Transactions, **SMiRT-26**
Berlin/Potsdam, Germany, July 10-15, 2022
Division IV

HARD ROCK SITE RESPONSE VALIDATION WITH EMPIRICAL DATA

Philippe Renault[1]

[1] Managing Director, swissnuclear, Olten, Switzerland (philippe.renault@swissnuclear.ch)

ABSTRACT

The operating Swiss nuclear power plants (NPP) are located in the Alpine Foreland and sitting on hard to very hard rock (Vs = 1500-2500 m/s). Since around 2015 a special seismic monitoring network for research on ground motion at surface and deep boreholes has been installed at the NPP sites. One of the boreholes has a depth of 670 m and is currently the deepest borehole in Europe equipped with an accelerometer. The other two boreholes have a depth of 120 and 160 m. Over the course of the years a relevant amount of data has been collected, which now allows a representative evaluation of various ground motion and site parameters.

The collection of ground-motion data has now been used for the detailed investigation of site response and other site-specific parameters such as transfer function between rock and surface, V/H on the surface and depth and kappa (attenuation in the high-frequency range). The statistical evaluation of those parameters allows a realistic quantification of the mean values and associated variability. The empirical results are compared to the models based on expert judgment and numerical simulations. The empirical evidence and associated challenges are discussed in order to guide the analyst in making decisions for structural assessments and soil-structure interaction analyses. The comparison and their interpretation allow to draw conclusions on the shear-wave velocity profiles at the sites and the range of former decisions.

INTRODUCTION

A total of approximately 1400 earthquakes were recorded in the year 2020 in Switzerland and its neighbouring countries. Today, the Swiss Seismological Service uses modern seismic instruments to record usually an average of two earthquakes a day in Switzerland and neighbouring countries. For this the national network uses more than 200 seismic stations. On average, 10 to 15 of these events can be strong enough (magnitude of 2.5 or more) to be felt by the local population. On 25 October, Switzerland's largest earthquake of the year 2020 was recorded with a magnitude of 4.3, which was felt as far away as Ticino (Italy) and Lake Constance (Germany).

The monitoring and determination of the ground motion caused by earthquakes is important for the determination of the earthquake risk at any site of interest, but especially at sites with critical infrastructures. For this reason, an extended earthquake monitoring system was installed in 2014-2016 at the sites of the Swiss nuclear power plants Beznau (KKB), Gösgen (KKG) and Leibstadt (KKL); see Figure 1. The so-called ERBIUM project. The system at each site consists of a triaxial surface acceleration sensor and one or two triaxial acceleration sensors installed at different depths ranging from 120 m to 670 m depending on the location. The borehole of 670 m at the site of the Gösgen NPP is currently the deepest borehole in Europe equipped with an accelerometer. The seismic instruments record continuously (24 hours a day) with a sampling rate of 250 sps and thus, allows a spectral resolution up to 100 Hz.

The seismic instrumentation of critical infrastructures such as nuclear power plants worldwide is regulated and has often been based in the early 70's on the ANSI/ANS-2.2 and US NRC Regulatory Guide 1.12. Most countries have meanwhile formulated their own national requirements for seismic instrumentation. The recent major earthquakes, such as the Japanese NGO Earthquake, have prompted the authorities to propose an extension of existing seismic instrumentation with modern technologies (see, for example, draft of the US Guideline DG-1332.

Beside the recorded ground motion itself, site-specific data help to predict the seismic hazard more accurately and to improve the modelling of site-effects. On the other hand, measurements during and after an earthquake are important for decisions making, such as the continuation of the operation or a shutdown/restart of the facility. As data is the key to improve the knowledge about ground motions and reduce uncertainties the Swiss utilities invested in a long-term monitoring network at the sites (see Renault et al. 2018 for a more detailed description).

Figure 1. Construction work for drilling the boreholes and visible casing for one borehole in the front before shortening to ground level.

COLLECTED DATA

The stations of the network are running since 2016 in a viable and automated way. Since then, the data has been collected by swissnuclear and compiled in a database at the "Olten Data Centre" (ODC). The data collection is of course ongoing, but now, after several years enough data is available to start systematic evaluations and compute statistics of the recordings. In total 2266 events have been recorded at all three sites together until the end of 2019. Figure 2 provides an overview over time at the three locations. As can be seen, the sites KKL and KKG are collecting a significant amount of event data and in a more or less consistent way. The distance between both sites is approximately 30 km. On the other hand, the site KKB is recording a lower number of events. KBB and KKL are only approximately 6 km apart. The issues related to the data transmission causing missing records have been resolved.

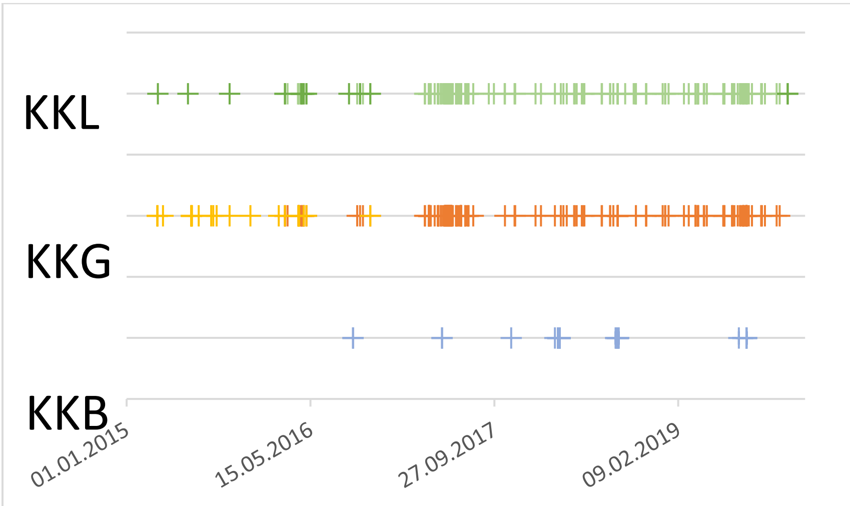

Figure 2. Overview of recorded events until end 2019.

Comparison with national earthquake database

The Swiss Seismological Service (SED) in charge of the national network is maintaining a database of all recorded earthquakes in Switzerland and neighbouring countries. As a first assessment of the data recorded at the NPP stations a comparison with the SED database is performed based on the date-time stamp. The comparison is down in an automatic PowerShell-Script and connects to the database via http://seismo.ethz.ch/de/earthquakes/switzerland/all-earthquakes/ and parsing the html to compare the events based on the date and time. In case of a match this is added to the list and the output is a CSV file of recognized events. The so far matched events are listed in Table 1. There are 16 until end of 2019 and their magnitudes (Mw) range from 0.9 to 4.6. Almost all were recorded at the KKG and KKL site, only 6 of them also at KKB. But it should be noted that out of those only one was automatically identified as event by the recorder. The others were manually retrieved and processed from the continuous recordings at the KKB site.

Table 1: List of earthquakes recorded at the NPP sites with matching identification in SED database.

Date (local)	**Time (local)**	**Location**	**Magnitude**	**Depth [km]**	**Lat/Lon**
04.11.2019	00:59	Albstadt (D)	3.9	6.6	48.24 / 9.00
29.08.2019	14:22	Konstanz (D)	3.4	3.1	47.74 / 9.11
10.08.2019	03:17	Bellinzona (IT)	1.6	16	46.20 / 8.92
30.07.2019	00:42	Konstanz (D)	3.2	3.9	47.74 / 9.11
29.07.2019	23:17	Konstanz (D)	3.7	3.7	47.74 / 9.11
04.05.2018	21:36	Muellheim (D)	3.3	15.3	47.77 / 7.52
11.03.2018	23:29	Laufenburg	3.1	17.3	47.67 / 8.01
01.02.2018	01:47	Montafon (A)	4.1	1	47.15 / 10.00
17.01.2018	19:07	Montafon (A)	4.1	1	47.15 / 9.99

21.11.2017	09:22	Zug	3.3	31.7	47.15 / 8.55
01.06.2017	00:56	Muellheim (D)	0.9	15.1	47.76 / 7.76
17.05.2017	09:57	Imst A	1.2	-0.9	47.39 / 10.60
21.11.2017	10:22	Zug	3.3	31.3	47.15/8.55
23.03.2017	15:52	Laufenburg	1.3	-0.6	46.91/8.93
06.03.2017	21:20	Linthal - Glarus	4.6	5	46.91/8.94
11.04.2016	10:47	Poschiavo (GR)	3.2	7.6	46.43 / 10.02

Implementation of automatic data completeness checks

Given the issues with the internet connection and transfer of data a concept and automatic process was implemented to check the completeness of records. Furthermore, a permanent state-of-health monitoring system was installed in form of a screen display in the offices of swissnuclear in order to immediately see if a sensor is deficient (traffic light type, see Figure 3).

As there are four sets of data reaching the ODC and need to be stored, adequate approaches had to be developed in order to account for the different types of data and interval. The data recorders provide continuous recordings (24/7) in MiniSeed format, event files in case the trigger threshold is exceeded and once a day a state-of-health (SOH) file. The event files are made available in a binary format (EVT) and ASCII format.

The four different data types are arriving in the ODC server and a VBA program has been developed to check on a daily basis if the hourly transmitted continuous recordings as MiniSeed files are complete. Also, there is a daily check of the SOH files. The program can also be used to check the data completeness over a certain time period, e.g. a full year. The advantage of such a continuous monitoring of the files is also that data completeness over time can be visualized in form of "sensor/channel availability" for all 21 channels and maintenance down times can be tagged. The program also takes care of renaming the folders and files in a consistent way on the ODC, while it checks for the completeness.

As the event files are only generated and transmitted in case of a triggering, no daily check is performed. Beside the transmitted ASCII and EVT files in case of an event the ODC processes them to generate a ZIP-archive with ASCII files for the raw data, processed accelerations, velocities, displacements, FFT, response spectra and spectral amplification from downhole to surface. The self-developed program has been extended to check the consistency of those three file types which should always exist for a specific event and can be converted into each other. In the past this consistency check was not done and after the effort to try to recover some events for the KKB site it was noticed that not in all cases all three types were available in the database. The reason for this is unknown.

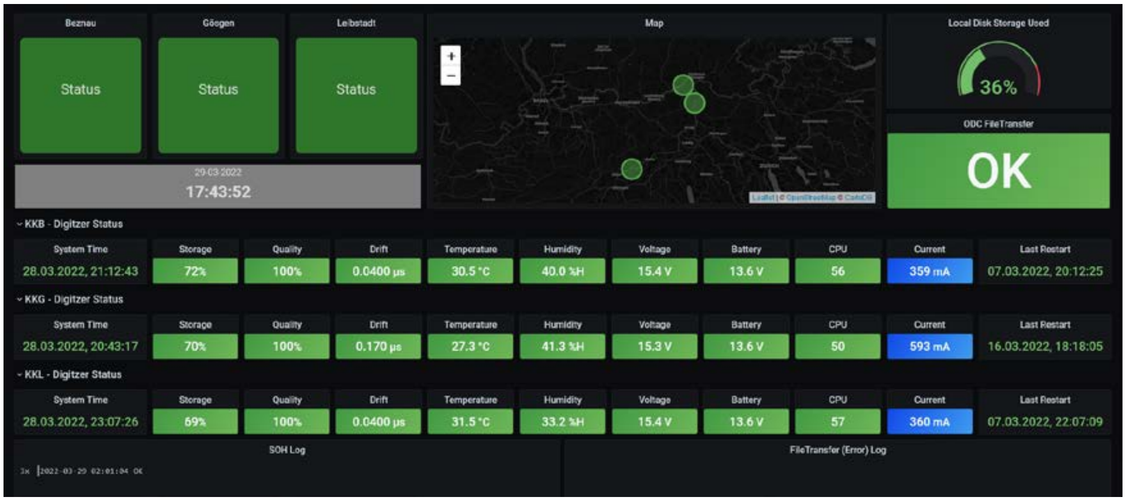

Figure 3. Overview dashboard for seismic monitoring stations at the ODC.

DATA EVALUATION

After reviewing the full database of records collected so far it became very quickly evident that a large amount of "events" were fake events and not ground motions due to earthquakes. Many could be immediately identified as man-made (e.g. construction work, single spikes). Even though this was expected for the surface sensor (as sitting next to an industrial facility with running machines and day/night activity), it was not to that extent for the borehole stations.

Given the large amount of data to process (of course not in the order of magnitude as "big data" or the one which has to be dealt with when operating a national network) and from the practical point of view of an end user who wants to make use of the data and not research, a solution was necessary. As discussed above, the first attempt was the comparison of events with the national database, but given the fact that locally at each site there was much more data available –from very small magnitude events – the goal was to process also the additional data with no direct association with an official earthquake. But a manual screening and processing is for a utility neither practical, nor economic. Thus, we looked for an automatic way and advancements in the machine learning field. The methods and practical application examples have significantly increased in the last couple of years (Kong et al. 2019, Jiao et al. 2020).

Machine-learning approach for identifying records for further processing

It was decided to make use of a basic machine-learning approach using principal component analysis (PCA) and compare three machine-learning algorithms: support vector machine (SVM), K-Nearest Neighbours (KNN) and Random Forest (RF). For the model evaluation the dataset was randomly divided into a training, test and validation set. For the feature extracting mainly the Fourier transformation and spectrogram were used. As programming language Python was used, as all necessary machine-learning libraries exist as add-on packages and it can be integrated easily in any OS environment.

One of the challenges was to compile the set of true seismic signals for the training algorithm, as for example at one site one of the two sensors might have recorded a useful seismogram, but the other not. To complicate the task, one had to also identify if all three directions for the same event have a meaningful signal. One example of a case where surface and borehole sensor of the same event did not provide in all directions signals to be used for training is depicted in Figure 4. It is evident that the surface signal has a valid and useful signal where the P an S wave can be even distinguished. At depth the signal is more like noise and would not be suitable for a training algorithm. After preparation of the dataset for “typical” ground motion recordings, also a dataset for the “not an earthquake signal” is necessary. This should be the rest of the remaining recordings and is mainly a noise signal. But within those also some recordings had a very special shape which needed to be cleaned from the dataset.

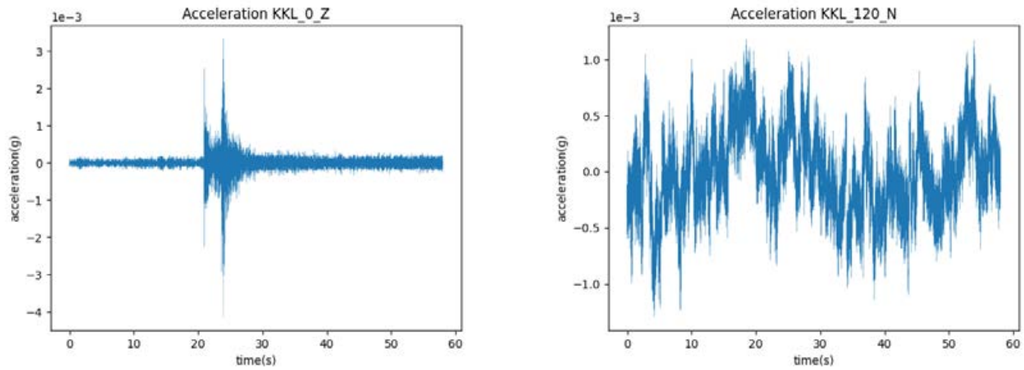

Figure 4: Recording of 2018-03-11, 23:29:06 at KKL site at surface (left) and depth (right).

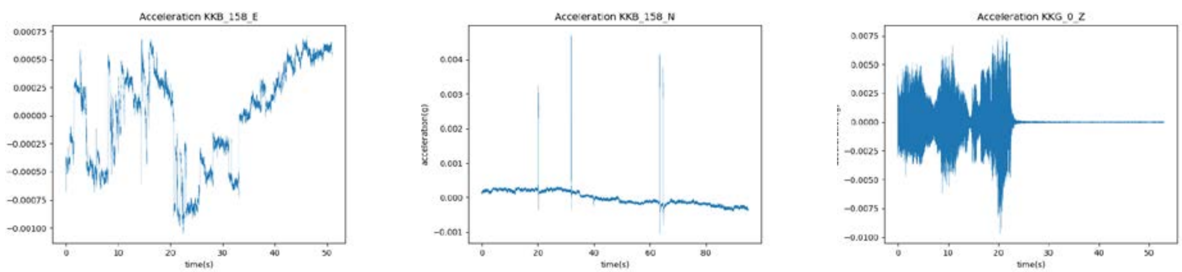

Figure 5: Examples of non-earthquake type recordings with special shape.

For models using K-Nearest Neighbours or Random Forest as classification algorithm, feature combinations with a PCA reduction generally perform better. The dimension reduction by a PCA has besides the effect of a reduction of the number of dimensions also the property that resulting Principle Components are independent from each other. This has a positive effect on the classification algorithms, since both benefit from an independent feature set.

After successful training the machine-learning model the program was applied to the ODC database. At the end 2255 events constituted the database. From those, 11 events were classified by the algorithm as “earthquakes”, 303 were classified as “maybe earthquakes” and the remaining 1941 as “non earthquake”. More details and a discussion of the developed model can be found in Sussmann et al. (2020).

Thoughts on improvements of the implemented approach are: In the current implementation, a site is defined by configurable limits. However, the accelerograms of a site differ in depth and orientation. For example, it could be argued that a classified signal registered by a downhole sensor is a stronger indication of an earthquake than a signal from a sensor at the surface. Additionally, an event registered by several sites is more likely to represent an earthquake than a event recorded at only one site. With sufficient data, this additional information could be represented by an enhanced machine-learning model. Furthermore, the issue of up- and down travelling waves in the casing of the borehole will be investigated.

Statistical evaluation of identified event records

Based on the results of the procedure described above, a total number of 394 accelerograms were retained for further processing and statistical evaluation. Those included recordings of the 11 clearly identified earthquake events and the “maybe earthquake” set at all three sites and at different depth levels. Nevertheless, a visual check of the 394 accelerograms revealed that only 150 accelerograms were satisfying the quality assurance parameters (as e.g. signal-to-noise level). Within those are 71 accelerograms for the site of KKG, 77 accelerograms for the site of KKL and only 2 for KKB. Thus, for KKB so far, no robust statistical evaluation is available.

For the systematic processing and subsequent statistical analysis, a program was developed in Matlab. This software was used mainly because its capability to easily produce figures (necessary for the visualization of the results) and generate a PDF report based on LaTex outputs generated by Matlab. It is intended to use and implement this code in an automatic routine to process new records automatically and update the statistical evaluation.

Starting from the accelerograms, a baseline correction is performed and the velocity and displacement time histories for all three directions are derived by integrating the corrected accelerations. While the accelerations are in g, the velocities and displacements are in cm/s and cm. Afterwards, Fourier spectra and response spectra (5% damping) are calculated for the three directions. Additionally, the geometrical mean of the horizontal response spectra is calculated and the ratio of the vertical direction to the horizontal mean is computed. Furthermore, the signal-to-noise ratio (SNR) is determined for the three directions Z, N and E.

Using the Fourier spectra, the empirical transfer functions for all three directions are derived by means of the ratio of Fourier spectra between the ground level and rock level. At the site KKG, the transfer function is additionally calculated between the ground level and the depth of 350 m as well as between the latter and the lowest level (669 m). The spectral amplification is estimated using the response spectra of the different levels. Though, the same combinations are investigated as in the case of the transfer functions.

Finally, several parameters are derived for every input signal. These parameters consist of peak ground acceleration (PGA), peak ground velocity (PGV), peak ground displacement (PGD), strong motion duration for 5-75% as well as 5-95%, cumulative absolute velocity (CAV) and the value for kappa (κ), which is a parameter typically used to quantify the high-frequency attenuation at a site.

Using the processed data for every single event, a statistical evaluation is performed for the different sites. Therefore, the mean value as well as 16%-fractile, 50%-fractile and 84%-fractile values are calculated for the site-specific response spectra and Fourier spectra as well as the transfer function and the spectral amplification. In addition, these statistical parameters are also calculated for the site-specific kappa values.

Before performing the actual statistical evaluation, all the accelerograms are checked once again, in order to investigate for every single input all the levels which may be used for the calculation of e.g. the site specific Fourier spectra or spectral amplification. In consequence, the list of events actually considered for the site-specific statistic may differ from the events evaluated independently.

Within the summary for the different sites, the plots of the different spectra as well as the transfer and amplification functions show the mean values as well as the 16%-fractile and 84%-fractile values, which are calculated by addition or subtraction of the standard deviation from the mean value. In the case of the 16%-fractile value, a special treatment is utilised, since mathematically negative values do not make sense physically and are not plottable in a lognormal plot. Therefore, the 16%-fractile curve in the plots corresponds to the maximum value of the regular 16%-fractile (mean value minus standard deviation) and the minimum value of all events considered for the calculation of the mean value.

RESULTS

In this contribution, for one of the sites the site effects (soil amplification, V/H ratio) are discussed. The site has a shallow soil layer of approx. 28 m and hard rock below (Figure 6. Expert models of shear-wave velocity profiles, soil stratigraphy and measured V_S profile for example site (left to right).Figure 6). At this site the surface and downhole measurements of ground motions are useful to compare and calibrate the theoretical inverted models (based on derived V_S profiles) with empirical measurements.

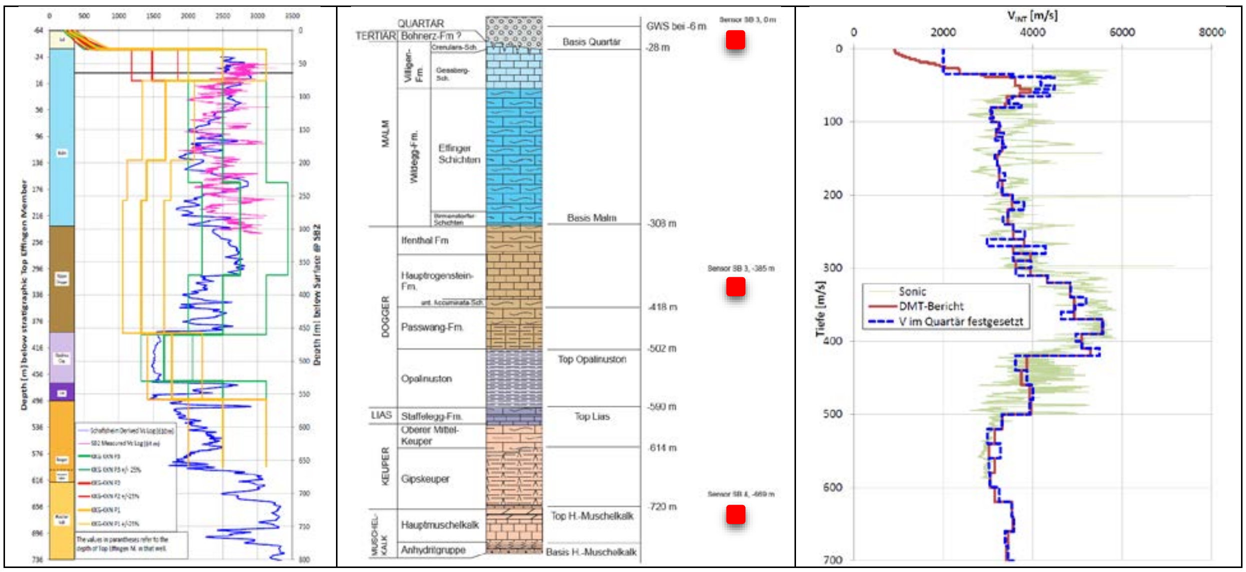

Figure 6. Expert models of shear-wave velocity profiles, soil stratigraphy and measured V_S profile for example site (left to right).

Figure 7 shows the overview of the statistical evaluation. In this example, it is interesting to note that the soil amplification between -699 to -350 m is the main contributor to the spectral amplification and from -350 m to surface there is only a minor increase. Furthermore, it is evident that the horizontal amplification is significantly larger than the vertical over the whole frequency range of relevance, as expected.

The empirical transfer function is also shown, but could be further processed with a smoothing filter by using e.g. a Konno-Ohmachi bandwith equal to 20. This would allow for a better comparison and an engineering perspective. The transfer function at very low frequencies (below ~3 Hz) seem doubtful with an increase of the amplification at low frequencies, likely due to poor SNR.

Comparing the empirical amplification function between -669 m and surface (0 m) for the horizontal component (graph in the first row and third column in Figure 7), the two peaks at around 4.5 and 11 Hz are consistent with the expert models developed for the NPP sites in the PRP project (swissnuclear 2013). The corresponding amplification function for low amplitude ground motions is shown in Figure 8. The amplitude of amplification in the relevant frequency domain is also consistent when comparing the mean of the empirical function with the expert model range.

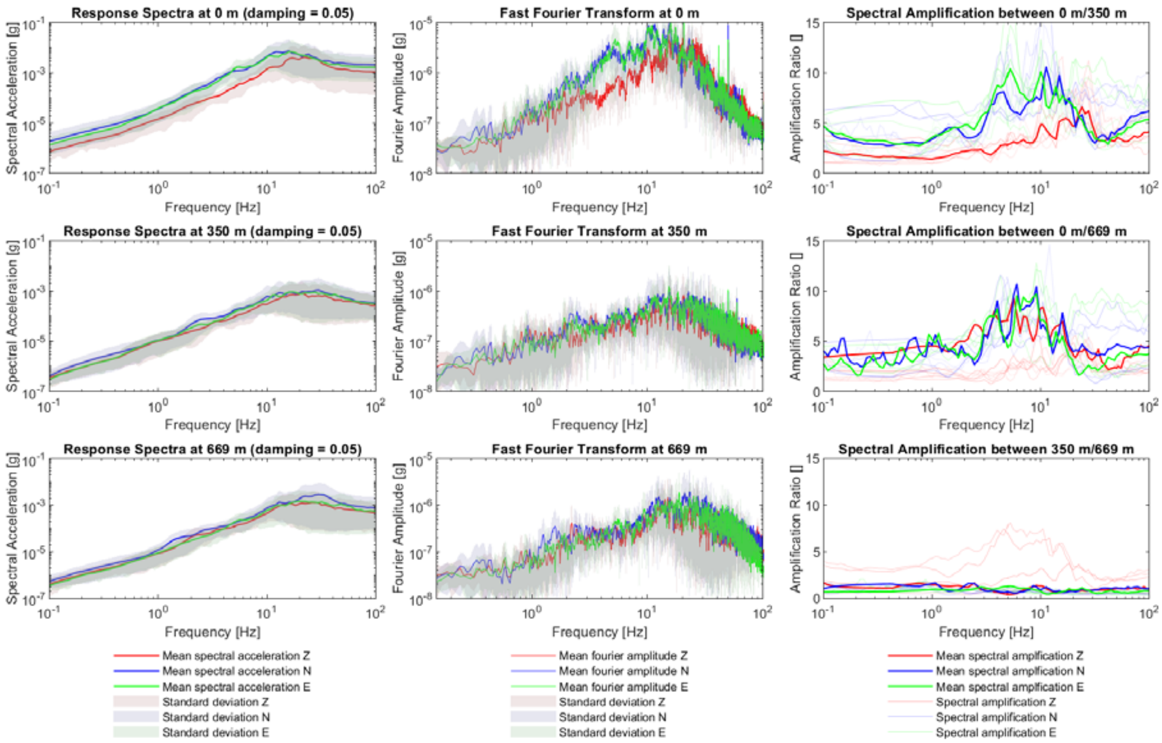

Figure 7. Statistical evaluation of response spectra (left column), Fourier spectra (center column) and amplification function (right column) for one site.

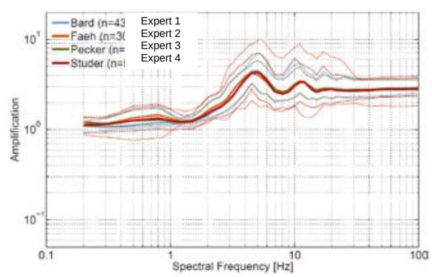

Figure 8. Simulation and expert based amplification models at surface and low PGA. Geom. mean horizontal component. Solid line represents the median models and thin lines min and max.

A direct comparison of the vertical amplification is not possible. Nevertheless, one can compare with the expert models based on the V/H ratio to be combined with the horizontal amplification. In Figure 9 the statistical evaluation of the collected empirical data with the expert models is shown. Also here there is a good consistency of the functions for the frequency range of interest. For very low frequencies below 1 Hz the scatter of the records is very large and its reason should be further investigated. The peak of the V/H function at around 20 Hz is also consistent. So far, the mean V/H ratio at PGA (here 100 Hz) is around 0.5, while the epistemic uncertainty of the expert models is quite large, but their median value is also supporting a factor of approx. 0.5.

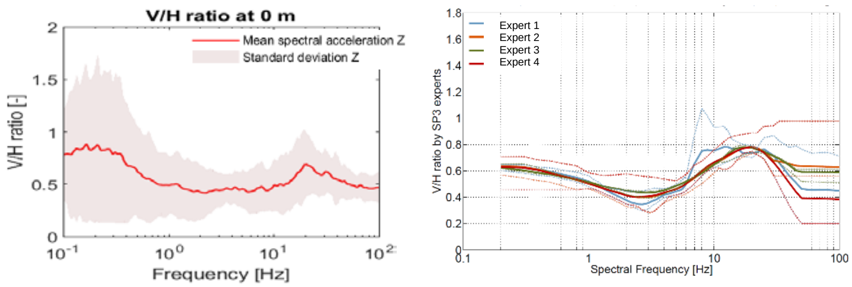

Figure 9. Empirical (left) vs. simulation and expert based (right) V/H models at surface and low PGA. and V/H ratio. On the right the solid line represents the median models and thin lines min and max.

CONCLUSION

From an end user and utility perspective it is a challenge to handle the hybrid datasets (continuous and event records), as the focus is clearly only on the real events, but the continuous data offers the possibility to process missed events or evaluate some ground motion parameters based on the noise. Thus, storage and processing need to be done for both. After some initial technical problems, the network and processing software is in place and operational. The amount of data collected in average per year is satisfying and promising. None of the data evaluated so far and conclusions derived from it is pointing to a safety issue and is consistent with previous assessments. Nevertheless, the robustness of the statistical evaluation can be and must be further improved and we are looking for more records in the next years to achieve this.

As it has been explained, the software environment used to collect, process and evaluate the data is very diverse and due to the historical step-by-step development. This is not ideal, as the underlying systems (e.g. Matlab) have dependencies, need to be maintained and can cause interface issues after e.g. an update. It is certainly a challenge to integrate everything in a fully automated and homogeneous IT environment, but requires an effort and is not a turn-key product.

REFERENCES

American National Standard Institute (1978): ANSI/ANS-2.2-1978 - Earthquake Instrumentation Criteria for Nuclear Power Plants.

Jiao P, Alavi AH (2020): *Artificial intelligence in seismology: Advent, performance and future trends*. Geoscience Frontiers, 11 (3), 739-744.

Kong Q, Trugman DT, Ross ZE, Bianco MJ, Meade BJ, Gerstoft P (2019): *Machine Learning in Seismology: Turning Data into Insights*. Seismological Research Letters, 90(1), 3-14.

Renault P, Dalguer LA, Skolnik D (2018): *Monitoring of ground motions with surface and deep borehole instrumentation at Swiss NPP sites.* 16th European Conf. on Earthquake Engineering, Thessaloniki, Greece.

Sussmann H, Schürch R, Graber M (2020): *Klassifizierung von Erdbeben aufgrund von seismologischen Daten*. Technical Report by Fachhochschule Nordwestschweiz, Brugg-Windisch

swissnuclear (2013): *Probabilistic Seismic Hazard Analysis for Swiss Nuclear Power Plant Sites - PEGASOS Refinement Project.* Final Report, Vol. 1-5. URL: www.swissnuclear.ch/en/downloads.html

US NRC (1997): Nuclear Power Plant Instrumentation for Earthquakes, Regulatory Guide 1.12, Rev. 2. U.S. Nuclear Regulatory Commission.

US NRC (2016): Nuclear Power Plant Instrumentation for Earthquakes. Draft Regulatory Guide DG-1332, Proposed Revision 3 of Regulatory Guide 1.12, dated March 1997. U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Division 1.