https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions

Title: Benefits of using convolutional neural networks for seismic data quality analysis.

Authors: Paolo Casale[1] and Alessandro Pignatelli[1]

[1] Istituto Nazionale di Geofisica e Vulcanologia, via di Vigna Murata 605, 00143 Rome, Italy
Correspondence to: alessandro.pignatelli@ingv.it

Keywords:
Seismic signal quality, Machine Learning, Convolutional Neural Networks

Abstract. Seismic data represent an excellent source of information and can be used to investigate several phenomena such as earthquake nature, faults geometry, tomography etc. These data are affected by several types of noise that are often grouped into two main classes: anthropogenic and environmental ones. Nevertheless instrumental noise or malfunctioning stations detection is also a relevant step in terms of data quality control and in the efficiency of the seismic network. As we will show, visual inspection of seismic spectral diagrams allows us to detect problems that can compromise data quality, for example invalidating subsequent calculations, such as Magnitude or Peak Ground Acceleration (PGA). However, such visual inspection requires human experience (due to the complexity of the diagrams), time demanding and effort as there are too many stations to be checked. That's why, in this paper, we have explored the possibility of "transferring" such human experience into an artificial intelligence system in order to automatically and quickly perform such detection. The results have been very encouraging as the automatic system we have set up shows a detection accuracy of over 90% on a set of 840 noise spectral diagrams obtained from seismic station records.

1.Introduction

Seismic data quality control is a key point to perform correct analyses of the signal produced by the seismic stations. For example: exact knowledge and regular check of instrumentation sensitivity are essential for a correct Peak Ground Acceleration evaluation and magnitude estimate after an earthquake; time marks on seismogram, number of gaps (data transmission efficiency check) are important for Early Warnings improvements (Picozzi et al., 2015); instrumental transfer function (Wielandt, 2012) (also called instrument response (Wielandt, 2012)) consistency and good signal-to-noise ratio plays a fundamental role in any analysis of low frequency signals (Pondrelli et al., 2020; Morelli et al., 2000; Custódio et al., 2014). Normally some checks are carried out in the operations rooms of seismic monitoring, first of all the connectivity check of the stations. For example, there are procedures that display the latency of each individual station in real time (Michelini et al., 2016). There are also programs that calculate other signal parameters in the time domain such as Offset, RMS (Root Mean Square of the signal) etc. In these cases, however, these are summary checks that do not provide sufficient information about the quality of the signal and above all do not ensure whether it is usable for calculating some important Earthquake quantities

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

such as Peak Ground acceleration (PGA), Richter Magnitude (M_L), Regional Centroid Moment Tensor (RCMT, (Morelli et al., 2000) etc. For example, the RMS, being a quadratic estimate of the counts (of the digitised signal), does not take into account the different sensitivities (Wielandt, 2012) between stations, nor if the sensitivities or instrumental responses (Wielandt, 2012) are incorrect. For the reasons mentioned above, currently most of the malfunction reports concern the lack of data. This is also because the signal has not yet been sufficiently analysed and the potential of spectral diagrams has not been fully explored. As identifying large gaps is generally already performed by automatic procedures, the main aim of this work is to capture signal anomalies (even serious and not rare) that often go unnoticed, and, in order to achieve such a result, a more in-depth systematic study of the input signal in the monitoring rooms is required.

Seismic noise analysis, especially using its several spectral approaches, is a powerful way to investigate stations' performance (McNamara and Boaz, 2006) in order to learn about noise sources as well as to check station data quality, detecting operational problems(McNamara and Boaz, 2006) with a view to the management of seismic networks. In particular, (McNamara and Buland, 2004) visually recognizing, in a noise spectral diagram, highlights some features due to specific malfunctions, for example too many calibration impulses, errors in the instrumental response (Wielandt, 2012) data loss (little gaps), frequent re-centering of the seismometer mass, etc. (Wielandt, 2012) and can suggest some methods for determining the self-noise of seismometers and for separating instrumental noise from seismic one. (Sleeman et al., 2006) uses three seismometers operating side by side and proposes a method to measure their self-noise using coherence analysis.

Due to the high number of stations and the different types of waveforms and relative spectra, visual identification of different event types (landslides, volcanic tremors, small earthquakes, explosions etc.) or possible precursor signals or, in many cases, instrumental malfunctions requires too much time and efforts to be systematically performed by a "human", even if he is an expert. So help from the neural networks can be desirable, especially if dealing with noise, as assuming its non-deterministic nature (Scales and Snieder, 1998; Bormann and Wielandt, 2013). For a general description of neural networks and their state of the art, see par. 4.

In literature, there are many papers about artificial intelligence applied to the analysis of seismic data (Pignatelli et al., 2021) and in some of them neural networks are specifically applied to noise handling but such works concern adjacent sectors, for example denoising (Bekara and Day, 2019) or detection of specific types of noise or poor data quality in the seismic prospecting topic (Thorp et al., 2020; Mejri and Bekara, 2020). For example, (Thorp et al., 2020) shows how deep learning can be used to classify if an image of a seismic stack can be considered in terms of quality level related to some specific noise types or geological features visibility .

The novelty of this paper is that a huge visual "human experience" focused on the entire noise spectra to detect acceptable signals from anomalous ones, has been transferred to artificial intelligence in order to automatize check operations for the efficiency of the seismic network. More specifically, we describe how some seismometer malfunctioning (or errors in metadata) can be reliably detected by an "expert eye" looking at noise spectral diagrams and how such an experience can be used to train convolutional neural networks to automatically perform such detection. This allows an efficient monitoring of the seismic station signal and is definitively a great support to improve data quality control of a seismic network. In this work we used diagrams obtained, for the most part, from the data of the Italian Seismic Network (ISN, code IV) (Michelini et al., 2016), in part from the Mediterranean Very Broadband Network (MedNet)(Pondrelli et al., 2020) and in part from some of the local networks (http://cnt.rm.ingv.it/en/instruments) whose data flows into the seismic monitoring rooms of the Istituto Nazionale di Geofisica e Vulcanologia (INGV) or however contribute to the INGV surveillance system(Michelini et al., 2016) .

2. **Seismic noise and its spectral representation**

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

Noise analysis is an important step for seismological data quality check. Although there is still discussion about what signal and noise are in seismology (Scales and Snieder, 1998; Bormann and Wielandt, 2013), generally seismic signal is defined as the earthquake induced ground motion recording, while seismic noise as the recording of ground vibration due to any other source(Holcomb, 1989). Actually, the noise (not only seismic) recorded by seismometers also includes several disturbances, often induced by instabilities and noise of the instrumentation (Bormann and Wielandt, 2013; Wielandt and Steim, 1986) or surrounding microclimate that can affect seismometers directly(Wielandt et al., 2002) even without generating ground vibration. Hereafter, we will therefore refer to noise in this last and more general acceptance and to seismic noise as ground noise.

In the waveforms recorded by a seismic station, the signal of seismological interest ranges over several orders of frequencies in relation to the physical phenomena to investigate or the result we want to extrapolate: for examples, the solid tide induced by the moon and the sun (Broucke et al., 1972) pertains to the sub-mHz region; the Earth's normal modes to the mHz one (Wielandt and Steim, 1986); tenths-hundredths of a Hz is the region useful to estimate the magnitude of large earthquakes (Morelli et al., 2000); 1-10 Hz is the range to localise earthquakes (Allen, 1982; Vassallo et al., 2012); tens of Hz for studies of tomographic detail and seismic prospecting (Ma et al., 2005). So the knowledge of phenomena that affect the useful signal, such as seismic noise and, more generally, noise over the whole frequency range is important.

The amplitude of seismic noise is different at different frequencies as shown in the noise spectrum of Figure 1. The physical quantity on the ordinate is the Power Spectral Density (PSD) (Bendat and Piersol, 2011; Peterson, 1993), in absolute decibel (dB), referred to the PSD calculated for the acceleration of 1 m/s^2. Definitely 0 dB corresponds to 1 $(m^2/s^4)/Hz$ (Rastin et al., 2012; Bormann and Wielandt, 2013). The figure reports the level of Low and High seismic noise (black curves), according to the (Peterson, 1993) empirical models for the ground acceleration: the Low Noise Model or LNM (bottom curve) represents the empirical minimum while the High Noise Model or HNM (top black curve) can be taken as a reference for the maximum acceptable noise for a useful seismic station.

Not only the amplitude but also the seismic noise sources are mostly different in different frequency ranges (Holcomb, 1989). As shown in Figure 1, three main regions can be distinguished. For frequencies higher than 1 Hz, seismic noise is mostly due to anthropic activities (trains and cars, industrial activities, engines, turbines etc) (Stutzmann et al., 2000), even though rain and wind may also contribute. At intermediate frequencies, ranging from 0.04 to 1 Hz, sea activity is recorded, showing two peaks (Darbyshire and Okeke, 1969). LNM shows that the highest one is around 0.2 Hz and this is called secondary frequency peak. The lowest one (primary frequency peak) is around 0.07 Hz. At frequencies lower than 0.04 Hz, the atmospheric pressure changes become more important as the frequency decreases (Sorrells, 1971). According to (Sorrells, 1971), the pressure changes induce a ground tilt and it has a greater effect on the horizontal components (Wielandt et al., 2002; Wielandt and Forbriger, 1999) (or see e.g. green and blue curves in Figure 1) of the seismometers. Moreover, in this frequency range, the temperature variations produce a direct effect on the seismometer(Wielandt et al., 2002) and this is an example of "not seismic" noise: also in this case, the lower is the frequency, the higher is the effect.

Figure 1 also exhibits the noise level in terms of PSD) recorded by AIO seismic station (Antillo, Sicily). We report the vertical component in red, the North-South (NS) in blue and the East-West (EW) in green. Since in the temporal waveforms used for the PSD, no appreciable earthquake has been observed, the diagrams represent the noise mean over a month. However, to get an estimate of the noise statistical distribution over a time interval (one month for example), other representations can be more useful, as described below.

2.1 Probability Density Function: description and example.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions

The INGV receives and stores signals from seismic stations. As the use of such a signal is not known in advance, the continuous waveform is permanently collected without discarding any time interval. The seismic waveforms are accessible online on the EIDA database (EIDA Italia, https://eida.ingv.it/it/). In addition to the waveforms, the database also contains information about instrumental parameters of the station (for example sensitivity, sampling rate, position, etc.). Such information is called "metadata". To better represent the noise, the SQLX (Seismic data Quick Look eXtended https://sqlx.science/) software package (McNamara and Boaz, 2006; McNamara and Buland, 2004; Marzorati and Lauciani, 2015) is useful. Basically, SQLX takes as input the waveforms (in the time domain) and metadata (in particular the instrument transfer function) contained in the database and produces a series of control diagrams such as spectrograms, PSDs, noise amplitude behaviour at selected frequencies, trace statistics etc. In addition to the averaged PSD, like the diagrams of Figure 1, SQLX also produces Probability Density Function (PDF) diagrams (McNamara and Buland, 2004). An example of PDF is shown in Figure 2. The concept of PDF is similar to the PSD, but it contains more information as it does not show the noise mean but its statistical distribution during the analysed time-interval. Other advantages of the PDF are the use of continuous data and the accessibility to accurate documentation, so it has been adopted around the world(Custódio et al., 2014). Details of the PDF calculation method are described in(McNamara and Buland, 2004) but here we highlight only some aspects:

a) a PDF is produced by an elaboration of multiple PSDs and the statistical distribution of the noise is graphically represented by a colour scale;

b) the power spectra obtained from station waveform (the records are in digital counts) are deconvolved, by the instrument transfer function, to remove the instrumental response (Wielandt, E., 2003) in order to get the true value of the PSD at different frequency ranges in particular at low frequencies, where the sensitivity of the seismometer is not constant as it decreases with frequency (Wielandt et al., 2002) ;

c) in the PDF calculation procedure, earthquakes, or other transient phenomena (such as occasional disturbances, brief anomalies in the flow of data) are not discarded. However, having these events a small duration (except for important and lasting seismic sequences), they appear in the PDF with colours associated with a low level of probability. On the other hand, the conditions of quasi-stationary seismic noise or lasting instabilities (permanent failures in the instrumentation, errors in the instrumental transfer function, excessive number of gaps in the waveform, etc.) produce a high level of occurrence(McNamara and Boaz, 2006) reflected by a corresponding colour in the PDF plot.

As a consequence, as shown in Figure 2, the areas where colours have greater probability of occurrence (red, yellow, green, cyan ...) represent the noise level "closest" to the average, while the parts with less likely colours, as magenta, substantially represent the fluctuations. In this way, it is possible to have a picture of how the noise level is distributed in PSD and frequency, during the whole inspected time interval.

As an example, Figure 2 displays the PDF distribution for the E-W component at TUE station (Stuetta, SO, Italy) in central Alps, from January 1st to February 8th, 2021. To plot this diagram, 1678 PSDs have been analysed. The used transfer function is associated with the instrumentation which, in this case, is constituted by a Very Broad Band (VBB) station. A seismic station is considered VBB if it has a high dynamic digitiser and a seismometer whose sensitivity is high and constant over a wide range of frequency (Wielandt et al., 2002), in order to cover almost the entire spectrum of seismological interes t(Wielandt and Steim, 1986). A Broad Band (BB) station is similar to a VBB one but with a slightly narrower bandwidth (Wielandt et al., 2002). TUE instrumentation (adopting a sampling rate of 100 sps) fulfils VBB requirements (Pondrelli et al., 2020, Anon; Mazza et al., 2008; Kinemetrics Inc., n.d.).

Figure 2 also reports the Peterson LNM and HNM reference curves in black. We observe that the noise at TUE site has an excellent behaviour in all the frequency range, even in the winter period, when the environmental seismic noise is greater than in summer (Custódio et al., 2014). In fact, PDF shows that the PSD highest occurrence is quite close to the LNM curve. No malfunctions of the instrumentation are observed. Nonetheless, two slight effects can be noted, both attributable to the

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

near dam for the production of electricity: A) a slight peak around 8 Hz due to the frequency (8.33 Hz = 500 rpm) of the turbine; B) a low-frequency bimodal trend, for frequencies below 0.03 Hz, probably due to the tilt of the ground (Wielandt et al., 2002) which, in this case, can occasionally be induced by the variation in the water pressure of the dam on the bottom when the water level changes to produce energy.

As we have just seen, PDF diagrams are very useful to observe the general trend of a seismic station and in particular to recognize minor (as in Figure 2) as well as important anomalies in the recorded signal. In the next paragraph we will give examples indicating some criteria to recognize a malfunction and possibly to hypothesise the causes of specific problems in the seismic station or in the metadata.

3. Criteria to detect problems in seismic data using noise spectra

In the previous paragraph we showed the noise spectrum of one of the best Italian stations (Figure 2). In general, any ISN station will usually show a higher noise level than TUE. However, as a general criterion, the spectrum shape, representing the noise trend as a function of the frequency, for a properly working station has to be similar to the one shown in Figure 2: more specifically, the secondary frequency peak at 0.2 Hz, due to marine activity, must be evident if we analyse a sufficiently long-time interval. In addition, too high or too low a noise level can highlight problems, for example in the instrumental sensitivity. Excluding large gaps, there are two general causes of degradation of seismic data and/or subsequent data processing:

A) Problems with the station's instrumentation (including little gaps in data transmission)

B) Errors in the metadata stored in the DataBase (for example the sensitivity value): they invalidate many subsequent processes.

Here we give just a few examples of how and how much (to what extent, to what level of detail) such problems can be recognized by analysing a noise spectrum, in particular a PDF diagram.

3.1 Problems at the seismic station.

An example of a malfunctioning station can be seen in Figure 3 which shows the PDF diagram at the RMP station (Monte Porzio, Rome). The instrumentation is very similar to the TUE station (par. 2) i.e. a 120 s very broad band Trillium seismometer (http://support.nanometrics.ca/) and a GAIA digitizer, which has almost 24 bits(Cirrus Logic Inc., n.d.) (https://www.mouser.com/datasheet/2/76/cs5371-72_f3-1160187.pdf) These slight differences in instrumentation are not notable in the spectral representation. In Figure 3a, the lower edge of the coloured area follows a sloped straight line almost everywhere, except for a slightly pronounced peak around 0.2 Hz and for the behaviour at the extreme frequencies. Moreover, from 0.02 to at least 3 Hz this lower edge is almost the mode i.e., the most statistically populated trend, as can be seen from the colours (red, yellow, green ...) associated with the highest probability of occurrence. In general, the shape of the curve looks very different from a normal VBB station, such as TUE (Figure 2). The shape of the diagram indicates problems with the seismometer, probably in its positioning. The subsequent on-site intervention revealed that the malfunction was due to a slight inclination of the seismometer ("out of bubble"). Once the seismometer was repositioned, the diagrams mirrored a normal VBB station (Figure 3b) with its typical shape. In fact, in Figure 3b (showing the PDF of the same station after the intervention) the lower part of the coloured area no longer has a straight edge and, especially at medium and low frequency, the trend has a maximum and a minimum. In particular, in a medium frequency zone (0.07--0.4 Hz), it follows the LNM curve and resolves the secondary peak well. Note that RMP shows a very noisy (although still acceptable) pattern at high frequencies, but the shape and the shades of colour in the diagram indicates site noise and not a serious instrumentation problem.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions

As other examples, Figure 4 (a and b) show two spectra that highlight problems at the seismic station. In Figure 4a the PDF of the TERO Broad Band station (Teramo, Italy, in Central Apennines) shows a bimodal trend at medium and low-medium frequency, the most frequent curve is far below the LNM curve. This means that the BB seismometer has a not-constant behaviour, revealing that it is often not powered or switches from BB to "Short Period" mode (Custódio et al., 2014). For practical purposes the two are substantially equivalent; in fact in both cases any subsequent analysis using Long Period amplitudes is invalidated (Custódio et al., 2014). Figure 4b shows the noise at ARRO station (Arrone TR, Central Apennines) where INGV has installed a short period seismometer Lennartz type Le5s whose corner frequency (Wielandt et al., 2002) is equal to 0.2 Hz (5 s period, https://www.lennartz-electronic.de/wp-content/uploads/2021/04/Lennartz-SeismometerManual.pdf). On the vertical component (indicated by the code EHZ in the title of the diagram; codes of seismic data channels are described in (Scott, Halbert, n.d.)) in Figure 4b, one can see a splitting of the blue curve into a bimodal trend. This behaviour reveals an intermittent problem in the vertical component of the seismometer. In fact, a correct trend would be the one shown in the lower blue curve, representing the typical trend of "short period" stations. This type of anomaly on the vertical component was also observed in other stations with the same seismometer (Le5s) such as ATCC (Casa Castalda, [PG], central Apennines), FOSV (Fossato di Vico [PG]) etc.

3.2 **Errors in the metadata (Instrumental Transfer Function)**.

A sensitivity error is shown in figure 5a reporting the noise trend at the GIGS VBB station (Gran Sasso, Central Italy) over a time interval of more than 5 months. The shape of the spectrum is correct because for frequencies higher than 0.07 Hz it follows LNM shape but its value is too low, i.e. the correct spectrum would be the same but translated at higher PSD values. An error (subsequently corrected) is evident in the memorization of the instrumental sensitivity while populating the Metadata archive.

The last example concerns a short period station: in Figure 5b we see the noise at the BARO station (Barbarano, VT, Central Italy). Here a 5 s seismometer (https://lunitek.it/seismic/seismic-sensors/tellus-5s/), is installed (similar to ARRO's Le5s). The shape of the spectrum is very different from the typical one for this short period seismometer, shown in the lower blue curve of Figure 4b. In fact, an error was readily found in the transcription of the polynomials describing the instrumental transfer function. After the correction, the diagram exhibited the correct trend.

So, in this section we have shown that PDF visual analysis can potentially detect seismic station malfunctions and errors in metadata. However the shape of the noise spectra depends on many variables (Bormann and Wielandt, 2013) that can be environmental (Custódio et al., 2014), anthropogenic (Stutzmann et al., 2000), instrumental(Bormann and Wielandt, 2013) or related to the quality of the station installation(Anglade et al., 2015), therefore the variety of the PDFs diagrams is much wider than the few examples shown here. There are several other examples, leading an expert user to detect malfunctioning and often to understand its causes too (McNamara and Buland, 2004; McNamara and Boaz, 2006). However, the stations and the diagrams produced by automatic spectra generators are too many to be checked every time by “human” operators. On the other hand, as we have seen, the noise spectra have a complex typology and therefore transferring the criteria of a complex visual recognition into an adequate deterministic algorithm is not easy. On the contrary, the simplified criterion “good/bad” based on noise average which must remain between a minimum and a maximum threshold (Massa et al., 2022) (even if on selected frequency bands) does not represent the complexity of human visual evaluation. In fact, for example, the expert evaluates not the single average trace (figure.1) and not in selected frequency bands but annual or monthly distribution diagrams of the noise spectra, as are PDFs (as in figure 2).

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023


Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

For all the reasons described above, in this work, we have considered noise spectra of many stations to explore the possibility of using artificial intelligence able to automatically detect diagrams revealing malfunctioning stations. In the following paragraphs we describe how we used a pre-trained convolutional neural network to perform "transferring" of human experience to an automatic system.

4. Machine learning general description

The main benefits in using Machine learnings are:

1) Machine learning algorithms are very general and, without specific changes, except the input data, they can solve many different problems. That's the reason why such techniques are becoming very popular in many different areas as medical diagnosis, business decisional processes, face authentication, scams recognition and many others;
2) Machine learning application requires, as only effort, to provide enough data to let the machine learn how to solve the problem and then the system will be able to solve the same problem autonomously and automatically in different situations, assuming that such situations are similar to the provided ones.
3) Machines' emotions' lack can be very advantageous to avoid statistical bias. For example, medical research procedures require double blinded protocols to avoid biases.

As well known, machine learning algorithms are divided into two groups: supervised methods and unsupervised ones.
A supervised method tries to predict a variable (*answer* or *label*) when some others (*predictors*) are known. As an example, in "mortgage lending", the predictors could be personal data such as age, income, job position etc….. and the answer could be the knowledge if the person requiring the mortgage will default or not.
In order to apply a supervised technique, the first step is to collect data including both the predictors and the answer (*labelled data*). Such data will feed the algorithm that will "learn by itself" if predictors show patterns (said *features*) able to predict the answer. This machine learning stage is named *training*. This stage is probably the most difficult one as the user has to collect labelled data. However, after the training, if successful, the algorithm is able to automatically predict the answer for any other datasets including all the provided predictors.
In order to check how the model is able to "generalise" its validity and how it is accurate when predicting labels for data not included into the training, a common practice(Lantz, 2013) is, before doing the training, randomly choose a part of the available dataset and remove it from the training process. So, when the training has been performed, the excluded data (named *test data* and containing the labels we are looking for) provide a powerful source to verify the model accuracy as this dataset includes both the true labels and, when used as input for the model, the predicted labels. This allows the accuracy percentage computation *Accuracy = matched results/total test data*. When dealing with machine learning, the best way to show methods accuracy is using the "confusion matrix" (Lantz, 2012). A confusion matrix is a table counting where the predicted classes agree or disagree with the true values as described in the par. 5 and in Figure 7.
The unsupervised machine learning techniques (the most important one is clustering) are basically methods to group similar records of a dataset. How many records must be considered similar depends on different mathematical metrics related to the specific algorithm. A "real life" example is clustering used by retailers to divide customers into different groups according to their personal data (gender, income, age, etc...). Such groups are used to assess different promotions.

4.1 Artificial neural networks description

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023


Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

Artificial neural networks (ANNs) represent a part of supervised machine learning techniques (Lantz, 2013). ANNs "architecture" is inspired by the structure of the primate cerebral cortex, in order to learn increasingly abstract features of the input and best support the desired output (O'Shea and Nash, 2015; Rawat and Wang, 2017). Basically, ANNs consist of a set of layers of neurons. Input data is inserted into the first layer (each input data into a single neuron) and each neuron elaborates the input value according to a transfer function depending on the ANNs used. Once each data input has been processed by the neuron it is "passed" to the next neuron's layer where the input is a linear combination of the previous layer outputs. Such linear combinations are determined by a set of coefficients (named neuron *weights*). The process is iterated along all the layers to the last layer (named "output layer") and the outputs are compared to the provided labels. According to the discrepancy between the *true* labels and the got ones a *backward propagation* process starts modifying the weights. Such propagation involves the iterative adjustment of a single parameter vector with the goal of minimising the differences between the true and predicted values (Cao and Parry, 2009).

4.2 Convolutional neural networks

Convolutional neural networks (CNNs) are a specialized type of artificial neural network that excel at extracting local features from matrices. Their structures are more intricate, and their layers primarily consist of filters applied to matrices (Cao and Parry, 2009). Unlike traditional neural networks that recognize global features, the main task of CNNs is to identify local features. This means that each neuron focuses on a limited number of inputs (or neurons from the previous layer) instead of collecting data from all of them.

A detailed explanation of CNNs would be too lengthy and specific for the scope of this paper. To gain a deep understanding of CNNs, we recommend referring to (O'Shea and Nash, 2015; Rawat and Wang, 2017). In this paper, we will provide an overview of the main features of CNNs, which consist of four types of layers: convolutional layers, pooling layers, fully connected layers, and a single softmax layer (O'Shea and Nash, 2015).

Convolutional layers are the most essential component for examining local features, as only a limited number of inputs/neurons are connected to each neuron of the next layer. This approach enables each neuron to inspect specific areas. Pooling layers are generally added after convolutional layers to reduce their size. For instance, each 2x2 convolutional output matrix is replaced by its maximum value.

The fully connected layers play an important role in the CNN by modifying the output matrix. Typically, the first fully connected layer in a CNN is followed by additional fully connected layers that transform the output from a 2D to a 1D structure (Pignatelli et al., 2021). For instance, a fully connected layer might represent a 6x4 matrix as a single vector of 24 components.

After passing through the fully connected layers, the output is fed into the softmax layer, which converts the numbers in the last fully connected layer into probabilities for each label. Finally, the predicted label is the one that corresponds to the highest probability value.

4.3 Pre-trained convolutional neural networks

Before training a convolutional neural network, a very difficult step is deciding its architecture (in other words the sequence of layer types and their parameters). However previous works (Pignatelli et al., 2021) have shown how using images of data rather than data can be very advantageous as it's possible using pre-trained networks.

Pre-trained networks have very efficient architectures already tested for general purposes (such as images classification). A very famous one is Alexnet (Han et al., 2017; Krizhevsky et al., 2012).

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

The primary function of Alexnet is to extract meaningful features from images, which can be used by the final layers of the network for various classification or regression tasks. One key advantage of using Alexnet is that many of the weights are already pre-calculated to identify basic image features (such as geometric shapes, segments, etc.). Consequently, the network is largely pre-trained and optimized for use, and the training process mainly involves fine-tuning the pre-existing parameters. This results in a faster and more efficient training process compared to designing a neural network from scratch. Additionally, a smaller amount of input data is required compared to a new neural network.. Alexnet model is shown in Figure 6.
As our main task was to discriminate between spectra coming from well functioning stations and from malfunctioning or broken stations, we realised that using Alexnet was a possible great approach.

5.Method application

As in our study the goal is to discriminate regular signals from signals recorded by damaged stations (bad metadata can also "damage" spectra), it's mandatory to collect both human classified "ok signals" and "broken signals". In our study we find that just images of data fulfil our goal, so basically we have collected two groups of images and labelled them "ok" or "broken" according to our knowledge. In our specific case, as a set of data, we have chosen PDF spectral diagrams, similar to those seen in paragraphs 2 and 3 (Figures 2 to 5). We recall that they substantially represent the level of seismic and/or instrumental noise of a seismic station. As a first step we have selected broadband (BB) stations with 100 sps sampling and flat response to ground velocity at least up to 40 seconds, that is with a band from no more than 0.025 to at least 25 Hz (see par. 2), whose channels are conventionally named HH*(Scott, Halbert, n.d.) , for the three cartesian components of the ground motion: HHZ (Vertical), HHN (North-South) and HHE (East-West). The choice of limiting the analysis to the broad (and very broad) band was meant to ease learning, as the diagrams of these stations are similar to each other but significantly different from those of the short period stations (channels EH*, SH*,...) especially at low frequency, as is evident when comparing Figure 2 with the lower part of Figure 4b. In the first approach we used data of the year 2019. The BB and VBB considered stations are a subset of about 300 stations over about 500 ones that in 2019 provided data to the INGV monitoring centers (Michelini et al., 2016) and to the INGV database. For each station, the SQLX package (paragraph 2) produces a large variety of diagrams among which we have chosen the annual diagrams, as the seismic noise (and sometimes also instrumental one) can undergo considerable variations not only day/night but also weekly, monthly or seasonal. Using the annual averages, it is easier to divide the diagrams into "OK" and "Broken", as within the "OK" class the annual diagrams are more homogeneous and similar. In summary, the chosen data set for the first training has the following characteristics:

a) Broadband (BB) or VBB stations (HH* channels) belonging to Italian, Mediterranean or local networks.
b) PDF spectra (similar to figures 2—5)
c) Annual diagrams for the year 2019
d) All three components of motion are analysed (a diagram for each component).

5.1 Training and first test

We collected 280 PDF spectral diagrams of the year 2019, respectively 140 "OK" and 140 "Broken", the latter showing trends that, according to the human eye, definitely belong to defective stations. Once such data had been gathered, we used them to feed the Alexnet network. As described in the machine learning general description section, to check the method accuracy we randomly split the dataset into two groups: 80% (224 diagrams) of training data and 20% (56 diagrams) of test data using

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

the “stratification” option (Kubat and Kubat, 2017) to keep the classes as much possible distributed as the total dataset. We performed the training step providing Alexnet with 224 diagrams and declaring for each one whether it was "OK" or "Broken". Once the training process ended, we applied the trained model to the test data, i.e. the 56 diagrams not used for training, specifically 28 “Broken” diagrams and 28 “OK”. The results show an accuracy of 92% and are detailed in figure 7.

5.2 Second test: Increase images number to confirm network accuracy and improve its robustness

After the first test, additional labelled data have been provided to the Alexnet to improve the learning phase. More specifically 61 additional “OK” and 106 additional “Broken” diagrams belonging to years other than 2018 were provided including their labels. So, the total of the diagrams used for the learning phase was 447. The reason why we preferred to increase mostly the number of "Broken" diagrams is that they present a greater variety as there are many possible causes of malfunction in a seismic station, while broadbands “OK” are more homogeneous. Results are summarised in the confusion matrix in Figure 8. As one can see, Increasing the number of labelled data has resulted in a significant increase in accuracy in the test (96.6%). However, we considered data used for the test still not enough (89 diagrams) and so we have performed a more robust test on a much larger number of diagrams.

5.3 Third test: Verification with data from the year 2018

As a more robust test, we selected 840 diagrams of the year 2018, therefore never previously analysed by the neural network Alexnet. Actually, the human operator first examined 877 diagrams of 2018 and discarded 37 of them because they were considered “uncertain”. 494 (58.8%) of the remaining 840 were classified as “OK” and 346 (41.2%) as “Broken”. This was the work that took most of the time. Before letting the machine analyse the 840 diagrams of the 2018, a new training was performed with the same 447 diagrams of the previous second test. In this case, as the test was meant to be applied to such new data, we decided to increase the training data percentage with respect to the previous training. So, 90% of diagrams (and not 80%) were used for the training as the significant test was successively performed with new data. We needed to spare 10% of data for the validation(Lantz, 2013) process.
The results are shown in figure 9. In total, the machine guessed 758 diagrams out of 840, equal to an accuracy of over 90%. Compared to the individual classes, the accuracy is 88.4% for the “Broken” class (306 out of 346) and 91.5% for the “OK” class (452 out of 494). Throughout this analysis, we have taken for granted the infallibility of the human operator and this is not an established fact, but the hypothesis of operator error is remote precisely because the "uncertain" diagrams have been discarded a priori by the human.

5.4. Fourth test: introducing a third class.

In the last test we noticed a high percentage of successes (over 90%), and a not similar percentage of false positives (8.1%) and false negatives (12.1%) indicating a “precautionary” trend of the neural network. The choice of having a "precautionary" or alternatively "unscrupulous" neural network depends on whether we want a slightly suspicious station to be reported to us as broken (with the risk of having a consistent number of false alarms) or if we want to deal with stations that are definitely malfunctioning (with the risk of losing the signalling of a number of stations that are not working well). It is a choice that often depends on external conditions, there are many elements to decide a strategy or another: for example the total number of stations, the density of the seismic network, the relative position of the damaged stations and our capacity for rapid intervention on the station site or remote intervention.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023


Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

Having a very or little precautionary neural network depends on how we train the neural network, for example on how we choose to consider "uncertain" diagrams. Every choice we make has pros and cons. In fact, compared to the "uncertain" diagrams, we can: a) discard them all a priori in learning --but Alexnet will sooner or later meet "uncertain" spectra among the many diagrams that we will give it to analyse and in those cases how will Alexnet compute?; b) insert them all in the OK spectra --less precautionary choice; c) insert them all in the Broken spectra --more precautionary choice; d) decide and classify diagram by diagram --a more demanding choice for the human operator and perhaps producing different results from operator to operator, e) create a third intermediate class of "uncertain" diagrams but, in the 2019 data set, they were not many and machine learning would be based on a few diagrams.

The INGV database also receives data from other local networks. Recently the number of stations received has significantly increased and, with it, also the percentage of uncertain diagrams, in many cases because there are temporary failures due to the non-connection between station and INGV, therefore the continuity of the data is lost; in these cases the choice whether to declare a malfunctioning station or not depends on the quantity and length of the gaps (time intervals without data or telemetry drop-out) but the evaluation is still subjective and there will always be a percentage of questionable cases. Furthermore, in local networks the occurrence of non-standard installations is more frequent than in the ISN and sometimes cheap instruments are used; in these cases the documentation to control all parameters is not always complete or clear. So the 2021 spectra data set shows a number of "uncertain" diagrams adequate to train Alexnet with three classes: two extreme (OK and BAD) and one intermediate (Dubious).

For the three-class learning we used a much larger number of diagrams than in the previous steps. We selected the year 2021 and collected 1865 PDF spectral diagrams. The intervention of a single human expert lasted about three weeks to classify them into 3 classes, respectively 554 "OK", 476 "Dubious" and 835 "BAD". As in the previous analyses, for training, (learning) Alexnet used 80% of the diagrams and for testing the remaining 20% (373 diagrams). The results are summarised in figure 10 (a and b).

With the introduction of the 3rd class (intermediate) the total accuracy decreased up to 85.5% but on the other hand a very comforting result is that no BAD diagram has been judged OK and no OK diagram has been judged BAD by the neural network. Thus no "absolute" false positives or "absolute" false negatives resulted out of the processing. However, there are some BAD or OK diagrams that have been judged Doubtful and conversely there are some Doubtful diagrams that have been placed in the OK or BAD classes.

The result can however be considered useful overall because the network has shown that it can restrict the set of diagrams on which to deepen the analysis and on which we should probably intervene (no OK diagrams was judged BAD). Then, according to the opportunities and the above considerations (number and density of the stations etc.) the operators will evaluate if and when to investigate the doubtful cases.

The reasons why the number of damaged stations increased in 2021 are similar to those why the uncertain diagrams increased: a growth of connected stations and an increase (in absolute and percentage) of stations with degraded data transmission. Pending improvement of the data transmission systems, being able to narrow down the field of malfunctioning stations is in any case very useful for scheduling maintenance interventions.

6. Conclusions

Good quality seismic data collection is very important in terms of geophysical research. More specifically, spectral analysis, in its various forms, is a very important investigation step for the purpose of checking the quality of seismic data. Unfortunately check and maintenance of many stations require a lot of "human" effort to keep the system running properly. In this paper, we have shown how visual inspection represents a powerful tool to achieve a high standard check level and

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

how artificial intelligence can be helpful in such tasks. More specifically the experience and knowledge of an expert in this field can be "transferred" into a neural network able to automatically discriminate signals coming from malfunctioning stations from the ones collected from working ones, with an accuracy of more than 90%, obtained in the third test on 840 diagrams. Previous tests were performed with a smaller number of diagrams and gave higher accuracy. In any case, a good policy is to try to increase the accuracy of the network as much as possible by providing even more data for learning. The accuracy has increased from the first to the second test from about 93% to over 96% as additional learning diagrams (label) were provided before the second verification test.

To further restrict the set of stations on which to intervene, we have introduced a third intermediate class which contains the diagrams showing a dubious operation of the station (the number of this type of diagrams increased in the last two years). The idea is to intervene initially only on stations judged BAD, reserving future insights on stations judged "uncertain". For the three-class learning we used a much larger number of diagrams than in the previous steps. With the introduction of the 3rd class, the total accuracy decreased up to 85.5% but on the other hand no BAD diagram has been judged OK and vice versa. The errors concern having placed an element in a class rather than in the next one. This result shows the reliability of Alexnet and allows us to concentrate the first interventions only on the stations judged as BAD and subsequently to investigate the Dubious ones. Being able to narrow the number of very suspicious stations to intervene on is still a very useful result for scheduling maintenance intervention. The applicability of the trained network may be not limited just to the Italian network but to every international one having BB or VBB seismometers.

The overall results look very promising for future developments where we hope to extend the analysis to short period stations and to be able to carry out the analysis on lower time intervals (for example one or two months), possibly also using other types of diagrams.

References

Allen, R.: Automatic phase pickers: Their present use and future prospects, Bull. Seismol. Soc. Am., 72, S225–S242, 1982.

Anglade, A., Lemarchand, A., Saurel, J.-M., Clouard, V., Bouin, M.-P., De Chabalier, J.-B., Tait, S., Brunet, C., Nercessian, A., and Beauducel, F.: Significant technical advances in broadband seismic stations in the Lesser Antilles, Adv. Geosci., 40, 43–50, http://dx.doi.org/10.5194/adgeo-40-43-2015, 2015.

Anon: STS-2 Legacy Product datasheet, 2020.

Bekara, M. and Day, A.: Automatic QC of denoise processing using a machine learning classification, First Break, 37, 51–58, http://dx.doi.org/10.3997/1365-2397.n0055, 2019.

Bendat, J. S. and Piersol, A. G.: Random data: analysis and measurement procedures, John Wiley & Sons, http://dx.doi.org/10.1002/9781118032428, 2011.

Bormann, P. and Wielandt, E.: Chapter 4: Seismic signals and noise, in: New manual of seismological observatory practice 2 (NMSOP2), Deutsches GeoForschungsZentrum GFZ, 1–62, https://doi.org/10.2312/GFZ.NMSOP-2_ch4, 2013.

Broucke, R. A., Zürn, W. E., and Slichter, L. B.: Lunar tidal acceleration on a rigid Earth, Wash. DC Am. Geophys. Union Geophys. Monogr. Ser., 16, 319–324, http://dx.doi.org/10.1029/GM016p0319, 1972.

Cao, Q. and Parry, M. E.: Neural network earnings per share forecasting models: A comparison of backward propagation and the genetic algorithm, Decis. Support Syst., 47, 32–41, http://dx.doi.org/10.1016/j.dss.2008.12.011, 2009.

Cirrus Logic Inc.: Low-power, High-performance ΔΣ Modulators https://www.mouser.com/datasheet/2/76/cs5371-72_f3-1160187.pdf, n.d.

Custódio, S., Dias, N. A., Caldeira, B., Carrilho, F., Carvalho, S., Corela, C., Díaz, J., Narciso, J., Madureira, G., and Matias,

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

492 L.: Ambient noise recorded by a dense broadband seismic deployment in western Iberia, Bull. Seismol. Soc. Am., 104, 2985–
493 3007, http://dx.doi.org/10.1785/0120140079, 2014.

494 Darbyshire, J. and Okeke, E. O.: A study of primary and secondary microseisms recorded in Anglesey, Geophys. J. Int., 17,
495 63–92, http://dx.doi.org/10.1111/j.1365-246X.1969.tb06379.x, 1969.

496 Han, X., Zhong, Y., Cao, L., and Zhang, L.: Pre-trained alexnet architecture with pyramid pooling and supervision for high
497 spatial resolution remote sensing image scene classification, Remote Sens., 9, 848, http://dx.doi.org/10.3390/rs9080848,
498 2017.

499 Holcomb, L. G.: Seismic noise, Encycl. Solid Earth Geophys., 1089–1092, http://dx.doi.org/10.1007/0-387-30752-4_132,
500 1989.

501 Kinemetrics Inc.: EQUIPMENT NEWS, SRL, 76, 84–85, https://doi.org/10.1785/gssrl.76.1.84, n.d.

502 Krizhevsky, A., Sutskever, I., and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, Adv.
503 Neural Inf. Process. Syst., 25, 1097–1105, http://dx.doi.org/10.1145/3065386, 2012.

504 Kubat, M. and Kubat: An introduction to machine learning, Springer, 2017.

505 Lantz, B.: Machine learning with R, Packt publishing ltd, 2013.

506 Ma, Y., Zhu, X., Guo, T., Rebec, T., and Azbel, K.: Reservoir characterization using seismic data after frequency bandwidth
507 enhancement, J. Geophys. Eng., 2, 213–221, https://doi.org/10.1088/1742-2132/2/3/005, 2005.

508 Marzorati, S. and Lauciani, V.: SQLX: Test di Installazione e Funzionamento, Rapp. Tec. INGV, 297, 25,
509 http://hdl.handle.net/2122/9435, 2015.

510 Massa, M., Scafidi, D., Mascandola, C., and Lorenzetti, A.: Introducing ISMDq—A Web Portal for Real-Time Quality
511 Monitoring of Italian Strong-Motion Data, Seismol. Res. Lett., 93, 241–256, http://dx.doi.org/10.1785/0220210178, 2022.

512 Mazza, S., Olivieri, M., Mandiello, A., and Casale, P.: The Mediterranean broad band seismographic network anno 2005/06,
513 in: Earthquake monitoring and seismic hazard mitigation in Balkan countries, Springer, 133–149,
514 http://dx.doi.org/10.1007/978-1-4020-6815-7_9, 2008.

515 McNamara, D. E. and Boaz, R. I.: Seismic noise analysis system using power spectral density probability density functions:
516 A stand-alone software package, US Geol. Surv. Open-File Rept, 2006.

517 McNamara, D. E. and Buland, R. P.: Ambient noise levels in the continental United States, Bull. Seismol. Soc. Am., 94,
518 1517–1527, 2004.

519 Mejri, M. and Bekara, M.: Application of Machine Learning for the Automation of the Quality Control of Noise Filtering
520 Processes in Seismic Data Imaging, Geosciences, 10, 475, http://dx.doi.org/10.3390/geosciences10120475, 2020.

521 Michelini, A., Margheriti, L., Cattaneo, M., Cecere, G., D'Anna, G., Delladio, A., Moretti, M., Pintore, S., Amato, A., and
522 Basili, A.: The Italian National Seismic Network and the earthquake and tsunami monitoring and surveillance systems, Adv.
523 Geosci., 43, 31–38, http://dx.doi.org/10.5194/adgeo-43-31-2016, 2016.

524 Morelli, A., Ekström, G., and Olivieri, M.: Source properties of the 1997–98 Central Italy earthquake sequence from inversion
525 of long-period and broad-band seismograms, J. Seismol., 4, 365–375, https://doi.org/10.1023/A:1026587817690, 2000.

526 O'Shea, K. and Nash, R.: An introduction to convolutional neural networks, ArXiv Prepr. ArXiv151108458,
527 https://doi.org/10.48550/arXiv.1511.08458, 2015.

528 Peterson, J.: Observations and modeling of background seismic noise, US Geol Surv Open-File Rept, 93, 1993.

529 Picozzi, M., Elia, L., Pesaresi, D., Zollo, A., Mucciarelli, M., Gosar, A., Lenhardt, W., and Živčić, M.: Trans-national
530 earthquake early warning (EEW) in north-eastern Italy, Slovenia and Austria: first experience with PRESTo at the CE 3 RN
531 network, Adv. Geosci., 40, 51–61, https://doi.org/doi:10.5194/adgeo-40-51-2015, 2015.

532 Pignatelli, A., D'Ajello Caracciolo, F., and Console, R.: Automatic inspection and analysis of digital waveform images by
533 means of convolutional neural networks, J. Seismol., 25, 1347–1359, http://dx.doi.org/10.1007/s10950-021-10055-8, 2021.

534 Pondrelli, S., Di Luccio, F., Scognamiglio, L., Molinari, I., Salimbeni, S., D'Alessandro, A., and Danecek, P.: The first very

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions

broadband Mediterranean network: 30 yr of data and seismological research, Seismol. Res. Lett., 91, 787–802, http://dx.doi.org/10.1785/0220190195, 2020.

Rastin, S. J., Unsworth, C. P., Gledhill, K. R., and McNamara, D. E.: A detailed noise characterization and sensor evaluation of the North Island of New Zealand using the PQLX data quality control system, Bull. Seismol. Soc. Am., 102, 98–113, http://dx.doi.org/10.1785/0120110064, 2012.

Rawat, W. and Wang, Z.: Deep convolutional neural networks for image classification: A comprehensive review, Neural Comput., 29, 2352–2449, http://dx.doi.org/10.1162/neco_a_00990, 2017.

Scales, J. A. and Snieder, R.: What is noise?, Geophysics, 63, 1122–1124, http://dx.doi.org/10.1190/1.1444411, 1998.

Scott, Halbert: Appendix A, SEED reference manual, Incorporated Research Institutions for Seismology, n.d.

Sleeman, R., Van Wettum, A., and Trampert, J.: Three-channel correlation analysis: A new technique to measure instrumental noise of digitizers and seismic sensors, Bull. Seismol. Soc. Am., 96, 258–271, http://dx.doi.org/10.1785/0120050032, 2006.

Sorrells, G. G.: A preliminary investigation into the relationship between long-period seismic noise and local fluctuations in the atmospheric pressure field, Geophys. J. Int., 26, 71–82, http://dx.doi.org/10.1111/j.1365-246X.1971.tb03383.x, 1971.

Stutzmann, E., Roult, G., and Astiz, L.: GEOSCOPE station noise levels, Bull. Seismol. Soc. Am., 90, 690–701, http://dx.doi.org/10.1785/0119990025, 2000.

Thorp, J., Davies, K., Bluteau, J., and Hoiles, P.: Implementation of seismic data quality characterisation using supervised deep learning, APPEA J., 60, 784–788, http://dx.doi.org/10.1071/AJ19040, 2020.

Vassallo, M., Satriano, C., and Lomax, A.: Automatic picker developments and optimization: A strategy for improving the performances of automatic phase pickers, Seismol. Res. Lett., 83, 541, http://dx.doi.org/10.1785/gssrl.83.3.541, 2012.

Wielandt, E.: Seismometry in: IASPEI International Handbook of Earthquake and Engineering Seismology, Academic Press., W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinge, 283–304 pp., 2003.

Wielandt, E.: Seismic sensors and their calibration, in: New Manual of Seismological Observatory Practice 2 (NMSOP-2), Deutsches GeoForschungsZentrum GFZ, 1–51, 2012.

Wielandt, E. and Forbriger, T.: Near-field seismic displacement and tilt associated with the explosive activity of Stromboli, Ann. Geophys., 42, http://dx.doi.org/10.4401/ag-3723, 1999.

Wielandt, E. and Steim, J. M.: A digital very-broad-band seismograph., Ann Geophys Ser B, 4, 227–232, 1986.

Wielandt, E., Bormann, P., and Bribach, J.: New Manual of Seismological Observatory Practice (NMSOP). Chapter 5: Seismic sensors and their calibration, in: 62. Jahrestagung der Deutschen Geophysikalischen Gesellschaft (Hannover 2002), 2002.

5 Declarations

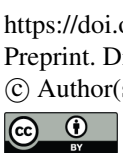
5.1 Funding: This work was partially supported by grant FISR 2021 Ricerca Libera (INGV, Istituto Nazionale di Geofisica e Vulcanologia).

5.2 Authors' contribution: all the authors have contributed to all the paper content parts and its drafting. Moreover, the manuscript is the result of continuous interaction, proposal and discussion among them. More specifically:
CP developed the idea, collected the data for the first and second test, decided the "human" interpretation for the third test data
PA has given most of his contribution in designing, tuning the neural network architecture and in coding the procedures to run them with data;

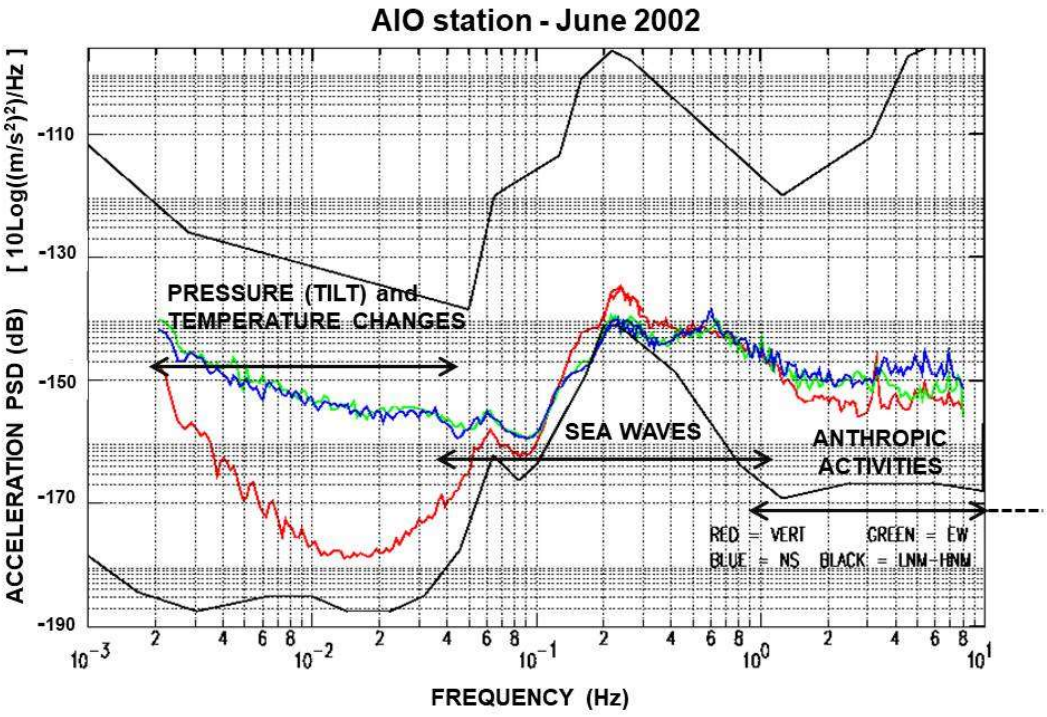
5.3 Competing interests: The authors declare no competing interests

5.4 Code/Data availability: Data used to get the shown results have been compressed and put at the following link: https://drive.google.com/drive/folders/1ncYig6TKZ1U8S404IIllIdO3vSrWWsS5L

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023



Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU



581
582 **Figure 1: Noise Spectra recorded in Antillo (ME, Italy) station (AIO code). Where thermal and electronic effects are negligible (in**
583 **this case for f > 0.04 Hz approximately), the colored curves represent Vertical (Red), N-S (Blue) and E-W (Green) components of**
584 **ground acceleration level spectrum at various frequencies. In black the empirical noise models of Peterson**

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

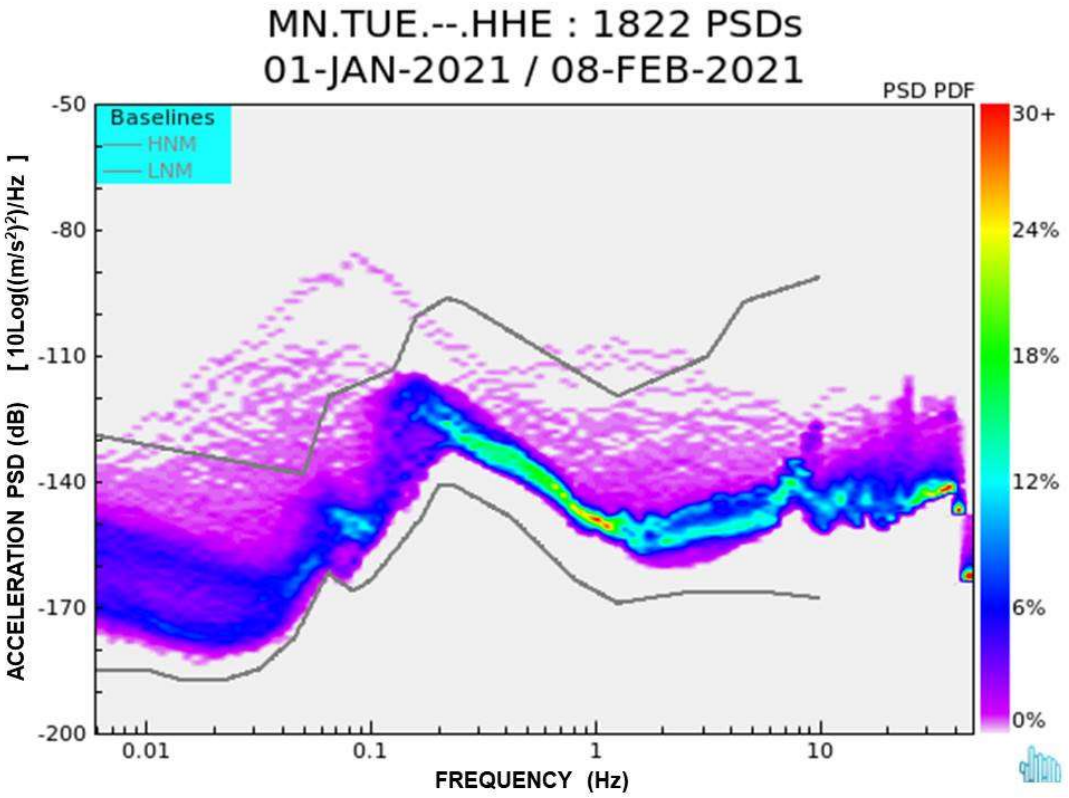


Figure 2: Seismic noise at TUE station (Central Alps) for the horizontal E-W components. The colored zone represents the Probability Density Function of the Power Spectral Density. The Peterson curves LNM and HNM are reported in black. The label "30+" refers to a probability greater than 30%. TUE station has a STS2 seismometer(Anon, 2020) and a Q330HR(Kinemetrics Inc., n.d.) digitiser/acquisitor. Instrument response is flat in velocity from about 0.01 to about 50 Hz(Pondrelli et al., 2020)

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions

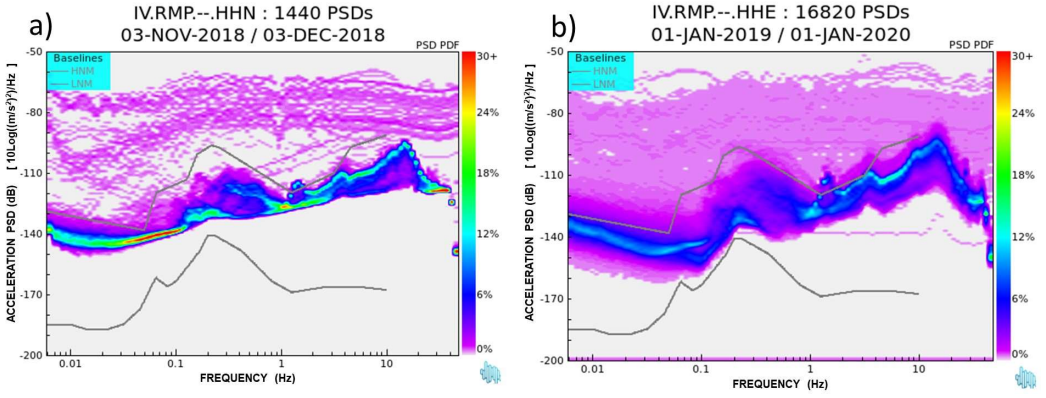


Figure 3: PDF noise level at the RMP seismic station (Monte Porzio, Rome) in the period 3 Nov - 3 Dec 2018 (a, before the intervention) and 1 jan - 18 Apr 2020 (b, after the intervention). On the left, the lower part of the colored area (statistically the most frequented) follows an sloped straight line almost everywhere, except for a very slight peak around 0.2Hz and at extreme frequencies. This abnormal behaviour indicates problems in the seismometer or in its positioning.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

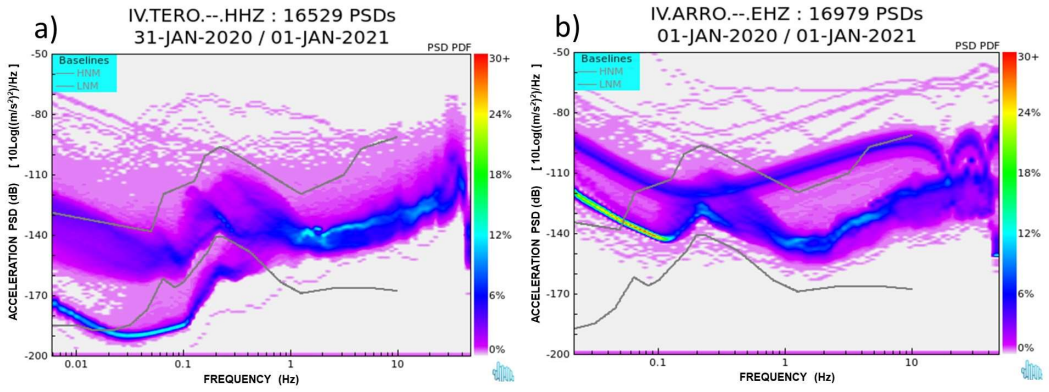


Figure 4 (a and b) a): Noise spectrum at TERO station (Teramo, Central Italy). The station consists of a BB seismometer, Trillium 40 s (http://support.nanometrics.ca/) and a GAIA digitizer/acquisition system(Michelini et al., 2016; Cirrus Logic Inc., n.d.) . At medium and low-medium frequency the noise is much lower (about 20 dB) than the LNM curve. b): Noise at the ARRO station (Arrone, Terni, Central Italy). The station consists of a Short Period seismometer (https://www.lennartz-electronic.de/wp-content/uploads/2021/04/Lennartz-SeismometerManual.pdf) and a GAIA. The bimodal trend on the vertical component denotes a malfunction (sometimes) of the seismometer.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

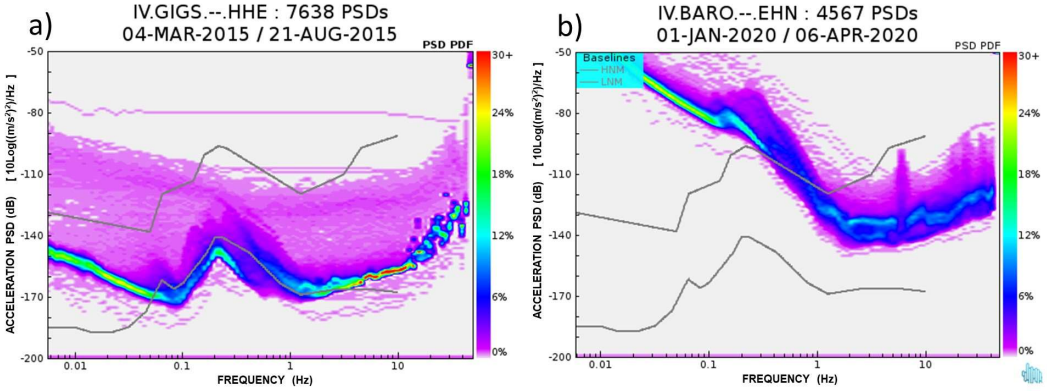


Figure 5 (a and b) a): noise at the GIGS station (Gran Sasso, central Italy) for more than 5 months time interval. The VBB station consists of a Trillium 240s seismometer (http://support.nanometrics.ca/) and a GAIA digitizer/acquisition system(Michelini et al., 2016; Cirrus Logic Inc., n.d.). The PDF shows a noise lower than LNM on about half of the analyzed band b): Noise at the BARO station (Barbarano, VT, Central Italy) with a Short Period *Tellus 5s* seismometer (https://lunitek.it/seismic/seismic-sensors/tellus-5s/) and a GAIA(Michelini et al., 2016; Cirrus Logic Inc., n.d.). The shape of the spectrum is very different from the typical trend of a 5 s seismometer which instead follows the lower blue curve of Figure 4b.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions

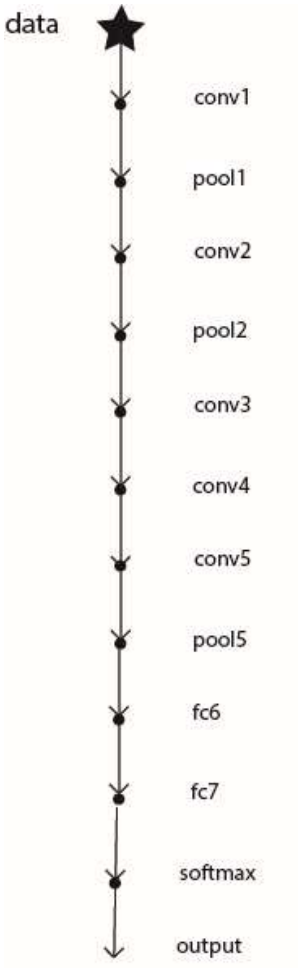


Figure 6 Alexnet classification architecture

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

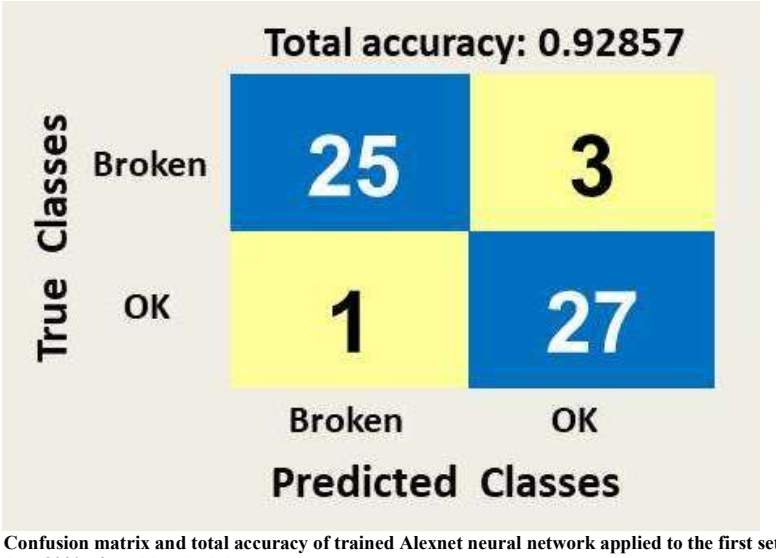


Figure 7: Confusion matrix and total accuracy of trained Alexnet neural network applied to the first set of test data. The test is performed on 20% of total labelled data.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

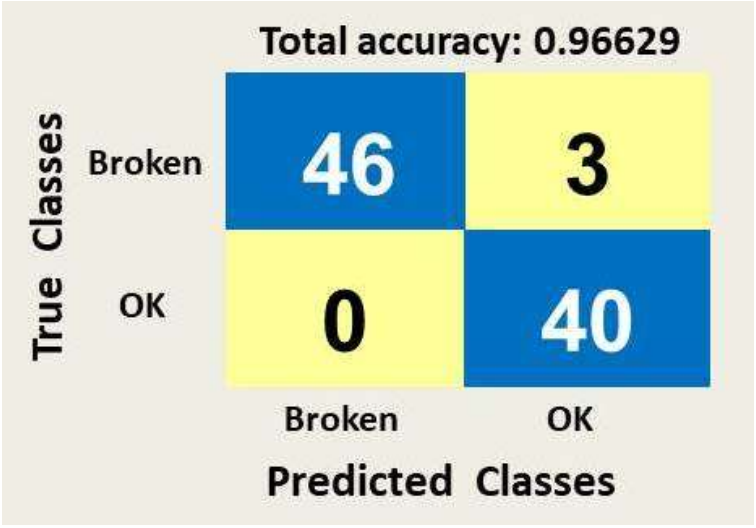


Figure 8: Confusion matrix and total accuracy of trained Alexnet neural network relative to second learning test. The test is performed on 20% of total labelled data.

https://doi.org/10.5194/gi-2023-4
Preprint. Discussion started: 8 May 2023

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

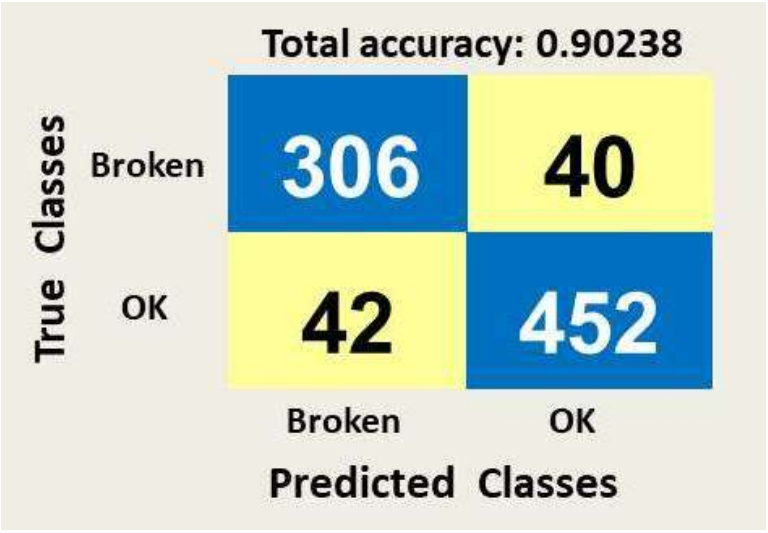


Figure 9 Confusion Matrix for the test on 840 diagrams. The main diagonal indicates the successes.

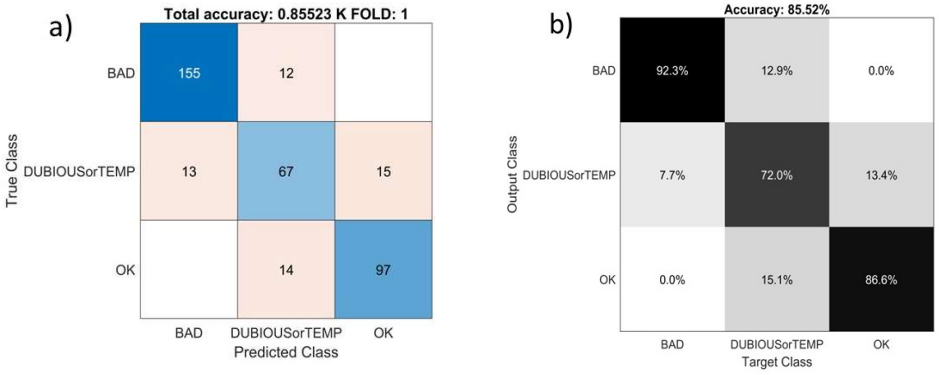


Figure 10 a) Confusion Matrix for the test on 373 diagrams. (The remaining 1492 of 1865 diagrams were used for the training). The training diagrams have been divided into three classes. The main diagonal indicates the successes. b) Percentage of successes and failures of the previous 3 x 3 matrix (three classes)