most cases, we were able to identify the errors. To emphasize how a wrong deconvolution of each single station can affect the measurement, we removed 25 of these 26 suspicious stations. In the middle panel of Fig. 11, only 1 wrong station is left, close to the center of the AlpArray region. We investigated the StationXML of that particular station, found the error, asked the network operator for correcting the issue, deconvolved the raw data with the updated StationXML and measured the phase wavefront again to verify that the new version of the StationXML was correct. The right panel shows a clean map with the corrected StationXML used for the wrong station from the middle panel. The test using phases of the waves is capable to detect also polarity flips, which we indeed found for units of stations of the AlpArray network. The gradual loss of data in that example is as follows: We downloaded 550 data files. 15 stations did not pass the check of the vertical components according to Section 3.4, 4 stations were removed by the visual inspection, 5 stations were lost during the fundamental mode picking. Remaining 526 stations were used for the wavefront-tracking test. Initially, 26 did not pass it. Hence, out of 550 downloaded stations, after all the checks, we were left with 500 stations to be used for the surface-wave analysis, which represents a loss of 9.1%. Later, most of the records from those 26 stations suffering from the issues with StationXML were successfully deconvolved using the corrected StationXML.

A similar wavefront tracking test was done for the Taiwan earthquake recorded by the AdriaArray Seismic Network, examined already in Section 3.4 above. In this case, we picked the group velocity arrival times. The term “wavefront” is thus generalized here, because group velocity arrival times refer to the time of the envelope maximum and do not have a meaning of the wavefront, while phases shown in the AlpArray example (Fig. 11) do. Group velocity times are easier to pick, as there is only one maximum of the envelope at the fundamental mode surface wavetrain. In case of the phase picking, we need to correct for the cycle skips, which is more difficult for larger regions. While tracking the phase wavefront works efficiently for the AlpArray, it is more laborious for AdriaArray due to its larger extent. In Fig. 12, we show the results of group (energy) propagation measurement. On the other hand, group velocity suffers more by the interference effects brought by the wavefield from outside due to propagation in generally 3D structure between the epicenter and the network footprint (Kolínský et al., 2020). Hence, the shape of the group propagation times could be potentially more complex than the phase wavefronts even in case of an ideal measurement without any errors. Note also, that tracking the group travel times does not allow for polarity check, as the envelope remains the same even when the polarity is wrong. However, for the purpose of identifying the outliers caused by wrong PaZ, group propagation times are sufficient.

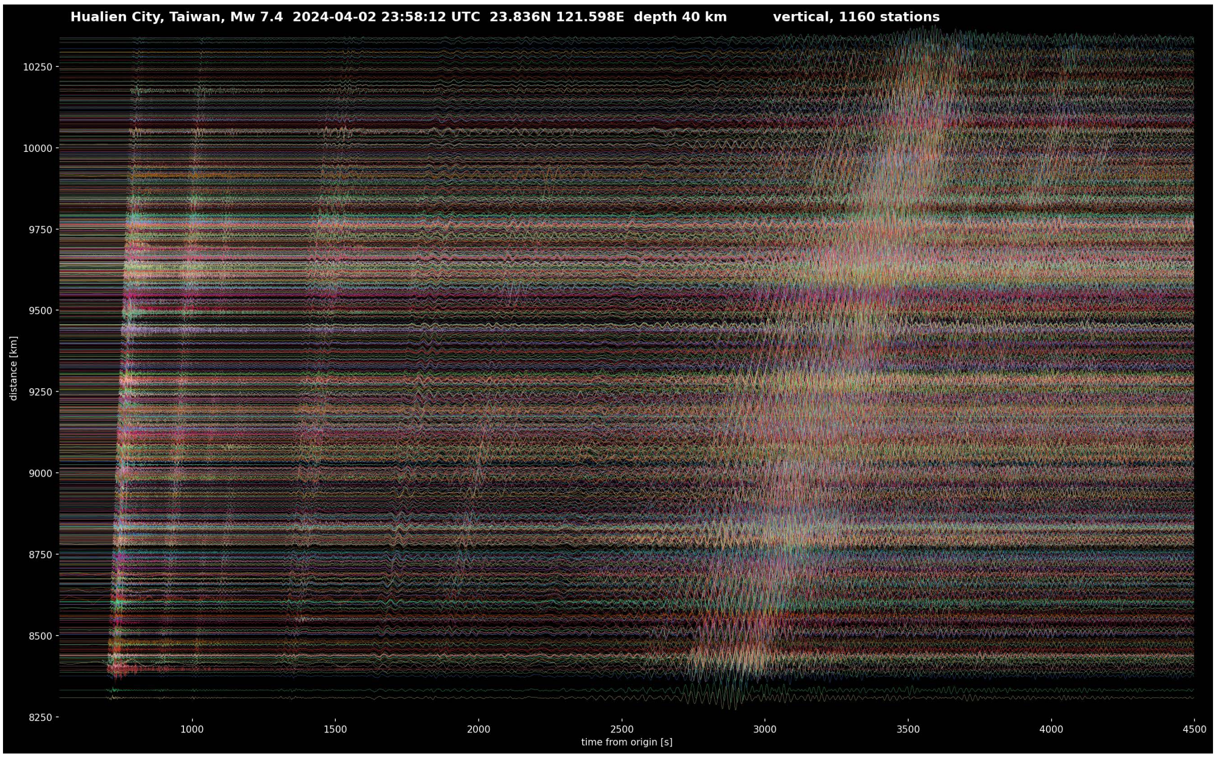

Figure 13. Records of vertical component, Taiwan M7.4 earthquake from 2024-04-02.

After removing 13 stations by visual inspection (non-seismic signals), we performed the group velocity measurement using the remaining 1160 stations (vertical component). Records of these stations are plotted in Fig. 13. The left map in Fig. 12 shows a generally coherently propagating wavefield, with many small-scale “bubbles”, similar to the left map in Fig. 11 – note the wavelength scale in the left map of Fig. 12. The contours are again plotted by 5 s of propagation time. Note, that here we use a shorter period of 90 s (compare it with 102 s used in Fig. 11). The wavelengths are related to the phase velocity calculated according to the PREM (Preliminary reference Earth model, Dziewonski and Anderson, 1981) in both cases. The group velocity is smaller than the phase velocity for the given period, and hence also the contours in Fig. 12 are denser.

Manually, we removed all stations (85) causing the distortion of the wavefield. The result is on the right map in Fig. 12. We still see stronger distortions compared to the phase wavefronts in Fig. 11, which could be caused by the higher sensitivity of the group velocity to the interference effects, as mentioned above. Beside apparently wrong StationXMLs, we identified other causes of the wrong group arrival times. Some records are shifted in time, suggesting a timing issue. There are also records stretched in time, suggesting a wrong sampling rate. The total loss from 1208 downloaded stations to the resulting 1075 stations usable for the surface-wave analysis, which are by green triangles in the right map of Fig. 12, is 11%, slightly more than the 9.1% of the AlpArray example above. There is still a high chance of correcting many issues and hence increasing the overall data usability.

Wavefront tracking could be provided for a broad frequency range. This helps to understand the nature of the metadata issues. In case the arrival time at a given station is offset from its neighboring stations by the same time difference across all frequencies, the likely cause is the wrong timing or incorrect sampling rate. In contrast, if the distortion varies with frequency, the most probable cause is incorrect PaZ values. Since the instrument transfer function is frequency dependent, errors in PaZ introduce phase shifts that vary across frequencies. As the phase of the instrument transfer function increases with period, longer waves are more affected. Therefore, we present here the wavefront tracking for rather long periods of 90 and 102 s in Figs. 11 and 12, because the distortion is well visible.

Preliminary use of the wavefront tracking for the AdriaArray data in 2022 and 2023 proved to be ineffective in terms of detecting outliers, as the number of stations with wrong (meta)data was comparable to the number of the correct ones, making it impossible to identify outliers. With a gradual increase of both the number of stations deployed and data available and many issues being corrected, the wavefront tracking test is able to detect outlying stations at the time of elaboration of this paper, January 2025.

Numbers of stations removed in every step of the tests described in Section 3.4 – Earthquake data quality checks and in this Section 3.5 – Wavefront tracking, are given in Fig. 14 for both the AlpArray and AdriaArray examples. Colors of pie chart portions correspond to Fig. 10. Charts show a total number of 618 operational stations for the Japan earthquake recorded by AlpArray (generalized area shown in Fig. 11) and 1462 operational stations for the Taiwan earthquake recorded by the AdriaArray. Percentages are given with respect to that number of operational stations. Diagonally and vertically hatched portions represent stations removed by the visual inspection and fundamental mode picking, and by the wavefront tracking, respectively. These hatched portions overlay the colored regions as wavefront tracking is applied to the green (no problem) and orange (issue fixed) stations. We see that although the right plot for AdriaArray encompasses more than two-times more stations than the left plot for AlpArray, the charts are similar. The useful number of stations reaches 80.9% and 73.5% for AlpArray and AdriaArray, respectively. Note that for AlpArray, data has been processed later with respect to the date of the earthquake (4 years after the earthquake). For AdriaArray, the percentage will get higher when all the backfilled offline data will be collected (we processed the data only less than a year from the earthquake). We see, that the biggest "loss" of data is due to stations which did not provide any data – see the white regions of the charts. The vertically hatched data has already been corrected for AlpArray after the test had been done. In case of AdriaArray, the correction is still expected. Once done, it would further increase the percentage of useful data for AdriaArray.

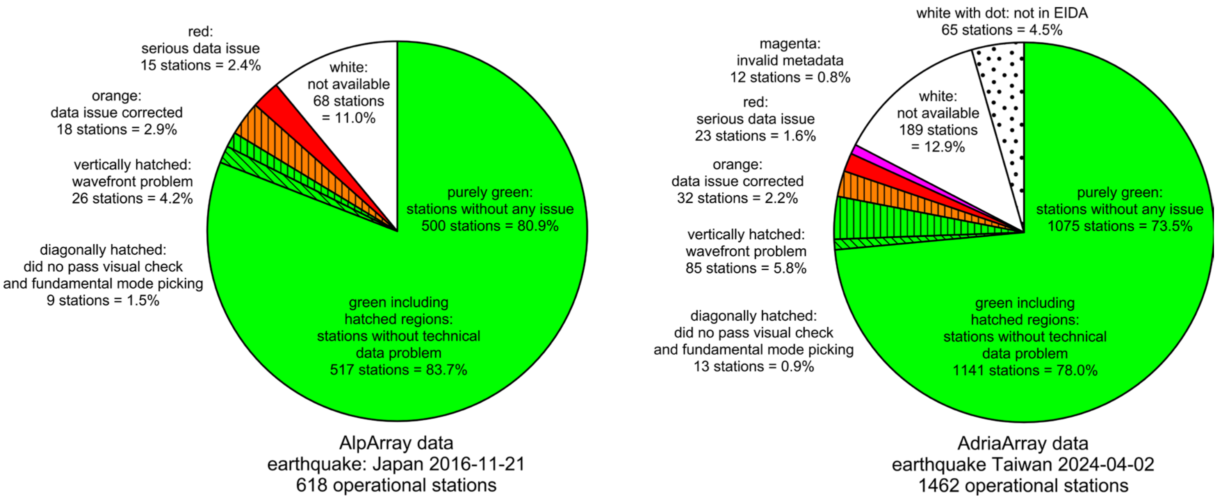

Figure 14. Numbers and percentages of stations removed during the earthquake data and wavefront tracking tests. Colors correspond to Fig. 10. All stations were in EIDA for the AlpArray case (left chart) and there were no invalid metadata found. Vertically and diagonally hatched portions overlay the orange and green regions, as these hatched stations were removed in the second test of wavefront tracking. Both charts relate to the vertical component analysis. We only show the operational stations to allow for comparison between AlpArray and AdriaArray.

4. Online repositories of the tests

4.1 Summary sheet

As mentioned before, we have been performing the tests described above since the beginning of the AdriaArray deployment in 2022. When communicating the results at the AdriaArray workshops and other meetings, the seismological community raised the question of whether and how to publish the results of these tests, so that both the station operators and researchers can easily access them. Publishing test results does not just mean sharing the maps, as shown in our paper here, which give an overview but are of limited use when it comes to checking a particular station. We needed an accessible and readable format, which would allow anyone to see at what time and at what station(s) a specific test identified a suspicious behavior, warning, or clear error. Different tests may focus on different time periods of the (meta)data. Some are conducted for a specific day or a few days, others for a longer time span. Different tests are performed using different codes and by different researchers. We developed a procedure that summarizes the results of all the tests in a single sheet. To make these results accessible, we set up a GitHub repository with an archive of these sheets. All the tests described in our paper (retrievability, noise levels, formal checks, corner periods, earthquake data and wavefront tracking) are summarized there, see https://github.com/PetrColinSky/DataQuality/tree/master/summary.

Every row of the sheet corresponds to one station. Every column then shows the result of a particular test. This sheet does not replace the detailed results of the individual tests, which may evaluate separate channels or more epochs, or provide more extensive output than can be implemented in a simple sheet. Instead, this sheet serves as a first look, pointing users and station/network operators to the particular test they need to look at to identify the issue at their station of interest. Whenever a new additional test is performed, the sheet is updated by adding that new test to the right side of the sheet. In case the same test is repeated, it is replaced in the sheet. In case the new test includes more stations, the respective rows are added to the sheet and kept empty for the previous tests. In GitHub, we keep the summary sheets with the dates of publishing and updated sheets are added with the new versions of the tests. If only one or several particular tests were newly performed or added, the sheet includes the updated or added versions of the new tests while previous versions of other tests are kept. This way, the most recent sheet will always show the latest results of all tests. The older versions of the sheet enable us to track the tests' history and to see how the number of issues decreases with time.

In the sheet, the retrievability tests are listed for every month of the operational period of AdriaArray from June 2022 to March 2024. Eight columns for each month correspond to the four parameters of the BH and HH channels respectively. These parameters are: WFCatalog availability, FDSNWS availability, metadata issue and retrievability. In every column, we give the decimal fraction (as an interval from 0 to 1) of the data available or retrieved. For easier navigation, the cells are colored. Cells with values below 60% are marked red, cells with values over 95% are green and those between are orange – similarly to the color-coding in Figs. 1, 2 and 3. Cells for stations which were not queried are empty, corresponding to the white triangles in the latter three figures.

The noise tests are summarized after the retrievability results on the right side of the sheet. Three columns show the noise levels for the three frequency bands (3 Hz, 5 s and 20 s), as presented in this paper. Instead of reporting in nanometers, we use simplified categories: mid-range "M" (green), low "L" and high "H" (orange), and above "A" and below "B" (red), based on the histograms in Fig. 7. The color coding of the cells exactly corresponds to the colors of the maps in the right panels of Figs. 4, 5 and 6.

The next column summarizes the StationXML formal checks. Similarly, as in the map in Fig. 8, we show the worst issue grade found for a given station, among all its components. Again, no problem and notification are colored in green (grades [-1] no problem, [0] and [1] notifications), grades with warnings ([2] and [3]) are orange, and serious issues (grades [4] and [5]) are red. The same grading and color scheme is also used for the other column summarizing the test of corner periods – see Fig. 9.

The last two columns of the sheet are for the earthquake data quality tests (Japan 2024-01-01 M7.5 earthquake and Taiwan 2024-04-02 M7.4 earthquake shown in this paper). In addition to the color coding, again the same as in Fig. 10, the cells also contain keywords pointing to the issue found. Besides the three traffic light colors, we also use magenta if a StationXML issue was detected. Suspicious stations revealed by the wavefront tracking test are shown in additional columns.

A newer version of the sheet (from 2024-06-10) is updated with four tests of the formal StationXML properties, performed at four different dates. These new tests are given at the end of the sheet as eight columns. Each two columns correspond to one date of performing the tests. The first of the two columns always shows the test of the epochs for only the AdriaArray time period since May 2022, and the second column shows the results for all the epochs for the given station. The corner period test is not shown separately as it is included in the summarizing result of all eight tests.

In general, for all tests, a white cell means the given parameter was not tested for the particular station, green means a positive result, red is negative, and orange is an intermediate issue or warning. The summary sheet contains all the tested stations. Besides the AdriaArray stations, we also included short-period, strong-motion, and broadband stations with corner periods shorter than 30 s, altogether more than 2500 stations. The figures in our paper only show the results for the AdriaArray Seismic Network (stations with 30 s and longer corner periods).

4.2 Repositories of individual tests

The summary sheet introduced in the previous section gives a simplified overview of the most serious issue for each station. For many stations, however, the individual channels and epochs were tested, and also several parameters were checked for example during the metadata formal consistency evaluation. A detailed description of all the tests discussed in our paper is available at respective GitHub sites.

Retrievability: The GitHub page at https://github.com/doukutsu/eida-data-monitoring contains the description of the retrievability procedure. The Python code for performing the download testing procedure and for creating the result maps is available there, along with some example maps. Directly related to our paper is the folder months/, where there are maps, similar to Figs. 1-3, of retrieved data month by month for the AdriaArray Seismic Network starting with June 2022 and – at this moment – ending with March 2024.

Noise levels: The GitHub page at https://github.com/felix-eckel/AdriaArrayQC contains the description of the noise level test, with some figures similar to those presented in our paper and also additional plots. The maps of noise levels show all the tested stations, while those presented in our paper only show the AdriaArray stations.

Metadata formal properties and corner periods: The GitHub page at https://github.com/PetrColinSky/DataQuality/tree/master/ludekvecsey contains detailed sheets with all the tested properties for all stations, channels and epochs for all the parameters given in Table 2 of this paper. There are also maps showing the distribution of stations with particular issues, similarly as done for the corner periods here in Fig. 9. The repository contains four versions of the tests performed in four different dates from February to June 2024.

Earthquake data check and wavefront tracking: The GitHub page at https://github.com/PetrColinSky/DataQuality/tree/master/petrkolinsky contains results of the earthquake data check together with maps for two earthquakes, similar to Fig. 10. There are also plots of the records of all three components for both earthquakes, similar to Fig. 13. Here as well as in the summary sheet, the respective issues are specified by keywords, pointing to the cause of the issue, if known. Results of the wavefront tracking are included in that repository too.

These individual tests as well as the summary sheet were introduced at the EGU poster, see Kolínský et al. (2024) available at the GitHub repository (link in the next Section). In addition, we also posted a complementary issue at the EIDA issue tracker (links also below). Some of the issues mentioned in the issue tracker can overlap with those marked in the particular tests, some are given in addition.

5. Links related to data availability and quality

Summary of the tests: https://github.com/PetrColinSky/DataQuality/tree/master/summary
Individual test repositories:
Retrievability: https://github.com/doukutsu/eida-data-monitoring
Noise levels: https://github.com/felix-eckel/AdriaArrayQC
Metadata formal checks: https://github.com/PetrColinSky/DataQuality/tree/master/ludekvecsey
Earthquakes and wavefronts: https://github.com/PetrColinSky/DataQuality/tree/master/petrkolinsky
Presentations on data quality 2023-2025:
https://github.com/PetrColinSky/DataQuality/tree/master/presentations
EGU 2024 poster introducing the summary table (Kolínský et al., 2024):
https://github.com/PetrColinSky/DataQuality/blob/master/presentations/2024_Kolinsky_EGU_poster.pdf
AdriaArray webpage: https://orfeus.readthedocs.io/en/latest/adria_array_main.html
ORFEUS main page: https://orfeus-eu.org/
EIDA webpage: https://www.orfeus-eu.org/data/eida/
List of the AdriaArray temporary networks, FDSN webpage: https://www.fdsn.org/networks/?search=adriaarray
FDSN StationXML schema: https://www.fdsn.org/xml/station/
FDSN StationXML validation schema: https://xmlschema.readthedocs.io/en/latest/usage.html#validation
IRIS StationXML validator: https://github.com/iris-edu/StationXML-Validator/wiki

with a list of the tests: https://github.com/iris-edu/stationxml-validator/wiki/StationXML-Validation-Rule-List
FDSN availability webservice specification: http://www.fdsn.org/webservices/fdsnws-availability-1.0.pdf
EIDA test report: https://www.szgrf.bgr.de/eidaqc_report/
EIDA issue tracker: https://github.com/EIDA/userfeedback/issues
Announcement of the summary sheet at the issue tracker: https://github.com/EIDA/userfeedback/issues/166
a10y availability application: https://github.com/EIDA/a10y/releases/tag/2024.040
AutoStatsQ GitHub page: https://github.com/gesape/AutoStatsQ
ORFEUS User Advisory Group reports: https://polybox.ethz.ch/index.php/s/ApZ875YaohovoOj

6. Summary

Regular checks of the availability and retrievability of data, of the seismic noise in the recordings, and of the metadata correctness are essential to ensure the effective usage of large data sets like the one obtained by AdriaArray. Results of these checks are regularly updated and available at GitHub. We encourage users to explore results of these tests and to develop them further. Anyone is welcome to contribute as any additional test could be easily added to the summary sheet. Station and network operators are encouraged to curate in particular the metadata in case of detected issues. We note that the data quality checks were done in close cooperation with the ORFEUS User Advisory Group (UAG) established in 2018. UAG contributed considerably to the stabilization and improvement of EIDA services over the last years.

Retrievability tests evaluate which data users can download for particular time epochs, benchmarking the chain of steps between network operators, EIDA and the end user. Noise level maps identify stations with outlying gains. Checks of StationXML formal properties point to both warnings and serious issues that prevent proper transfer function deconvolution. Additionally, earthquake data quality tests and wavefront tracking reveal potential issues with wrong PaZ. All these tests are introduced and demonstrated using examples for a particular time period and for earthquakes recorded by the AdriaArray Seismic Network. This paper presents methods which can be used for large dense networks of passive seismic experiments beyond AdriaArray, offering insights and tools to improve the quality and reliability of data. We do not point to individual issues at particular stations, because we tested more than 2500 stations, we repeat the tests regularly and their results will be updated. Instead, we introduce an online GitHub repository with the test results, where users as well as the stations and network operators can find the issues for their stations. Each presented test has its own repository with detailed results. Thanks to our efforts in testing the data and metadata since the beginning of the AdriaArray deployment, issues at tens of stations have been identified and already corrected. However, as our examples show, as of February 2025, there are still additional tens of stations within the AdriaArray region that require an operator's attention.

The importance of testing data and metadata quality has already been shown in the case of the AlpArray Seismic Network (2016-2019). Here, we also show one example of the wavefront tracking method based on AlpArray data from 2016. All proposed methods can be generally applied to any dense network. Our aim is to continue with these tests for the AdriaArray region over the next few years and expand the region to include all stations within the ORFEUS EIDA archive across the whole Europe and Mediterranean.

We found that most of the issues are related to wrong metadata information, which can be easily corrected. Newer, mainly temporary stations, have their metadata freshly updated and usually contain less errors. Older, mainly permanent stations, might have undergone hardware upgrades that are not reflected in the metadata. The set of StationXMLs will never be completely error-free, as new stations are being built and older ones are being upgraded continuously. Sometimes, the same error in metadata is found for more stations of a particular network. The more data is recorded correctly and deconvolved using accurate metadata, the better and easier it will be to identify the outliers by multi-station methods applied to large dense networks.

As there are many more data quality tests than described in our paper (see Sections 1 and 2), we expect that researchers will perform their own tests before analyzing seismic data. The users are encouraged to report issues to the EIDA issue tracker (see the link above). Our paper is intended to serve as an overview of possible approaches. In a similar sense, our online test repository is intended as a guide for the community to understand what can be expected when working with the AdriaArray dataset. Our ultimate goal is to contribute to the overall better quality of the AdriaArray Seismic Network, which only then will allow us to answer the open research questions about the structure, evolution, seismicity, geohazards and geodynamics of the Adriatic plate.

Data availability statement. Waveform data from all AdriaArray stations is available through ORFEUS EIDA. Data from the permanent stations and from temporary stations with network codes 4P, 7B, Y5 and XP is publicly accessible immediately. Data from temporary stations with network codes 1Y, 2Y, 9H, Y8 and Z6 is accessible to the AdriaArray Seismology Group participants. The rolling embargo allows this data to be publicly available two years after its acquisition. However, data from all AdriaArray temporary stations is immediately available for seismological observatories with monitoring and alerting duties within the AdriaArray region.

Acknowledgements. PK and LV are supported by the Czech Science Foundation grant No. 23-06370S. JS was a recipient of the 2022 ORFEUS software development grant which supported the improvement of the availability/retrievability tests described in this paper. TBT was supported by the Croatian Science Foundation under Project No. IP-2020-02-3960. The communication between data users, stations operators and ORFEUS is facilitated by the ORFEUS User Advisory Group (UAG). PK, LV and FE were members of the UAG when working on the tests and writing this paper. TM is a former UAG member and its first chair (2018-2020). Later, the UAG was chaired by Stéphane Rondenay (2020-2022) and PK (2022-2025). Data quality control was always on the agenda of the UAG and we appreciate the voluntary contributions of all former UAG members since its foundation in 2018.
A detailed overview of many tests is given in the ORFEUS UAG 2023 report, https://polybox.ethz.ch/index.php/s/ApZ875YaohovoOj/download?path=%2F&files=UAG_R4.pdf.
Earlier versions of the availability/retrievability tests were supported and performed by Florian Fuchs, Javier Quinteros and Máté Timkó. Within the AdriaArray initiative, data quality control is managed by the Working Group 3, see Kolínský et al. (2025).
Maps were plotted using Generic Mapping Tools (GMT), Wessel et al. (2013). The Python Toolbox ObsPy by Beyreuther et al. (2010) and Megies et al. (2011) was used for data and metadata downloading and pre-processing. Earthquake parameters were taken from the U. S. Geological Survey Earthquake Lists, Maps, and Statistics, https://www.usgs.gov/natural-hazards/earthquake-hazards/lists-maps-and-statistics.
The altitude and bathymetry data were plotted using the ETOPO1 Global Relief Model provided by the NOAA Physical Sciences Laboratory, Boulder, Colorado, USA, from their website at https://www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/data/bedrock/grid_registered/netcdf/, see also NOAA (2009) and Amante and Eakins (2009). The tested AdriaArray Seismic Network stations and the virtual _ADARRAY network are composed of the following networks, listed in FDSN: **1Y** – AdriaArray Temporary Network: Greece, North Macedonia (Friederich et al., 2022). **2Y** – AdriaArray Temporary Network: Italy – northeast (Pesaresi and Rossi, 2022). **4P** – AdriaArray Temporary Network: Italy – north, south. **7B** – AdriaArray Temporary Network: Austria, Croatia, Slovakia. **9H** – CRONOS – Croatia/Norway Contribution to AdriaArray Temporary Network. **AC** – Albanian Seismological Network (Institute of GeoSciences, Polytechnic University of Tirana, 2002). **BS** – National Seismic Network of Bulgaria (National Institute of Geophysics, Geodesy and Geography – BAS, 1980). **BW** – BayernNetz (Department of Earth and Environmental Sciences, Geophysical Observatory, University of Munchen, 2001). **C4** – CERN Seismic Network (CERN, 2016). **CH** – National Seismic Networks of Switzerland (Swiss Seismological Service (SED) at ETH Zurich, 1983). **CL** – Corinth Rift Laboratory Seismological Network (Corinth Rift Laboratory Team And RESIF Datacenter, 2013). **CR** – Croatian Seismograph Network (University of Zagreb, 2001). **CZ** – Czech Regional Seismic Network (Charles University in Prague, Institute of Geonics, Institute of Geophysics, Academy of Sciences of the Czech Republic, Institute of Physics of the Earth Masaryk University and Institute of Rock Structure and Mechanics, 1973). **FR** – Epos-France Broad-band network (RLBP), (Epos-France Seismology, 1995). **G** – GEOSCOPE, French Global Network of broad band seismic stations (Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre de Strasbourg (EOST), 1982). **GE** – GEOFON Seismic Network (GEOFON Data Centre, 1993). **GR** – German Regional Seismic Network (GRSN) (Federal Institute for Geosciences and Natural Resources, 1976). **GU** – Regional Seismic Network of North Western Italy (University of Genoa, 1967). **GX** – GFZ Affiliated Stations, Deutsches GeoForschungsZentrum GFZ (GFZ Potsdam), Germany. **HA** – Hellenic Seismological Network (University of Athens, 2008). **HC** – Seismological Network of Crete (Technological Educational Institute of Crete, 2006). **HL** – National Observatory of Athens Seismic Network (National Observatory of Athens, Institute of Geodynamics, Athens, 1975). **HP** – University of Patras, Seismological Laboratory (University of Patras, 2000). **HT** – Aristotle University of Thessaloniki Seismological Network (Aristotle University of Thessaloniki, 1981). **HU** – Hungarian National Seismological Network (Kövesligethy Radó Seismological Observatory (Geodetic And Geophysical Institute, Research Centre For Astronomy And Earth Sciences, Hungarian Academy Of Sciences (MTA CSFK GGI KRSZO), 1992). **IV** – Rete Sismica Nazionale (RSN) (Istituto Nazionale di Geofisica e Vulcanologia (INGV), 2005). **IX** – Irpinia Seismic Network (ISNet). **IY** – Rete Sismica Unical (Universita Della Calabria, Italy, 1981). **KO** – Kandilli Observatory And Earthquake Research Institute (KOERI) (Kandilli Observatory

And Earthquake Research Institute, Boğaziçi University, 1971). **LE** – Erdbebendienst Südwest (Erdbebendienst Südwest Baden-Württemberg and Rheinland-Pfalz, 2009). **MD** – Moldova Digital Seismic Network (Geological and Seismological Institute of Moldova, 2007). **ME** – Montenegrin Seismic Network (Sector for Seismology, Institute of Hydrometeorology and Seismology of Montenegro, 1982). **MK** – Seismological network of the Republic of North Macedonia (Seismological Observatory at the Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, Skopje, Republic of Macedonia, 1966). **ML** – Malta Seismic Network (University of Malta, 2014; Agius et al., 2025). **MN** – Mediterranean Very Broadband Seismographic Network (MedNet) (MedNet Project Partner Institutions, 1990). **MT** – Observatoire Multi-disciplinaire des Instabilités de Versants (OMIV) (French Landslide Observatory – Seismological Datacenter/RESIF, 2006). **NI** – North-East Italy Broadband Network (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS and University of Trieste, 2002). **OE** – Austrian Seismic Network (ZAMG – Zentralanstalt für Meteorologie und Geodynamik, 1987). **OT** – OTRIONS (University of Bari "Aldo Moro", 2013). **OX** – North-East Italy Seismic Network (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, 2016). **PL** – Polish Seismological Network (Institute of Geophysics, Polish Academy of Sciences, 1990). **RD** – CEA/DASE broad-band permanent network in metropolitan France (RESIF, 2018). **RF** – Friuli Venezia Giulia Accelerometric Network (University of Trieste, 1993). **RO** – Romanian Seismic Network (National Institute for Earth Physics (NIEP Romania), 1994). **SI** – Province Südtirol, ZAMG – Central Institute for Meteorology and Geodynamics, Austria. **SJ** – Serbian Seismological Network (Seismological Survey of Serbia, 1906). **SK** – National Network of Seismic Stations of Slovakia (ESI SAS; Former GPI SAS (Geophysical Institute Of The Slovak Academy Of Sciences), 2004). **SL** – Seismic Network of the Republic of Slovenia (Slovenian Environment Agency, 1990). **ST** – Trentino Seismic Network (Geological Survey-Provincia Autonoma di Trento, 1981). **TV** – INGV experiments network, Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italy. **UT** – Ukrainian National Seismic Network (Subbotin Institute, 2023). **VM** – Seismic Data acquired by Marche Seismic Network (MSN) (Istituto Nazionale di Geofisica e Vulcanologia (INGV), 2023). **VR** – Virgo Interferometric Antenna for Gravitational Waves Detection (European Gravitational Observatory, 2019). **XP** – MACIV-BB temporary experiment to carry out a seismic tomography of the crustal and mantle structures of the Massif Central, France (RESIF-SISMOB) (Paul et al., 2023). **Y5** – Swiss Contribution to AdriaArray Temporary Network (Obermann et al., 2022). **Y8** – AdriaArray Temporary Network: Bulgaria, Moldova, Poland, Romania, Ukraine (Neagoe, 2022). **Z6** – AdriaArray Temporary Network: Albania, Austria, Czech Rep., Germany, Hungary, Kosovo, Montenegro, Slovakia (Schlömer et al., 2022b).

Author contributions:
Petr Kolínský – conceptualization, writing, editing, earthquake data test, wavefront tracking test.
Johannes Stampa – retrievability test, writing, editing.
Luděk Vecsey – metadata checks, corner period tests, writing, editing.
Felix Eckel – noise level tests, writing, editing.
Tena Belinić Topić – coding, summary sheet, preparing data for figures, editing.
Thomas Meier – conceptualization, organization, writing, editing, supervision.

References

AdriaArray Temporary Network: Italy – north, south, Data set, network code 4P.

AdriaArray Temporary Network: Austria, Croatia, Slovakia, Data set, network code 7B.

AlpArray Seismic Network (2015). AlpArray Seismic Network (AASN) temporary component, AlpArray Working Group, doi:10.12686/alparray/z3_2015.

Amante, C. and B. W. Eakins (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-24, National Geophysical Data Center, NOAA, doi:10.7289/V5C8276M.

Anthony, R. E., A. T. Ringler and D. C. Wilson (2022). Seismic Background Noise Levels across the Continental United States from USArray Transportable Array: The Influence of Geology and Geography, Bull. Seism. Soc. Am., 112, 2, 646-668, doi:10.1785/0120210176.

Aristotle University of Thessaloniki (1981). Aristotle University of Thessaloniki Seismological Network, Data set, FDSN, doi:10.7914/SN/HT.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies et al. (2010). ObsPy: a Python toolbox for seismology, Seismol. Res. Lett., 81, 3, 530-533, doi:10.1785/gssrl.81.3.530.

Borleanu, F., L. Petrescu, C. Neagoe, H. Kampfová Exnerová et al. (2025). Romania-AdriaArray temporary broadband seismic stations deployment and data quality control, Ann. Geophys., 68, this issue, doi:10.4401/ag-9318.

Brune, J. N. and J. Oliver (1959). The seismic noise of the earth's surface, Bull. Seism. Soc. Am., 49, 4, 349-353, doi:10.1785/BSSA0490040349.

Busby, R. W., R. L. Woodward, K. A. Hafner, F. L. Vernon et al. (2018). The Design and Implementation of EarthScope's USArray Transportable Array in the Conterminous United States and Southern Canada, EarthScope report, 64.

Butler, R., T. Lay, K. Creager, P. Earl et al. (2004). The global seismographic network surpasses its design goal, Eos Trans., AGU, 85, 23, 225-229, doi:10.1029/2004EO230001.

Cauzzi, C., D. Bindi, D. Cambaz, F. Carrilho et al. (2022). Systematic comparisons of broad-band velocity and acceleration earthquake records as a quality assessment tool for European open strong-motion data, Seismol. Res. Lett., 93, 2B, 1297, doi:10.1785/0220220087.

Cauzzi, C., J. Clinton, W. Crawford, S. Custódio et al. (2024). Status and Outlook of ORFEUS Data Services, Products and Activities to Coordinate Access to Seismic Waveform Data in the Euro-Mediterranean Region and Beyond, EGU General Assembly 2024, Vienna, Austria, EGU24-3477, doi:10.5194/egusphere-egu24-3477.

CERN (2016). CERN Seismic Network, ETH Zurich, doi:10.12686/sed/networks/c4.

Charles University in Prague-Czech, Institute of Geonics, Institute of Geophysics, Institute of Physics of the Earth Masaryk University-Czech et al. (1973). Czech Regional Seismic Network, Data set, FDSN, doi:10.7914/SN/CZ.

Corinth Rift Laboratory Team and RESIF Datacenter (2013). CL – Corinth Rift Laboratory Seismological Network, CRLNET, Data set, RESIF – Réseau Sismologique et géodésique Français, doi:10.15778/RESIF.CL.

CRONOS, Croatia/Norway Contribution to AdriaArray Temporary Network, CRONOS, Data set, network code 9H.

Davis, P., M. Ishii and G. Masters (2005). An Assessment of the Accuracy of GSN Sensor Response Information, Seism. Res. Lett., 76, 6, doi:10.1785/gssrl.76.6.678.

Davis, P. and J. Berger (2012). Initial Impact of the Global Seismographic Network Quality Initiative on Metadata Accuracy, Seism. Res. Lett., 83, 4, 697-703, doi:10.1785/0220120021.

Department of Earth and Environmental Sciences, Geophysical Observatory, University of Munchen (2001). BayernNetz, Data set, FDSN, doi:10.7914/SN/BW.

Díaz, J., A. Villaseñor, J. Morales, the TopoIberia Seismic Working Group et al. (2010). Background Noise Characteristics at the IberArray Broadband Seismic Network, Bull. Seism. Soc. Am., 100, 2, 618-628, doi:10.1785/0120090085.

Dziewonski, A. M. and D. L. Anderson (1981). Preliminary reference Earth model, Phys. Earth Planet. Inter., 25, 4, 297-356, doi:10.1016/0031-9201(81)90046-7.

Ekström, G., C. A. Dalton and M. Nettles (2006). Observations of Time-dependent Errors in Long-period Instrument Gain at Global Seismic Stations, Seism. Res. Lett., 77, 1, 12-22, doi:10.1785/gssrl.77.1.12.

Ekström, G. and R. W. Busby (2008). Measurements of Seismometer Orientation at USArray Transportable Array and Backbone Stations, Seism. Res. Lett., 79, 4, 554-561, doi:10.1785/gssrl.79.4.554.

Epos-France Seismology (1995). Epos-France Broad-band network, RLBP, Data set, Epos-France Seismological Data Centre, doi:10.15778/RESIF.FR.

Erdbebendienst Südwest Baden-Württemberg und Rheinland-Pfalz (2009). Erdbebendienst Südwest, Data set, FDSN, doi:10.7914/SN/LE.

ESI SAS, Former GPI SAS, Geophysical Institute Of The Slovak Academy Of Sciences (2004). National Network of Seismic Stations of Slovakia, Data set, GFZ Data Services, doi:10.14470/FX099882.

European Gravitational Observatory (2019). Virgo Interferometric Antenna for Gravitational Waves Detection, Data set, FDSN, doi:10.7914/SN/VR.

Federal Institute for Geosciences and Natural Resources (1976). German Regional Seismic Network, GRSN, Bundesanstalt für Geowissenschaften und Rohstoffe, doi:10.25928/mbx6-hr74.

Federation of Digital Seismograph Networks-FDSN, Incorporated Research Institutions for Seismology-IRIS, United States Geological Survey-USGS (2012). SEED Reference Manual, Version 2.4, http://www.fdsn.org/seed_manual/SEEDManual_V2.4.pdf.

French Landslide Observatory, Seismological Datacenter, RESIF (2006). Observatoire Multi-disciplinaire des Instabilités de Versants, OMIV, Data set, RESIF, Réseau Sismologique et géodésique Français, doi:10.15778/RESIF.MT.

Friederich, W., C. Evangelidis, C. Papazachos, E. Sokos et al. (2022). AdriaArray Temporary Network: Greece, North Macedonia, Data set, FDSN, doi:10.7914/y0t2-3b67.

Fuchs, F., P. Kolínský, G. Gröschl, M. T. Apoloner et al. (2015). Site selection for a countrywide temporary network in Austria: noise analysis and preliminary performance, Adv. Geosci., 41, 25-33, doi:10.5194/adgeo-41-25-2015.

Fuchs, F., P. Kolínský, G. Gröschl, the AlpArray Working Group et al. (2016). AlpArray in Austria and Slovakia: technical realization, site description and noise characterization, Adv. Geosci., 43, 1-13, doi:10.5194/adgeo-43-1-2016.

GEOFON Data Centre (1993). GEOFON Seismic Network, Data set, Deutsches GeoForschungsZentrum GFZ, doi:10.14470/TR560404.

Geological and Seismological Institute of Moldova (2007). Moldova Digital Seismic Network, Data set, FDSN, doi:10.7914/SN/MD.

Geological Survey, Provincia Autonoma di Trento (1981). Trentino Seismic Network, Data set, FDSN, doi:10.7914/SN/ST.

GFZ Affiliated Stations, Deutsches GeoForschungsZentrum GFZ, GFZ Potsdam, Germany, Data set, network code GX.

Gibbons, S. J. (2006). On the Identification and Documentation of Timing Errors: An Example at the KBS Station, Spitsbergen, Seism. Res. Lett., 77, 5, 559-571, doi:10.1785/gssrl.77.5.559.

Gorbatov, A., K. Czarnota, B. Hejrani, M. Haynes et al. (2020). AusArray: quality passive seismic data to underpin updatable national velocity models of the lithosphere, in Exploring for the Future: Extended Abstracts K. Czarnota, I. Roach, S. Abbott, M. Haynes et al. (Eds.), Geoscience Australia, Canberra, 1-4, doi:10.11636/135284.

Gorbatov, A., B. Hejrani, J. Holzschuh, J. Zhao et al. (2024). AusArray continent-scale deployment, in Exploring for the Future: Extended Abstracts K. Czarnota, I. Roach, S. Abbott, M. Haynes et al. (Eds.), Geoscience Australia, Canberra, doi:10.26186/149640.

Govoni, A., L. Bonatto, M. Capello, the AlpArray Working Group et al. (2017). AlpArray-Italy: site description and noise characterization, Adv. Geosci., 43, 39-52, doi:10.5194/adgeo-43-39-2017.

Gráczer, Z., G. Szanyi, I. Bondár, C. Czanik et al. (2018). AlpArray in Hungary: temporary and permanent seismological networks in the transition zone between the Eastern Alps and the Pannonian basin, Acta Geod. Geoph., 53, 221-245, doi:10.1007/s40328-018-0213-4.

Heit, B., L. Cristiano, C. Haberland, F. Tilmann et al. (2021). The SWATH-D Seismological Network in the Eastern Alps, Seismol. Res. Lett., 92, 1592-1609, doi:10.1785/0220200377.

Hetényi, G., the AlpArray Seismic Network Team, the AlpArray OBS Cruise Crew, the AlpArray Working Group et al. (2018a). The AlpArray Seismic network: a large-scale European experiment to image the Alpine Orogen, Surv. Geophys., 39, 1009-1033, doi:10.1007/s10712-018-9472-4.

Hetényi, G., J. Plomerová, M. Bielik, G. Bokelmann et al. (2019). Pannonian-Carpathian-Alpine Seismic Experiment, Data set, FDSN, doi:10.7914/SN/ZJ_2019.

INGV experiments network, Istituto Nazionale di Geofisica e Vulcanologia, INGV, Italy, Data set, network code TV.

Institut de physique du globe de Paris – IPGP and École et Observatoire des Sciences de la Terre de Strasbourg-EOST (1982). GEOSCOPE, French Global Network of broad band seismic stations, Institut de physique du globe de Paris-IPGP, Université de Paris, doi:10.18715/GEOSCOPE.G.

Institute of Geophysics, Polish Academy of Sciences (1990). Polish Seismological Network, Data set, International Federation of Digital Seismograph Networks, doi:10.7914/90rh-0q80.

Institute of GeoSciences, Polytechnic University of Tirana (2002). Albanian Seismological Network, Data set, FDSN, doi:10.7914/SN/AC.

IRIS (2023). IRIS StationXML Validator, Version 1.7.5, Software, Incorporated Research Institutions for Seismology, http://iris-edu.github.io/stationxml-validator/.

Irpinia Seismic Network – ISNet, Data set, network code IX.

Istituto Nazionale di Geofisica e Vulcanologia – INGV (2005). Rete Sismica Nazionale-RSN, Data set, Istituto Nazionale di Geofisica e Vulcanologia, doi:10.13127/sd/x0fxnh7qfy.

Istituto Nazionale di Geofisica e Vulcanologia – INGV (2023). Seismic Data acquired by Marche Seismic Network-MSN, Data set, Istituto Nazionale di Geofisica e Vulcanologia, doi:10.13127/sd/z7hoi9u3ix.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS (2016). North-East Italy Seismic Network, Data set, FDSN, doi:10.7914/SN/OX.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS and University of Trieste (2002). North-East Italy Broadband Network, Data set, FDSN, doi:10.7914/SN/NI.

Juretzek, C. and C. Hadziioannou (2017). Linking source region and ocean wave parameters with the observed primary microseismic noise, Geophys. J. Int., 211, 3, 1640-1654, doi:10.1093/gji/ggx388.

Kampfová Exnerová, H., L. Dimitrova, T. Nagel, L. Dimova et al. (2025). AdriaArray experiment on the territory of Bulgaria, Ann. Geophys., 68, this issue, doi:10.4401/ag-9295.

Kandilli Observatory And Earthquake Research Institute, Boğaziçi University (1971). Kandilli Observatory and Earthquake Research Institute-KOERI, Data set, FDSN, doi:10.7914/SN/KO.

Kolínský, P., G. Bokelmann and the AlpArray Working Group (2019). Arrival angles of teleseismic fundamental mode Rayleigh waves across the AlpArray, Geophys. J. Int., 218, 114-144, doi:10.1093/gji/ggz081.

Kolínský, P., F. M. Schneider and G. Bokelmann (2020). Surface wave diffraction pattern recorded on AlpArray: Cameroon volcanic line case study, J. Geophys. Res., Solid Earth, 125, e2019JB019102, doi:10.1029/2019JB019102.

Kolínský, P., L. Vecsey, the AlpArray Working Group, the AdriaArray Seismology Group et al. (2024). Data quality of large dense seismic networks – lessons learnt from AlpArray and application to AdriaArray, EGU General Assembly 2024, Vienna, Austria, EGU24-12520, https://meetingorganizer.copernicus.org/EGU24/EGU24-12520.html.

Kolínský, P., T. Meier, M. R. Agius, the AdriaArray Seismology Group et al. (2025). AdriaArray – a Passive Seismic Experiment to Study Structure, Geodynamics and Geohazards of the Adriatic Plate, Ann. Geophys., 68, this issue, doi:10.4401/ag-9284.

Kövesligethy Radó Seismological Observatory, Geodetic And Geophysical Institute, Research Centre For Astronomy and Earth Sciences, Hungarian Academy Of Sciences – MTA CSFK GGI KRSZO (1992). Hungarian National Seismological Network, Data set, GFZ Data Services, doi:10.14470/UH028726.

Larson, E. W. F. and G. Ekström (2002). Determining surface wave arrival angle anomalies, J. Geophys. Res., 107, B6, 2127, doi:10.1029/2000JB000048.

Lu, Y., H. A. Pedersen, L. Stehly and the AlpArray Working Group (2022). Mapping the seismic noise field in Europe: spatio-temporal variations in wavefield composition and noise source contributions, Geophys. J. Int., 228, 1, 171-192, doi:10.1093/gji/ggab273.

Lynch, W. A. (1938). Traffic and other local disturbances registered at Fordham by the vertical Benioff seismometer, Bull. Seism. Soc. Am., 28, 3, 217-225, doi:10.1785/BSSA0280030217.

Megies, T., M. Beyreuther, R. Barsch, L. Krischer et al. (2011). ObsPy – What can it do for data centers and observatories?, Ann. Geophys., 54, 1, 47-58, doi:10.4401/ag-4838.

Meltzer, A., R. Carlson, T. Dixon, G. Ekström et al. (1999). USArray – A Synoptic Investigation of the Structure, Dynamics, and Evolution of the North American Continent, White Paper, 34.

Molinari, I., J. Clinton, the Swiss-AlpArray Field Team, the AlpArray Working Group et al. (2016). Swiss-AlpArray temporary broadband seismic stations deployment and noise characterization, Adv. Geosci., 43, 15-29, doi:10.5194/adgeo-43-15-2016.

McNamara, D. E. and R. P. Buland (2004). Ambient Noise Levels in the Continental United States, Bull. Seism. Soc. Am., 94, 4, 1517-1527, doi:10.1785/012003001.

McNamara, D. E. and R. I. Boaz (2005). Seismic Noise Analysis System Using Power Spectral Density Probability Density Functions – A Stand-Alone Software Package, U. S. Geol. Surv. Open File Rep., 2005-1438, 29.

MedNet Project Partner Institutions (1990). Mediterranean Very Broadband Seismographic Network, MedNet, Data set, Istituto Nazionale di Geofisica e Vulcanologia – INGV, doi:10.13127/sd/fbbbtdtd6q.

NOAA National Geophysical Data Center (2009). ETOPO1 1 Arc-Minute Global Relief Model, NOAA National Centers for Environmental Information, accessed in February 2013.

National Institute for Earth Physics – NIEP Romania (1994). Romanian Seismic Network, Data set, FDSN, doi:10.7914/SN/RO.

National Institute of Geophysics, Geodesy and Geography – BAS (1980). National Seismic Network of Bulgaria, Data set, FDSN, doi:10.7914/SN/BS.

National Observatory of Athens, Institute of Geodynamics, Athens (1975). National Observatory of Athens Seismic Network, Data set, FDSN, doi:10.7914/SN/HL.

Neagoe, C. (2022). AdriaArray Temporary Network: Bulgaria, Moldova, Poland, Romania, Ukraine, Data set, FDSN, doi:10.7914/b1sc-0n71.

Obermann, A., D. Jozinović, S. Cvijić and Swiss Seismological Service-SED at ETH Zurich (2022). Swiss Contribution to AdriaArray Temporary Network, ETH Zurich, doi:10.12686/SED/NETWORKS/Y5.

Paul, A., A. Mordret, C. Aubert, RESIF et al. (2023). MACIV-BB (backbone) temporary broadband experiment in the French Massif Central, France, RESIF-SISMOB, Data set, RESIF – Réseau Sismologique et géodésique Français, doi:10.15778/RESIF.XP2023.

Pesaresi, D. and G. Rossi (2022). AdriaArray Temporary Network: Italy – northeast, Data set, FDSN, doi:10.7914/1p36-6t87.

Petersen, G. M., S. Cesca, M. Kriegerowski and the AlpArray Working Group (2019). Automated quality control for large seismic networks: Implementation and application to the AlpArray Seismic Network, Seismol. Res. Lett., 90, 3, 1177-1190, doi:10.1785/0220180342.

Peterson, J. (1993). Observations and modeling of seismic background noise, U. S. Geol. Surv. Open File Rep., 93-322, 94, doi:10.3133/ofr93322.

Province Südtirol, ZAMG – Central Institute for Meteorology and Geodynamics, Austria, Data set, network code SI.

RESIF (2018). CEA/DASE broad-band permanent network in metropolitan France, Data set, RESIF – Réseau Sismologique et géodésique Français, doi:10.15778/RESIF.RD.

Ranasinghe, N. R., A. C. Gallegos, A. R. Trujillo, A. R. Blanchette et al. (2015). Lg attenuation in northeast China using NECESSArray data, Geophys. J. Int., 200, 1, 67-76, doi:10.1093/gji/ggu375.

Ringler, A. T., L. S. Gee, C. R. Hutt and D. E. McNamara (2010). Temporal Variations in Global Seismic Station Ambient Noise Power Levels, Seismol. Res. Lett., 81, 4, 605-613, doi:10.1785/gssrl.81.4.605.

Ringler, A. T., C. R. Hutt, K. Persefield and L. S. Gee (2013). Seismic Station Installation Orientation Errors at ANSS and IRIS/USGS Stations, Seismol. Res. Lett., 84, 6, 926-931, doi:10.1785/0220130072.

Ringler, A. T., T. Storm, L. S. Gee, C. R. Hutt et al. (2015). Uncertainty estimates in broadband seismometer sensitivities using microseisms, J. Seismol., 19, 317-327, doi:10.1007/s10950-014-9467-7.

Rueda, J. and J. Mezcua (2015). Orientation Analysis of the Spanish Broadband National Network Using Rayleigh-Wave Polarization, Seismol. Res. Lett., 86, 3, 929-940, doi:10.1785/0220140149.

Schlömer, A., J. Wassermann, W. Friederich, M. Korn et al. (2022a). UNIBRA/DSEBRA: the German seismological broadband array and its contribution to Alparray – deployment and performance, Seismol. Res. Lett., 93, 4, 2077-2095, doi:10.1785/0220210287.

Schlömer, A., J. Wassermann, J. Plomerová, L. Vecsey et al. (2022b). AdriaArray Temporary Network: Albania, Austria, Czech Rep., Germany, Hungary, Kosovo, Montenegro, Slovakia, Data set, FDSN, doi:10.7914/2cat-tq59.

Schlömer, A., G. Hetényi, J. Plomerová, the AlpArray-PACASE Working Group et al. (2024). The Pannonian-Carpathian-Alpine seismic experiment (PACASE): network description and implementation, Acta Geod. Geophys., doi:10.1007/s40328-024-00439-w.

Sector for Seismology, Institute of Hydrometeorology and Seismology of Montenegro (1982). Montenegrin Seismic Network, Data set, FDSN, doi:10.7914/SN/ME.

Seismological Observatory at the Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, Skopje, Republic of Macedonia (1966). Seismological network of the Republic of North Macedonia, Data set, FDSN, doi:10.7914/v6tz-xh04.

Sens-Schönfelder, C. (2008). Synchronizing seismic networks with ambient noise, Geophys. J. Int., 174, 3, 966-970, doi:10.1111/j.1365-246X.2008.03842.x.

Seismological Survey of Serbia (1906). Serbian Seismological Network, Data set, FDSN, doi:10.7914/SN/SJ.

Slovenian Environment Agency (1990). Seismic Network of the Republic of Slovenia, Data set, FDSN, doi:10.7914/SN/SL.

Stehly, L., M. Campillo and N. M. Shapiro (2007). Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts, Geophys. J. Int., 171, 1, 223-230, doi:10.1111/j.1365-246X.2007.03492.x.

Strollo, A., D. Cambaz, J. Clinton, P. Danecek et al. (2021). EIDA: The European Integrated Data Archive and Service Infrastructure within ORFEUS, Seismol. Res. Lett., 92, 3, 1788-1795, doi:10.1785/0220200413.

Subbotin Institute of Geophysics of the National Academy of Science of Ukraine (2023). Ukrainian National Seismic Network, Data set, FDSN, doi:10.7914/jcvs-1309.

Swiss Seismological Service – SED (1983). National Seismic Networks of Switzerland, ETH Zurich, doi:10.12686/sed/networks/ch.

Technological Educational Institute of Crete (2006). Seismological Network of Crete, Data set, FDSN, doi:10.7914/SN/HC.

Trani, L., M. Koymans, M. Atkinson, R. Sleeman et al. (2017). WFCatalog: A catalogue for seismological waveform data, Comput. Geosci., 106, 101-108, doi:10.1016/j.cageo.2017.06.008.

U. S. Geological Survey (2020). Earthquake Lists, Maps, and Statistics, accessed in July 2024, https://www.usgs.gov/natural-hazards/earthquake-hazards/lists-maps-and-statistics.

Università Della Calabria, Italy (1981). Rete Sismica Unical, Data set, FDSN, doi:10.7914/SN/IY.

University of Athens (2008). Hellenic Seismological Network, University of Athens, Seismological Laboratory, Data set, FDSN, doi:10.7914/SN/HA.

University of Bari Aldo Moro (2013). OTRIONS, Data set, FDSN, doi:10.7914/SN/OT.

University of Genoa (1967). Regional Seismic Network of North Western Italy, Data set, FDSN, doi:10.7914/SN/GU.

University of Malta, (2014). Malta Seismic Network, Data set, FDSN, doi:10.7914/SN/ML.

University of Patras (2000). University of Patras, Seismological Laboratory, Data set, FDSN, doi:10.7914/SN/HP.

University of Trieste (1993). Friuli Venezia Giulia Accelerometric Network, Data set, FDSN, doi:10.7914/SN/RF.

University of Zagreb (2001). Croatian Seismograph Network, Data set, FDSN, doi:10.7914/SN/CR.

Vecsey, L., J. Plomerová, P. Jedlička, the AlpArray working group et al. (2017). Data quality control and tools in passive seismic experiments exemplified on the Czech broadband seismic pool MOBNET in the AlpArray collaborative project, Geosci. Instrum. Meth., 6, 2, 505-521, doi:10.5194/gi-6-505-2017.

Vecsey, L., P. Środa, J. Plomerová, G. Bokelmann et al. (2025). Northern Promontory of AdriaArray: Network Design and Realization, Ann. Geophys., 68, this issue, doi:10.4401/ag-9327.

Weidle, C., R. A. Soomro, L. Cristiano and T. Meier (2013). Identification of response and timing issues at permanent European broadband stations from automated data analysis, Adv. Geosci., 36, 21-25, doi:10.5194/adgeo-36-21-2013.

Wessel, P., W. H. F. Smith, R. Scharroo, J. Luis et al. (2013). Generic Mapping Tools: improved version released, EOS Trans. AGU, 94, 45, 409-410, doi:10.1002/2013EO450001.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship, Sci. Data, 3, 160018, doi:10.1038/sdata.2016.18.

Zaccarelli, R., D. Bindi and A. Strollo (2021). Anomaly detection in seismic data–metadata using simple machine-learning models, Seismol. Res. Lett., 92, 4, 2627-2639, doi:10.1785/0220200339.

ZAMG – Zentralanstalt für Meteorologie und Geodynamik (1987). Austrian Seismic Network, Data set, FDSN, doi:10.7914/SN/OE.

***CORRESPONDING AUTHOR: Petr KOLÍNSKÝ**,
Institute of Geophysics of the Czech Academy of Sciences, Prague, Czech Republic
e-mail: petr.kolinsky@ig.cas.cz