

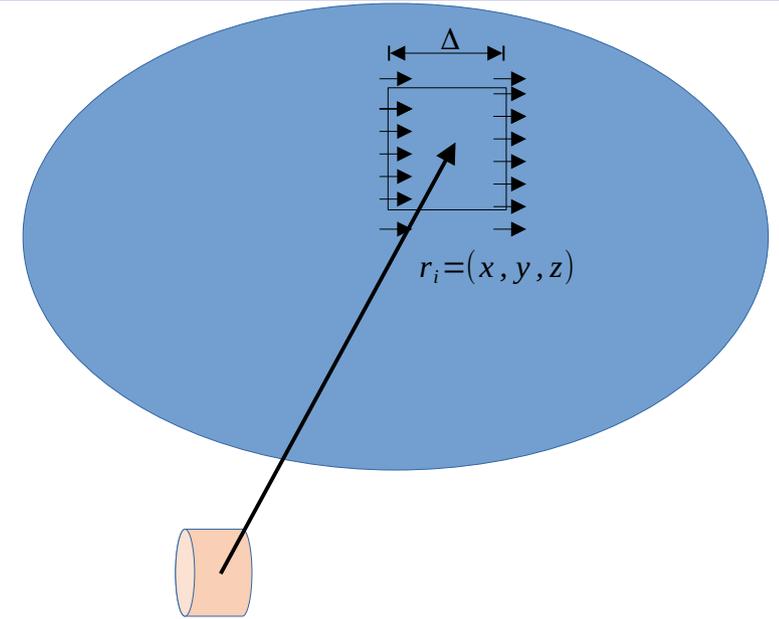
A study of deep neural networks for Newtonian noise subtraction



Vincent van Beveren, Maria Bader, Jo van den Brand, Henk Jan Bulten, Xander Campman en Soumen Koley,
for the ETpathfinder collaboration
henkjan@nikhef.nl

Newtonian noise

- Newtonian noise: seismic displacements may cause density fluctuations in the soil. These fluctuations cause a change in gravitational attraction.
- For a small volume element we have that the mass fluctuations are due to compression of p-waves; **the mass fluctuation is a scalar object** and the **resulting force is in the direction of the volume element, not in the direction of the displacement field**.



Newtonian noise

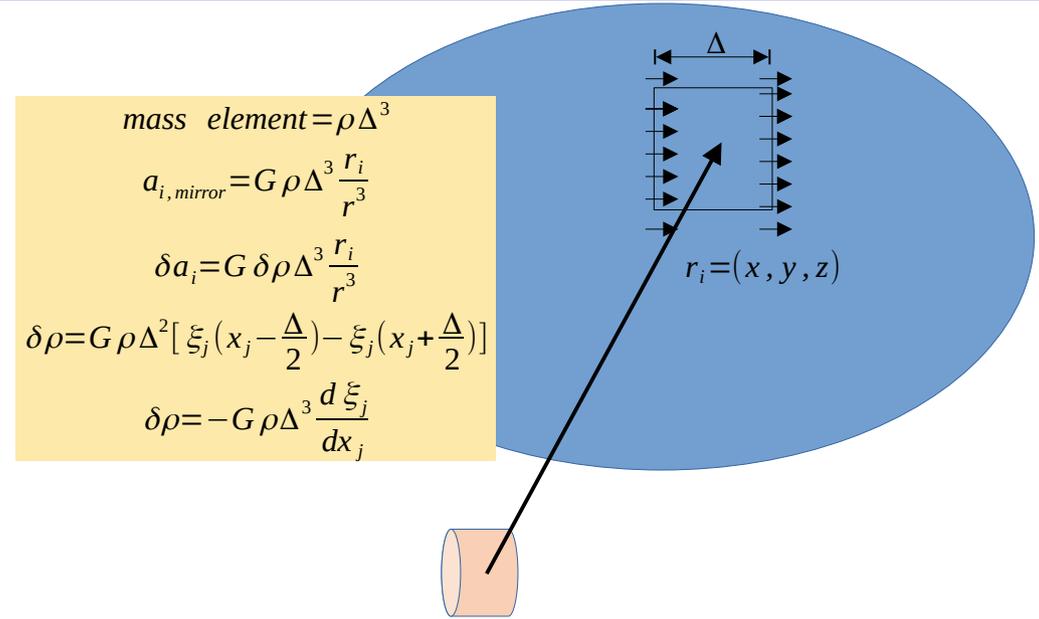
- Newtonian noise: seismic displacements may cause density fluctuations in the soil. These fluctuations cause a change in gravitational attraction.
- For a small volume element we have that the mass fluctuations are due to compression of p-waves; **the mass fluctuation is a scalar object** and the **resulting force is in the direction of the volume element, not in the direction of the displacement field**.
- One can integrate over the total volume inside a certain integration radius to obtain:

$$\delta a_{i,vol} = -G \rho \int_V \frac{r_i}{r^3} \frac{d\xi_j}{dx_j} dV$$

for surface element: $\delta\rho = (\rho - \rho_{out}) \xi_j \hat{n}_j dS$

$$\delta a_{i,surface} = G (\rho - \rho_{out}) \int_S \frac{r_i}{r^3} \xi_j \hat{n}_j dS$$

- Here the surface term takes into account that if there is a difference in density at the integration boundary, the change of mass in that element is modified.



Newtonian noise

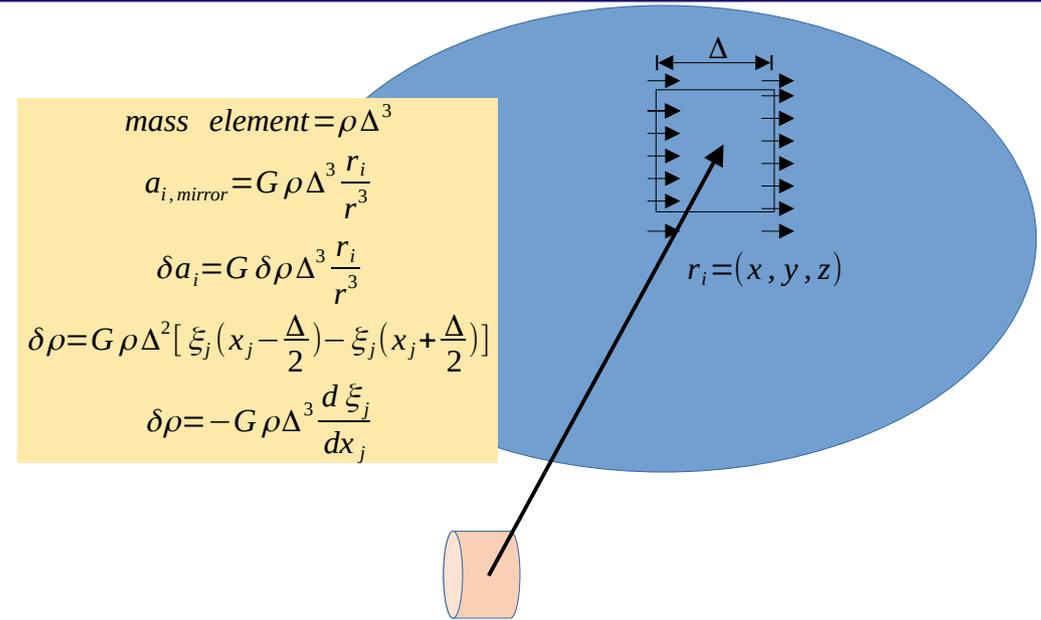
- Newtonian noise: seismic displacements may cause density fluctuations in the soil. These fluctuations cause a change in gravitational attraction.
- For a small volume element we have that the mass fluctuations are due to compression of p-waves; **the mass fluctuation is a scalar object** and the **resulting force is in the direction of the volume element, not in the direction of the displacement field.**
- One can integrate over the total volume inside a certain integration radius to obtain:

$$\delta a_{i,\text{vol}} = -G \rho \int_V \frac{r_i}{r^3} \frac{d \xi_j}{dx_j} dV$$

for surface element: $\delta \rho = (\rho - \rho_{\text{out}}) \xi_j \hat{n}_j dS$

$$\delta a_{i,\text{surface}} = G (\rho - \rho_{\text{out}}) \int_S \frac{r_i}{r^3} \xi_j \hat{n}_j dS$$

- Here the surface term takes into account that if there is a difference in density at the integration boundary, the change of mass in that element is modified.
- Since one measures ξ instead of $d\xi_j/dx_j$ one uses the divergence theorem, to arrive at an integral over the gravity gradients instead of the displacement field gradients:
- Note that the surface term when one integrates over gravity gradients only should be applied for ρ_{out} ; else you double-count.



Divergence theorem:

$$\int_V \frac{d}{dx_j} \left[\frac{r_i}{r^3} \xi_j \right] dV = \int_S \frac{r_i}{r^3} (\xi_j \hat{n}_j) dS$$

$$\int_V \left[\frac{d}{dx_j} \frac{r_i}{r^3} \right] \xi_j dV = \int_S \frac{r_i}{r^3} (\xi_j \hat{n}_j) dS - \int_V \frac{r_i}{r^3} \frac{d \xi_j}{dx_j} dV$$

$$\delta a_i = G \rho \int_V \left[\frac{d}{dx_j} \frac{r_i}{r^3} \right] \xi_j dV - G \rho_{\text{out}} \int_S \frac{r_i}{r^3} [\xi_j \hat{n}_j] dS = \dots$$

$$\dots = -G \rho \int_V \frac{r_i}{r^3} \frac{d \xi_j}{dx_j} dV + G (\rho - \rho_{\text{out}}) \int_S \frac{r_i}{r^3} [\xi_j \hat{n}_j] dS$$

Seismic fields at the surface in the EU Meuse-Rhine region

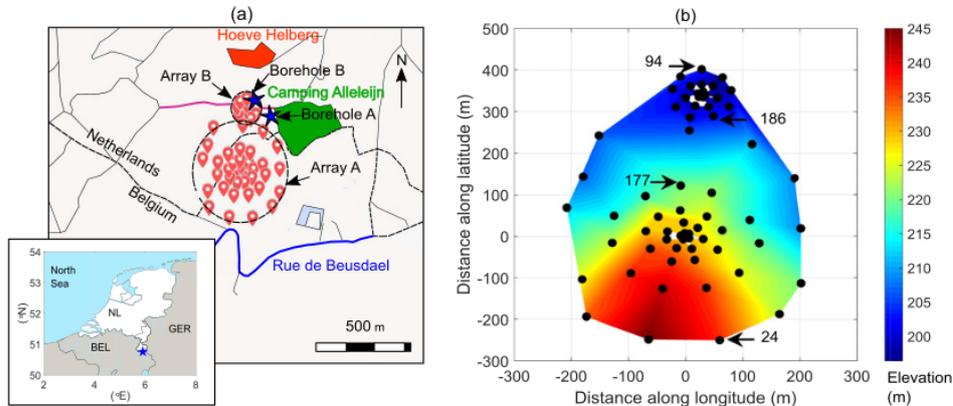


Figure 2. (a) Layout of array A and array B showing seismometers laid out on rings of different radii on a map of Terziet, Limburg. The two boreholes at the site are marked with stars. Map data ©2020 Google. (b) Sensor geometry shown on a Cartesian coordinate system and overlaid on the topography at the site. Black solid dots show four seismometers along with their sensor ID; their noise power spectral densities are shown in figure 3(a).

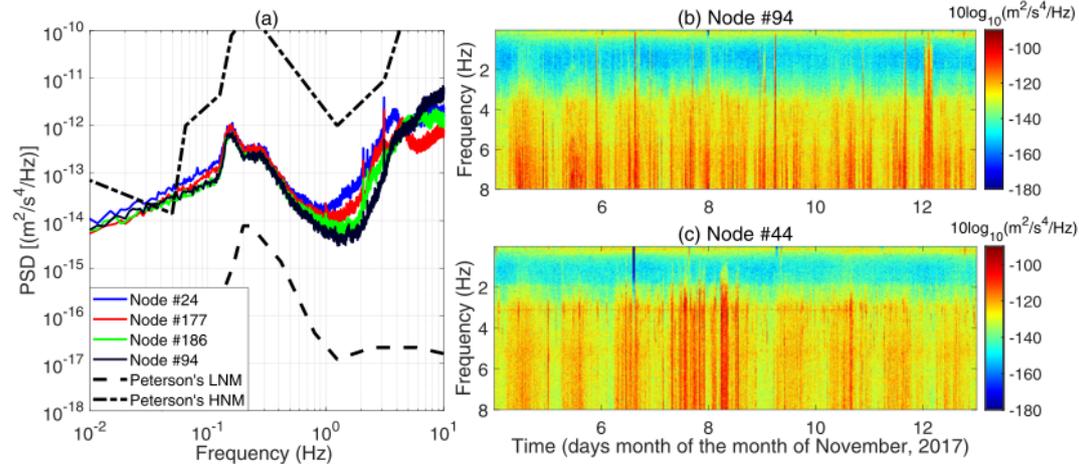


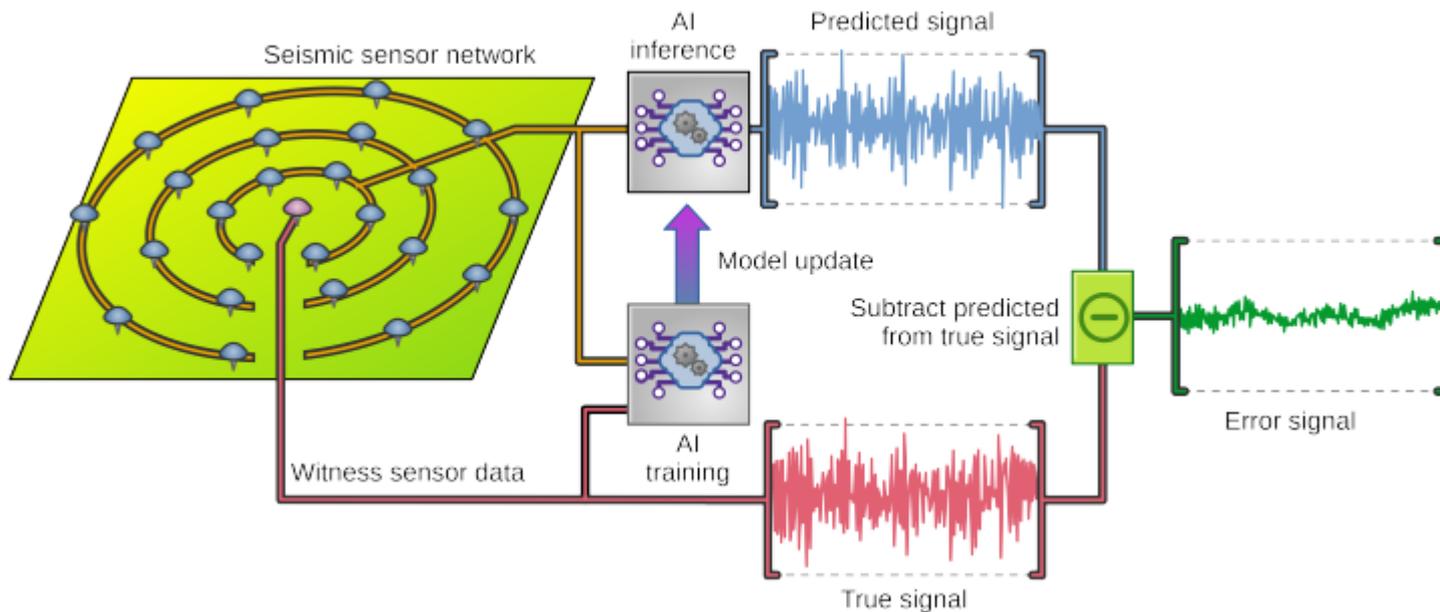
Figure 3. (a) PSD on Nov. 10, 2017 of the four seismometers marked by arrows in figure 2(b) showing microseismic noise at frequencies below 1 Hz and anthropogenic noise above 2 Hz. (b) and (c) Show spectrograms of the seismic noise featuring the diurnal variation for frequencies above 2 Hz.

- Survey data from the site study in Terziet, Nov 2017, was used to train and test the performance of the neural network
- We were interested in the band 0-10Hz. Data were downsampled to 25Hz and the instrument response was taken out.
- In the test, we use all sensors of the arrays, the witness channel is in the center of array A.

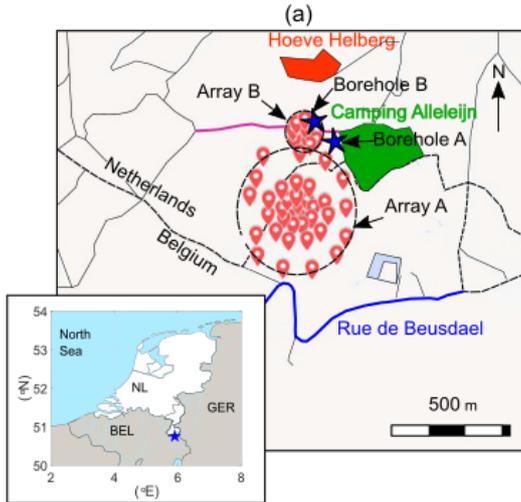
- Clearly visible – day-night differences. Note also that the PSDs between different nodes change: the ambient noise is not constant over the field.
- At any given time, the noise in any seismometer comes from different sources and follows different paths – one needs many sources simultaneously to describe the activity in the field, even over distances of 100 m.

Newtonian Noise – machine learning

- Newtonian noise: difficult to calculate from data: derivatives of the seismic displacement fields need to be known everywhere in a large volume (tens of cubic kilometers).
- Neural network: attempt to solve this problem by weighing/combining the full dataset (non-linear) – this approach may succeed if there is a way to train the network. ([thesis work Vincent van Beveren, submitted to GQC](#))



Reconstruction of displacement fields

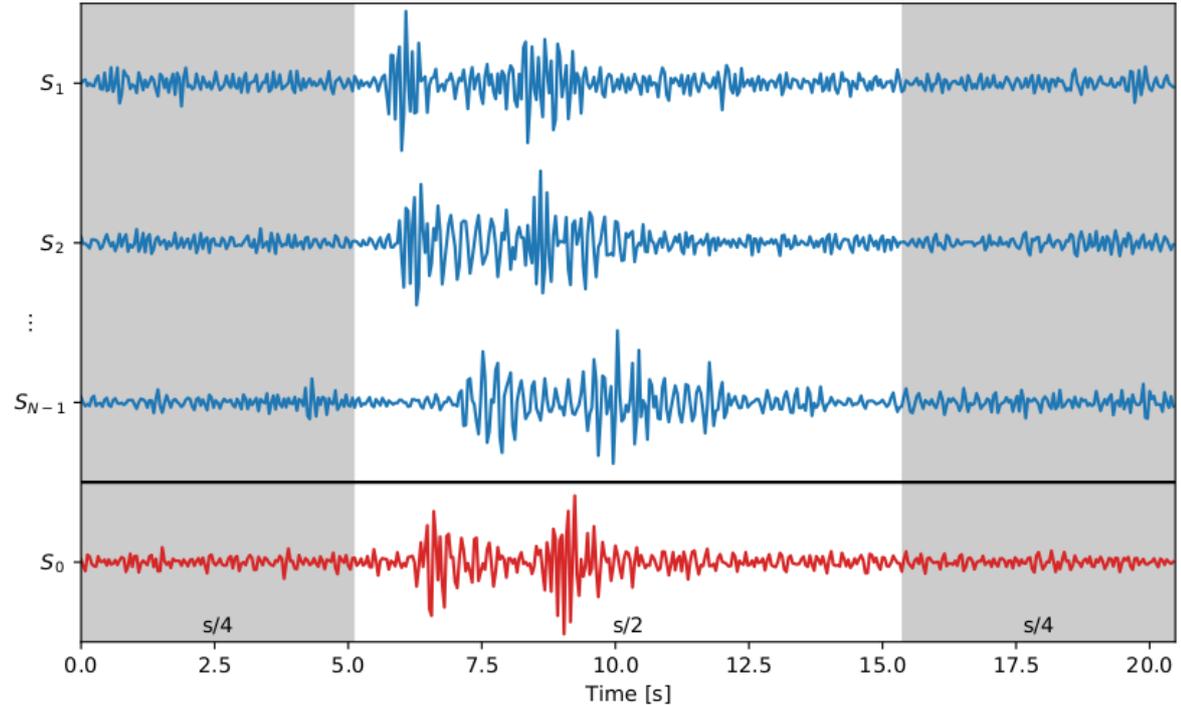


Time traces: relevant information in the band 0-10 Hz

- Downsample to 25 sps/ 12.5 Hz
- Remove instrument response
- Calculate maximal relevant lag: for lowest wave speeds of 100 m/s and max distances of 500 m: 5 seconds

Data (12 days of survey from Nov 2017): divide in 512 sample records (handy for FFT)

Records: staggered with an overlap of $s/2$ samples – in this manner the full time series can be analyzed, at the cost of having 2 times the number of records

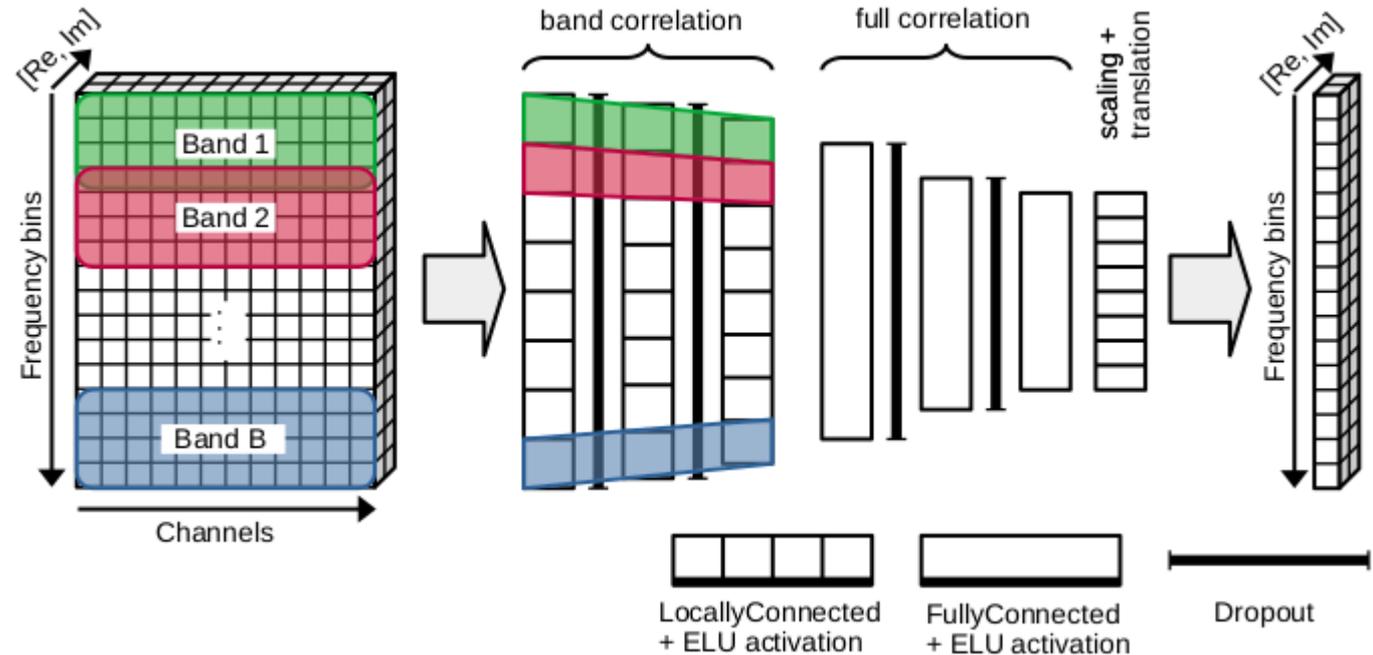


Typical time trace of 3 sensors and the witness sensor, containing a transient. The neural network makes a prediction for the witness sensor (red trace) using the other sensors (blue traces)

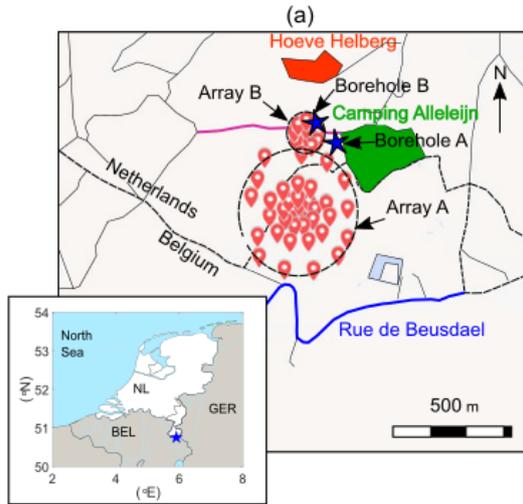
Data in the gray parts of the records may contain contributions from neighbouring segments and are spoiled; the central white part contains unspoiled results.

Neural Network – solve for displacement field

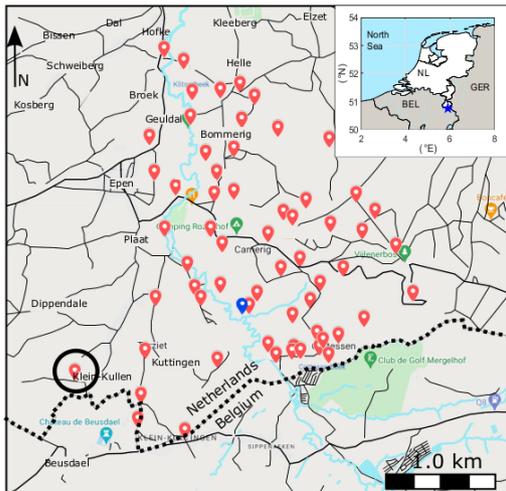
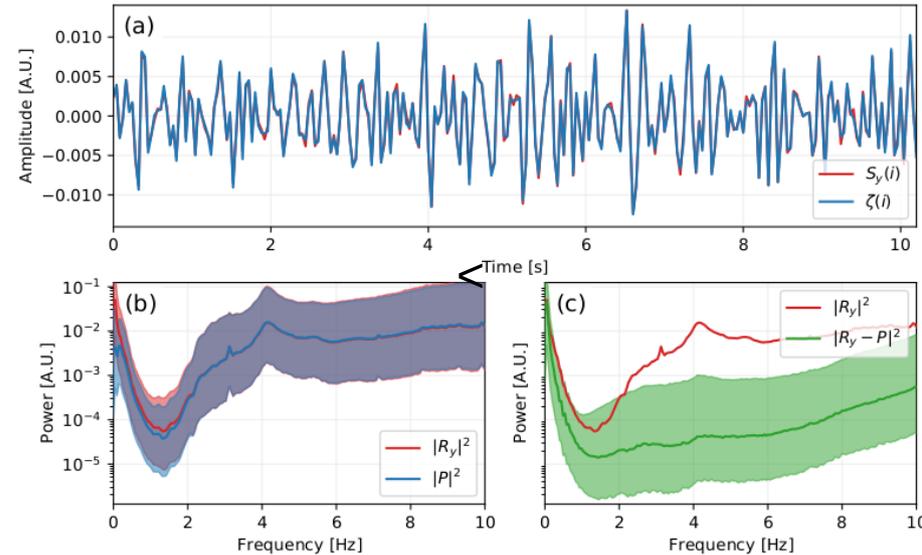
- Keras environment (<https://keras.io>) and Google Tensorflow are used. Python environment.
- Operate on FFTs of the data (we expect that intersensor correlations are stronger in frequency domain than in time domain)
- Data downsampled to 25 Hz, instrument response removed, and scaled to have similar amplitudes
- First three layers weigh the events based on correlations in the partly overlapping bands (18 bands, 11% overlap). Dropout occurs between the layers.
- Second 2 layers: full correlations between all bands. Final layer just takes care of scaling and translation.
- Hyper parameter optimization done with Keras hypertuning, using Nesterov (Nadam) algorithm.
- Training done on a Tesla NVIDIA V-100 GPU, 66M parameters used in the final configuration.
- 80 percent of the data is used for training, 20 percent for performance analysis.



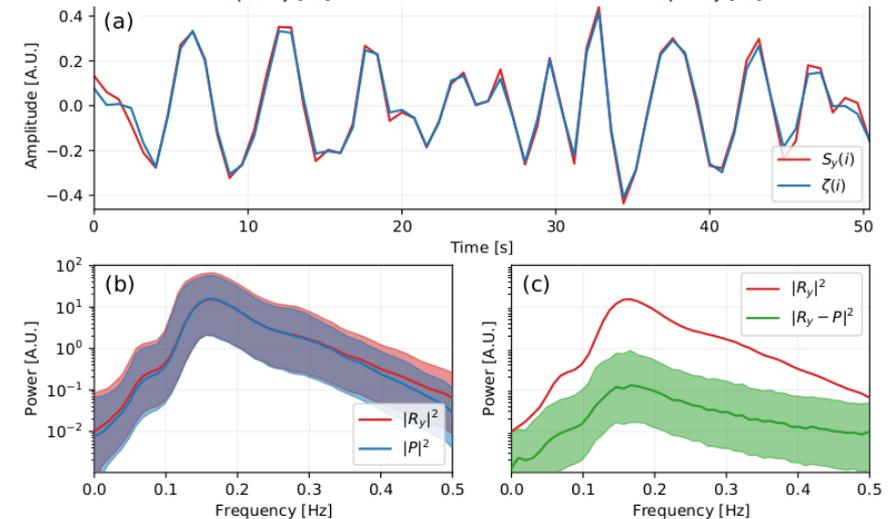
Displacement field, surface and borehole sensor



Reconstruction of the displacement of a witness sensor in the survey with surface sensors of Nov 2017. The sensor layout is shown on the left. The plot on the right show a typical time trace (red) and reconstructed trace (right) where the bottom shows the measured and reconstructed power and 10-90 percentile bands (red and blue), and the measured and residual power after subtraction (red and green). The Neural network reduces the power by about 2 orders of magnitude.



Reconstruction of the vertical displacement field of the borehole sensor at 250 m depth from the 5Hz geophone surface sensors (survey layout shown on the left). Due to the large distance between surface sensors, the band width was limited to 0-0.5 Hz. One tri-axial broadband seismometer at 3m depth in the Heimansgroeve was (blue marker) was also used, as well as a trillium-T240 at the top of the borehole (black circle). The neural network reduces the power by 2 orders of magnitude.



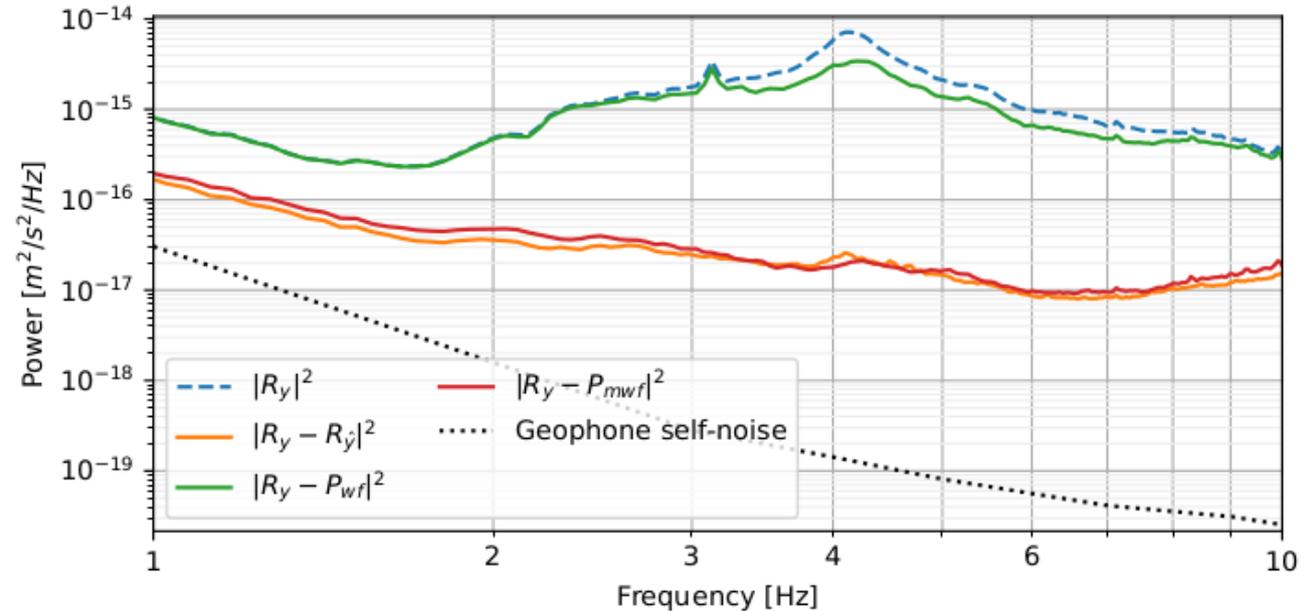
Wiener filter, surface displacement noise

- Wiener filter : optimal filter for linear problems

$$\text{corr}_{ij}(f) = \frac{S_i^*(f)S_j(f)}{|S_i(f)||S_j(f)|}$$

$$\text{WF}_{\text{Pred}}(S_{\text{witness}}) = \sum_i \sum_j \text{Corr}_{ij}^{-1} \text{corr}_{j,\text{witness}} S_j$$

- Built from cross-correlation matrix.
- Should give excellent predictions for displacement fields – Greens function between 2 points constant for a given frequency bin.
- However: noise spectrum not constant: transients, day/night variations, weather variations – it varies in time.
- e.g 1 FFT with 1 sensor 1000* average amplitude (from local transient) spoils the result for the matrix elements involving that sensor
- Without taking out transients, the Wiener filter performs quite poorly. We set a threshold and took out all segments where the 15-minute average power in a sensor was a factor of 3 or more above the running average of the previous 7 hours, for the Wiener filter (about 2 percent of the data).
- Neural network: outperforms Wiener filter: it can automatically correct for these changing conditions and does not need transient removal.



Blue: measured PSD.

Green: Wiener-filter residuals, without any data selection

Red: Wiener filter results with data reduction: noisy segments are discarded

Orange: Neural network results, all data

For the borehole sensor, such filtering is not needed: below 0.5 Hz we did not have any large transients in the surface elements that yield good correlation.

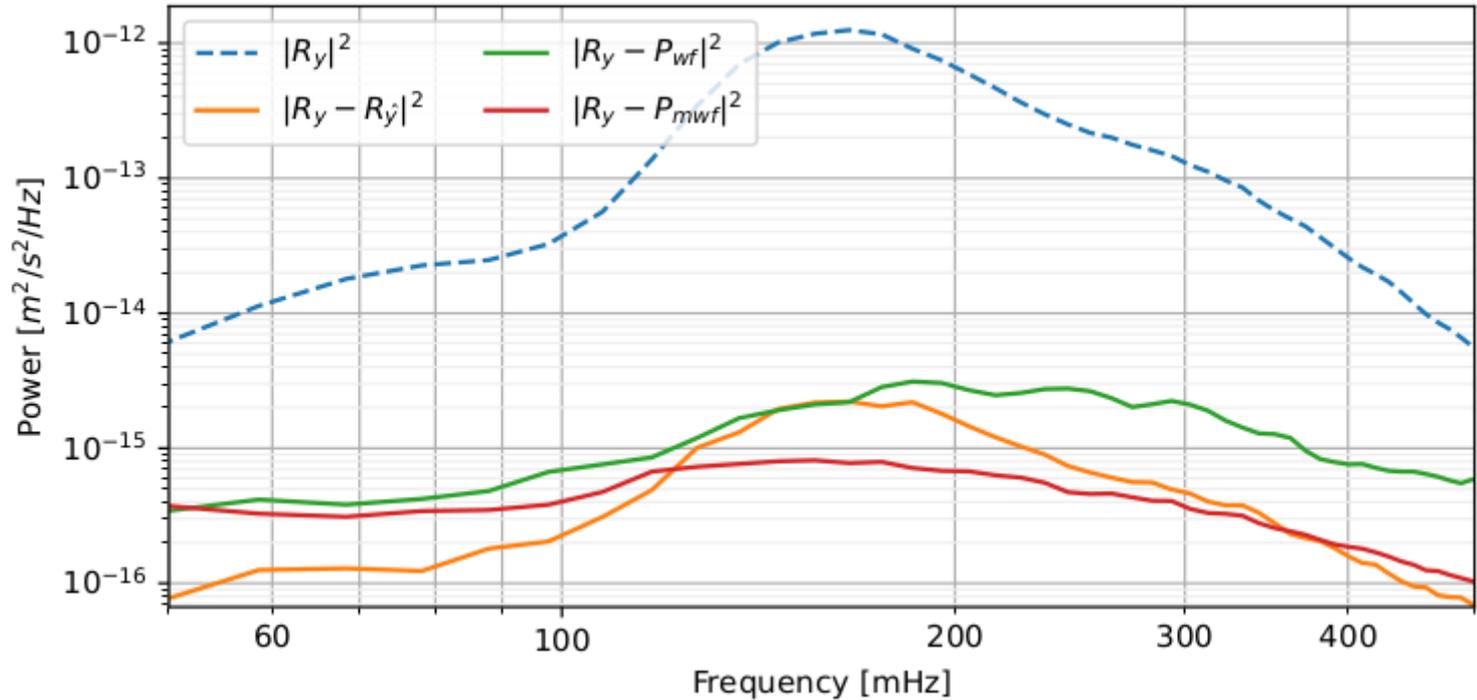
Wiener filter, borehole displacement noise

Neural-network and Wiener filter response;

Absolute performance very well: noise reduction integrated over the full band [50-500mHz] by 2 orders of magnitude.

Neural network is a bit better in estimating fluctuations below the microseismic peak. (Ground tilt? Sensor drift?)

Wiener filter needs some outlier removal for optimal performance above the microseismic peak, but the measured data has already low power in that range.



Blue: measured PSD.

Green: Wiener-filter residuals, without any data selection

Red: Wiener filter results with data reduction: noisy segments are discarded

Orange: Neural network results, all data

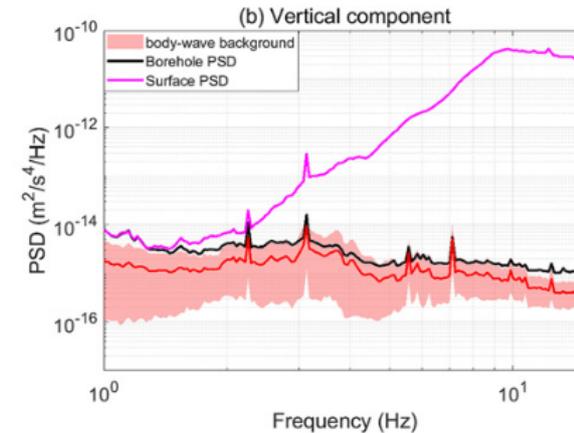
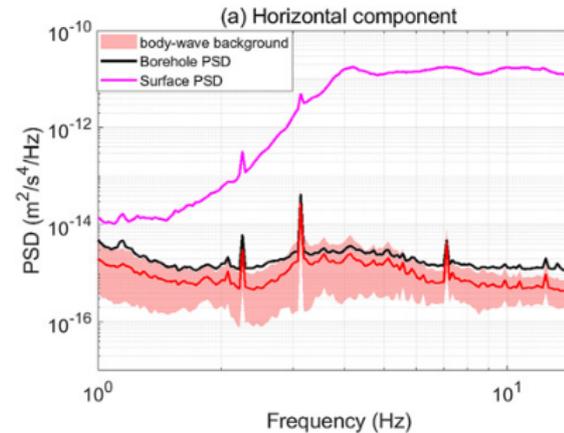
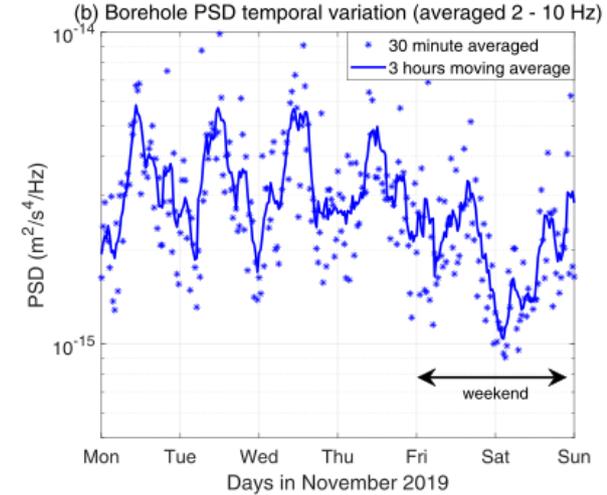
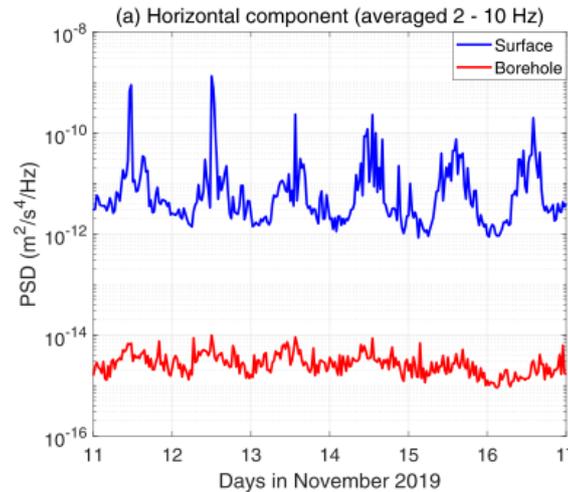
Newtonian-noise simulation: seismic field

No 3rd-generation GW detector available yet: train on synthetic data.

Active and passive seismic surveys used to characterize the soil and seismic noise in Terziet (Koley et al, GQC 35 (2022) 25008, Bader et al. GQC 39 (2022) 25009)

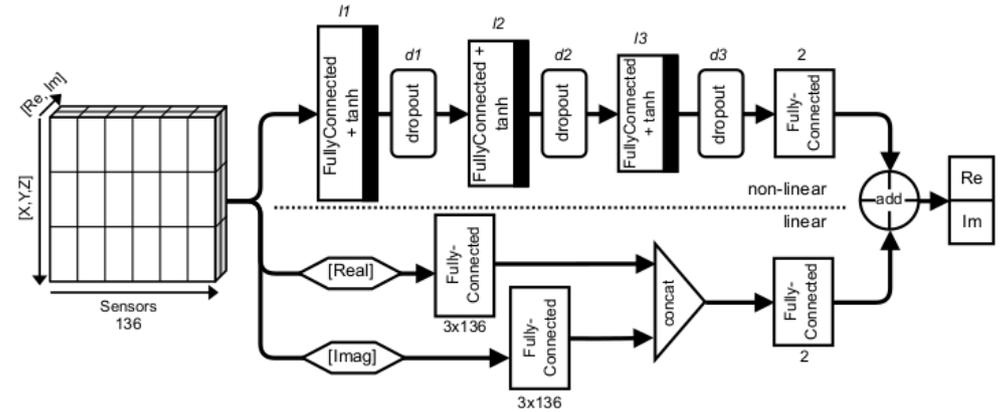
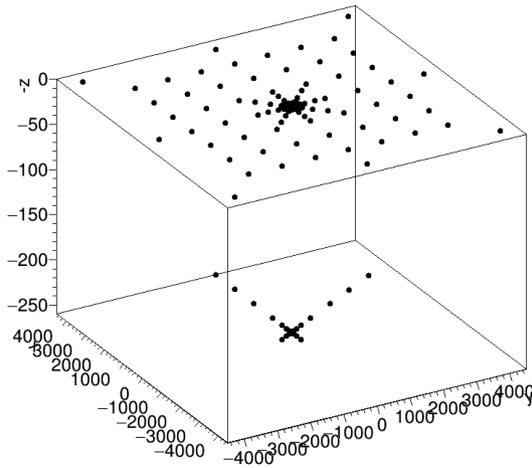
We used the Elastodynamic Toolkit to calculate displacement fields in a radius 4km around the mirror. Sources isotropically distributed 6m around mirror. The amplitudes are drawn from a Gaussian distribution with the mean such that the average observed PSDs are matched; observed features such as the first overtones and the ratio between underground and surface PSDs are reproduced.

The from the day-night observations we derived the presence of a common body-wave background. Since we did not have an underground network, we could not derive the constitution of that background. It has been modeled by isotropic plane waves with a p-wave speed of 4500 m/s and an s-wave speed of 2800 m/s; the speeds measured from the active and passive surveys in the area.



Newtonian Noise: neural network and Wiener filter predictions

- We generated 50000 trials per frequency; 40,000 were used for training the network and 10,000 for determining the performance.
- Newtonian Noise: 10-m cavern radius around a mirror at 250 m depth; 4 km integration radius.
 - The excavations of halls around the mirror and the presence of the arms are not included in the NN calculation, for ease of comparison.
- Sensor data: 132 sensors, tri-axial geophones at the surface, and trillium seismometers in the 2 arms. We assume that the arm extends 500 m downstream (we added sensor noise in the trials)
- Witness channel: the acceleration of the mirror in x and in y direction (we ignored acceleration in vertical direction).



Hyper parameters

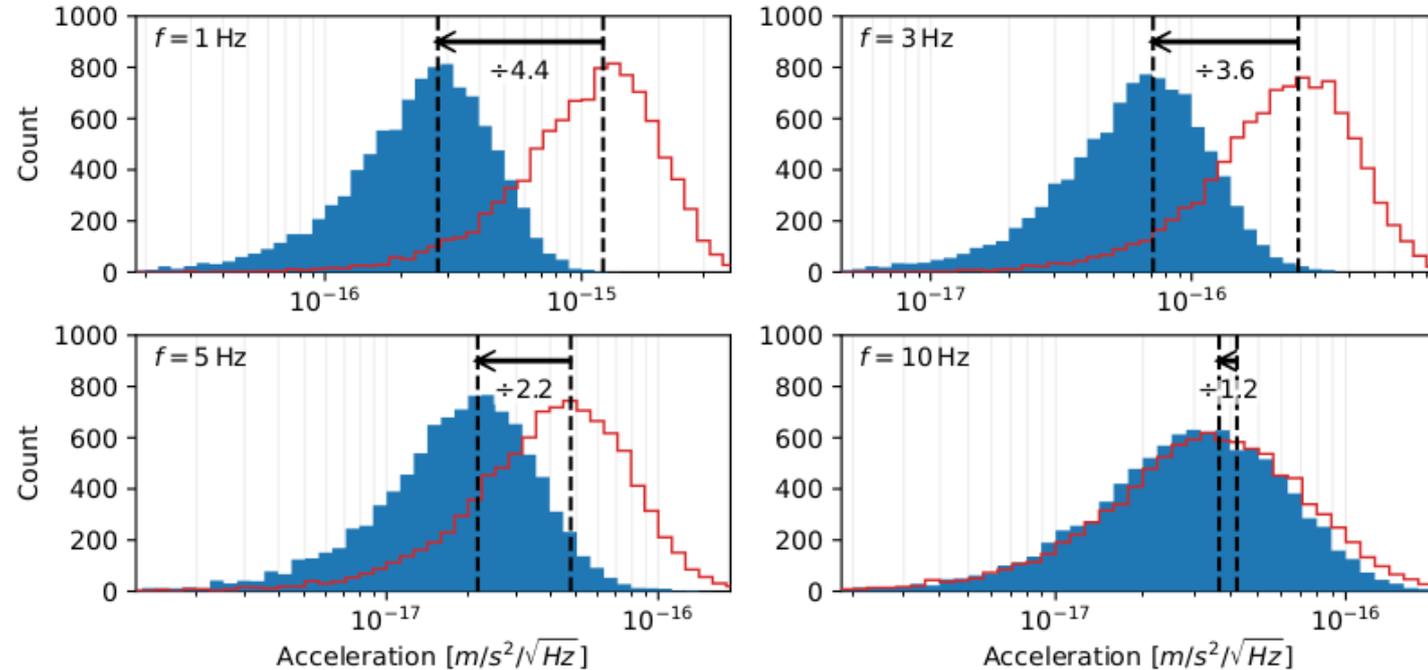
f [Hz]	I1	I2	I3	d1	d2	d3
1	1536	1228	982	30%	27%	24%
3	3072	2457	1965	10%	4%	2%
5	3072	1536	768	40%	20%	10%
10	3072	1843	1105	30%	27%	24%

Neural network: 2 pipelines; top with 3 fully connected layers each with 3 non-linear activation functions, bottom a linear pipeline without hyperparameters (same as Wiener filter: linear combinations of the input). Input: 1 frequency bin for 135 tri-axial sensors (displacement field) and 1 tri-axial witness (the mirror acceleration).

Non-linear pipeline: can attempt to extract volume density fluctuations (from gradients of the seismic displacement field) by using combinations of sensors. Linear pipeline: Wiener filter; may work for surface displacements since then the density fluctuation is depends linearly on the component of the seismic field normal to the surface.

Newtonian noise reduction; Wiener filter

- Wiener filter: expected to do well when surface terms dominate.
 - Density fluctuations depend on the derivatives of the displacement field- wash out of isotropic noise
 - Surface displacements: linear contribution.
- Largest cross-correlation value between mirror acceleration and sensor displacement: at 1Hz this occurs between the sensor in the tunnel 250m downstream and the mirror acceleration: 0.609 for the horizontal displacement and 0.346 for the vertical component.
 - P-waves under an angle have the attraction of the density around that position in phase with the density fluctuations at the mirror. The sensors in the cavern itself 6 m away from the mirror exhibit smaller correlations.
- At 10 Hz: maximal correlation with the vertical surface displacement at $x=-400\text{m}$, value of 0.14 only. Next-highest for the sensor at $x=-200\text{ m}$. The largest correlation for an underground sensor at 10 Hz occurs for a sensor in the 60-deg. Arms at $x= 7\text{m}$, $y=-11\text{m}$; with values of 0.098 and 0.089 for accelerations in the x and y directions, respectively.
 - Surface sensors dominate; at 10 Hz the PSD at the surface is 4-5 orders of magnitude above the PSD in the borehole.



The amplitude of the mirror acceleration in x-direction (red histogram) compared to the residuals using the Wiener filter prediction (blue histogram). The distance between the RMS of these amplitude distributions are shown; noise reductions between 26dB (1Hz) and 1dB (10 Hz) are obtained .

Newtonian noise reduction; Neural network

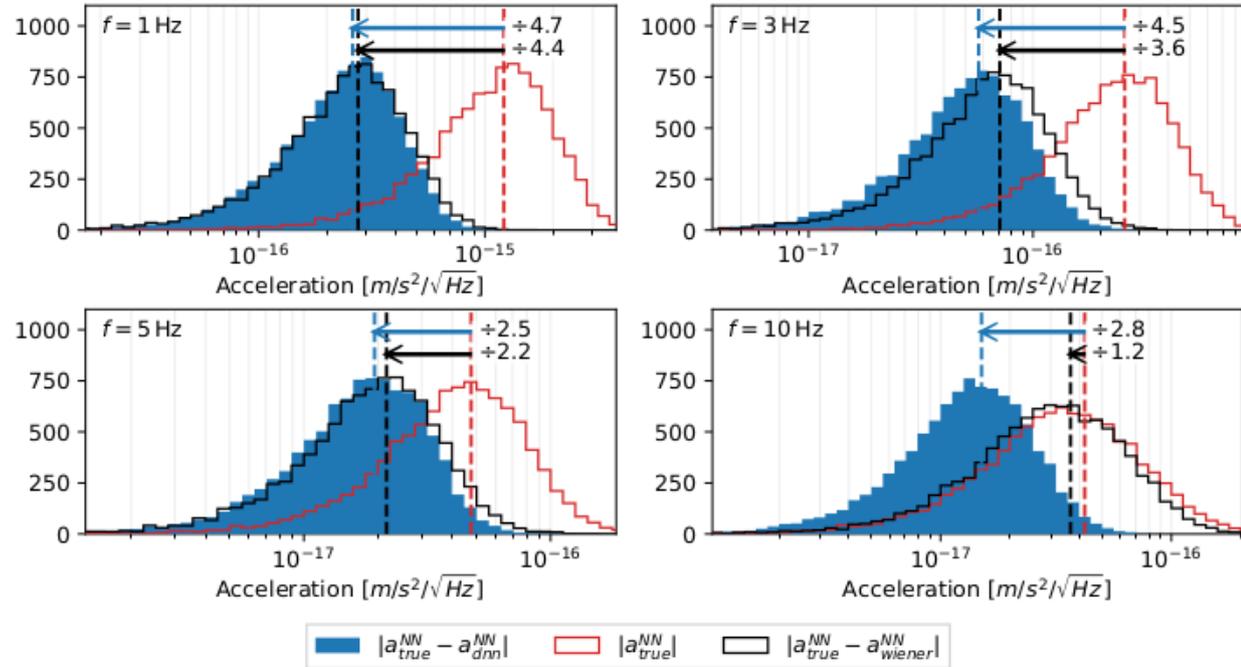
- The neural network has non-linear layers and activations, and outperforms a Wiener filter, especially at 10 Hz.

- We believe that this is due to the non-linear relation between displacement fields and mirror acceleration

$$\xi_j = Ae^{i(kx_j - \omega t)} \Rightarrow \delta\rho = ik \xi_j \rho dV$$

(the density fluctuation in a volume element is 90 (+k) or 270 (-k) degrees out of phase with the displacement field for p-waves; For a Wiener filter, one expects that the cross correlation approaches zero for integrating over an isotropic field)

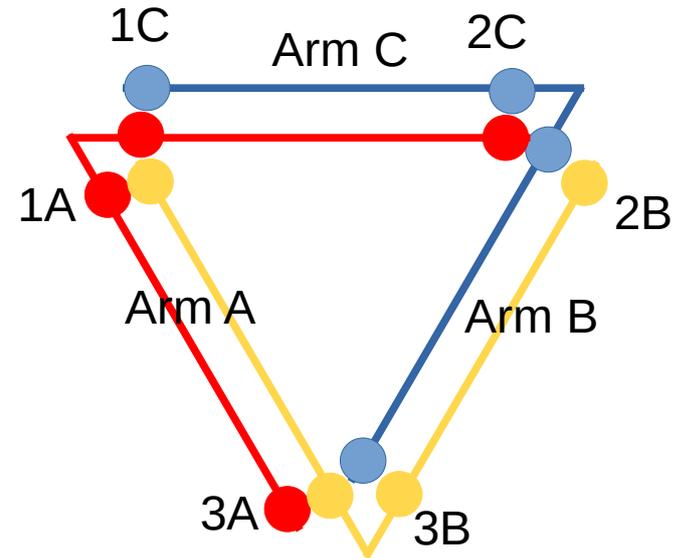
- We verified that the neural network can reduce the (uncorrelated) acceleration in both x and y direction with the same efficiency as for a single acceleration.



The amplitude of the mirror acceleration in x-direction (red histogram) compared to the residuals using the Wiener filter prediction (blue histogram). The distance between the RMS of these amplitude distributions are shown; noise reductions between 26dB (1Hz) and 1dB (10 Hz) are obtained.

Training

- The neural network (and also the Wiener filter) requires training; there must be a witness channel available.
- With 3 co-located arms, one can make an error signal for the change in a single arm length.
- The change in arm length depends only on the accelerations of the input and end mirrors in the direction of the arms; for co-located arms that is only on 2 variables (whereas the signal in a single interferometer depends the longitudinal accelerations at all FP mirrors; 4 variables)
- For the neural network, the expected performance to reduce the noise for 2 independent parameters (the longitudinal accelerations at the ends of 1 arm) is mathematically of the same complexity as reducing the mirror acceleration in 2 perpendicular directions.
- We therefore expect that a good NN reduction is feasible with a neural network and a modest number of seismometers in the tunnels and on the surface; In the order of 500 in our study



$$S_{CA} = \Delta C - \Delta A = (\Delta x_{2C} - \Delta x_{1C}) - 1/2(\Delta x_{3A} - \Delta x_{1A}) + 1/2\sqrt{3}(\Delta y_{3A} - \Delta y_{1A})$$

$$S_{AB} = \Delta A - \Delta B; S_{BC} = \Delta B - \Delta C$$

$$\Delta x_{2C} - x_{1C} = 1/2(S_{CA} - S_{AB} - S_{BC})$$

Summary and outlook

- We have shown that neural networks are a promising tool to subtract Newtonian noise:
 - The neural network can correctly predict seismic displacement data at a random location, if trained; we have shown that for both a borehole sensor and surface sensors
 - The neural network even outperforms a Wiener filter, often considered optimal, since it does not need outlier removal or pre-cleaning of the data set.
 - The neural network was tested against a synthetic data set. This set contained 135 hypothetical sensors on the surface and in the interferometer arms; and data were provided by solving the elasto-dynamic toolkit equations for the soil and noise as measured in the EU Meuse-Rhine region
 - Caveats: for body waves we do not have a good characterization at present, since we did not perform a survey with more than 1 underground sensor. The PSD of the underground sensor was matched with a plane-wave bodywave background which may be too simplistic
 - Furthermore, the underground cavern structures for a realistic interferometer were not modeled. We expect that the presence of excavations along the arms reduce the Newtonian noise significantly, but a realistic optimization needs a final design and also rescattering from the tunnel walls need to be numerically studied
- Newtonian noise depends on the gradients of the seismic displacement field. Therefore, we believe that a neural network is better suited than a Wiener filter for NN-subtraction.

Backup slide

- EDT toolkit calculations with the soil parameters as determined from the seismic surveys in the German-Belgium-Dutch EU-region [Koley et al, GQC35, 25008; Bader et al, GQC39, 25009].

shown are the horizontal (blue) and vertical displacements (red) from a horizontal source with a force of 1Nm as a function of radial distance of the source. The black curves show the same for vertical displacement resulting from a vertical force at the surface with a harmonic frequency of 5 Hz.

The displacement fields are plotted for depths of 0m (solid lines), 25m (dotted lines) and 250 m (hashed lines)

Note that the amplitudes do not linearly decrease as a function of distance, there are complicated interference effects from the layered structure of the soil.

