Technical acoustics - TTT4180 - A0

Erlend K. Berg[1)], NTNU Trondheim, Norway

1 Introduction

The natural soundscape contain many sources for noise. In this report the focus will be on the contribution a passing train has on the natural soundscape. A recording with length of 30 minutes were done at Marienborg train station where the contribution of a passing train took place once, but also contained noise from other sources. The agenda for this report is to familiarize how data can be processed and analyzed to visualize sound sources of interest in an informative way.

2 Materials and Methods

2.1 site and soundscape

The time of recording was done in the time zone 10:57-11:30(GMT+2) the 29th of September 2021 at Marienborg train station, Trondheim(N63°25′7″ E10°22′58″). During the time of recording the soundscape contained noise from light traffic, intercom messages from the train station speakers, dunking noises from a digging machine at a worksite across the river near Lerkendal, trains, wind and noise from trees, leaves and other natural obstacles that may produce sounds. The weather was sunny with a light wind at 3-5m/s from south-West, the humidity was at 56% and the temperature at 16^{o}C. The sound of interest was from a train that stopped at the station for then to accelerate and leave. Its soundscape consisted of noises from the engine and motors due to it being a hybrid between electric and diesel, friction between the wheels and the brakes, gasses released from pressure valves and the sound of mechanical parts when the train began to move. The passing train was at it's closest 3.2 m away from the microphone.

2.2 Setup

A ZOOM H5 with an attached microphone was used to record the data[1]. The position of the microphone was as illustrated in Figure 1 with its microphone pointing towards the railroad at an angle of 34^o towards the east, with north as reference. The microphone where set to 1.6 m above the ground using a microphone stand, and the distance from the railway and the microphone was 3.2 m. A small wind cover where used to cover the microphone due to wind having a very varying presence, no brand name was given for both the microphone stand and wind cover. Before and after the recording session, a calibration recording where done at NTNU, Gløshaugen using a calibration tool that produces a 94 dB sound at 1KHz. All equipment is borrowed from the faculty and acoustics lab.

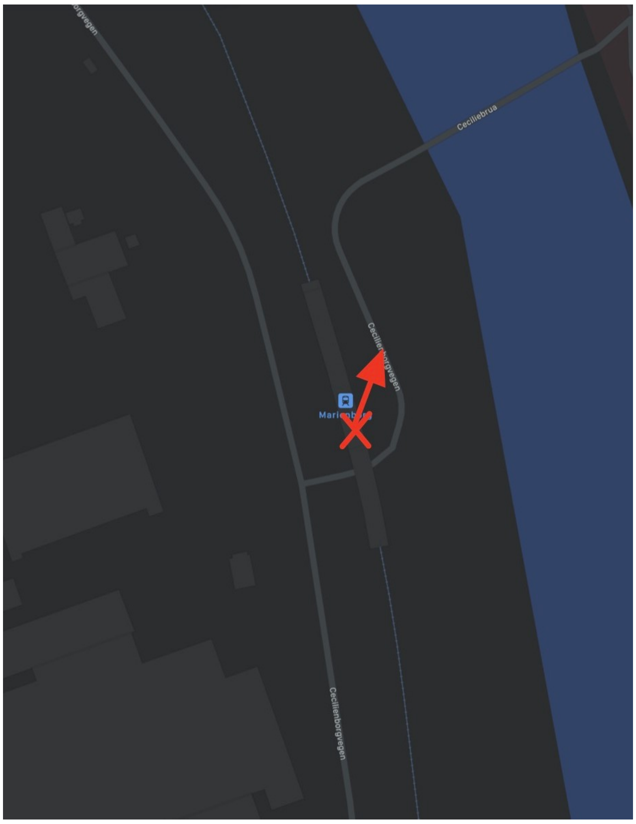

Figure 1: Microphone placement and direction at Marienborg station, Trondheim.

2.3 Methods

Python 3.10 was used to process the recorded data with the use of libraries such as matplotlib, numpy and scipy. The code was written in Microsoft visual studio code. The calibration signals from before and after the recording session was analyzed by calculating the total sound pressure, L_P, for both signals and comparing the values to see if any changes had occurred on the recorder or microphone. To calculate L_P we first must obtain the p_{rms} value for the calibration signal. To do this the average rms (root mean square) value will be calculated for a window of N samples. To obtain the p_{rms}- and L_P value the following equations are used

$$p_{rms,N}(n) = \sqrt{\frac{1}{N}\sum_{n=t}^{t+N} |n^2| \text{ in Pa}} \tag{1}$$

$$L_p = 10 \cdot \log_{10}\left(\frac{p_{rms}^2}{p_{ref}^2}\right) \text{in dB ref. } p_{ref} \tag{2}$$

where n is discrete values in the sound pressure array, t marks the beginning of the window and p_{ref} is the reference pressure value for the propagating medium, which will be air.

The signal coming from the recorder is a WAV file (Waveform Audio File Format) having discrete values that is not given in pascal(Pa), the unit of pressure. Therefore, the values need to be scaled with a scaling factor, A, so the amplitude of the signal can be presented in Pa. The scaling factor A is given as

$$A = \frac{n_{max}}{\sqrt{2} \cdot p_{max}} \tag{3}$$

where p_{max} and n_{max} are both the peak-amplitude of the real-and discrete signal. The parameters of the real signal are given by the calibration tool. Since the sound signal from the calibration tool represent a pure sinusoidal, the expected maximum value for the calibration signal is calculated, and the maximum value from the sampled signal is found by averaging a number of maximum values to ensure that the maximum is not a result of movement or disturbances produced when calibrating.

To analyze the recording signal of interest, the signal first needs to be converted to pressure values by applying the scaling factor value A as calculated using the calibration signals. Both A-weighting and third octave filters will be used when analyzing the sound pressure signal.

2.4 A-weighting

A-weighting is defined by the international standard IEC 61672:2003 and are used to account for the relative loudness human ears perceive sound. Human ear canals have a dip in perceived loudness at around 3-4KHz where the A-weighting will amplify the weighted signal and dampened for the lower frequencies. A signal weighted with the A-curve will be more correct towards perceived sound and noise for humans, and there for a more exact way of representing L_p when determining how loud a soundscape is for humans. The A-curve is not very steep and therefore not particularly good for frequencies over 12 KHz, due to the mismatch in the weighted curve and the perceived loudness for humans. A choice is to then use the G-curve which is steeper at the higher frequencies. Even so, the A-curve is the most used. For all intensity levels humans have normalization to 0 dB at 1KHz. The offset for the curves will therefore cross this point with a weighted factor of 0 dB. The A-weighted curve is given as

$$A(f) = 20 \cdot \log10(A_f) - A_{1000Hz} \tag{4}$$

where

$$A_f = \left(\frac{f_4^2 \cdot f^4}{(f^2 + f_1^2)\sqrt{(f^2+f_2^2)(f^2+f_3^2)}(f^2+f_4^2)}\right) \tag{5}$$

and f_1, f_2, f_3 and f_4 are determined by the IEC standard and A_{1000Hz} are the normalization constant given in decibels representing the gain needed to provide a frequency weighting of 0 dB at 1KHz.

The A-weighted curve can also be directly applied to the signal if it first is filtered into third octave bands. The spectral weighting adjustment factors to apply to each third-octave band is stated in the IEC61672:2014 . This is the preferred method due to computing time and accuracy of the analog filter.

2.5 Third-octave filter banks

Third-octave filters is a set of bandpass filters with center-, upper-and lower frequencies determined by

$$\text{Third-octave} = \begin{cases} f_{mn} = 10^{n/10} \\ f_{lower} = f_{mn}10^{-1/20} \\ f_{upper} = f_{mn}10^{1/20} \\ \dfrac{\Delta f}{f} = f_{upper} - f_{lower} = 0.23 \end{cases} \tag{6}$$

To apply the third-octave filter, a set of band-pass-filter banks are used with passband between f_{lower} to f_{upper}. This is done for all f_{mn} to create a filter bank. To do this in the digital domain we use Butterworth filters which have a relative steep slope decided by the order of the filter. The higher the order, the steeper the slope. The signal will be filtered in time domain by each individual bank and then transformed with the use of Fast-Fourier transform (FFT) to find its power spectrum. By iterating through the filtered power spectrum, a sound pressure level value can be given for each band in the third-octave register. This is done by calculating the total L_p value for each band. Compared to the FFT, which is a constant bandwidth analysis, the third-octave filter is a constant relative bandwidth filter that purely depend on its center frequency, f_c. The size of the pass-band and its center frequency is given in eq 6.

The even faster method is to first apply the FFT to the signal and then calculate the sound pressure level for each third-octave band.

3 Results

3.1 Calibration

By the use of equation 3 the scaling factor, A, was found to be $A = 354075323$, and then applied to both calibration signals to get a signal with magnitude measured in pascal. Both signals where plotted as a function of time to see for any distinct disturbances, and L_p for both calibration signals where calculated as shown in Figure 2.

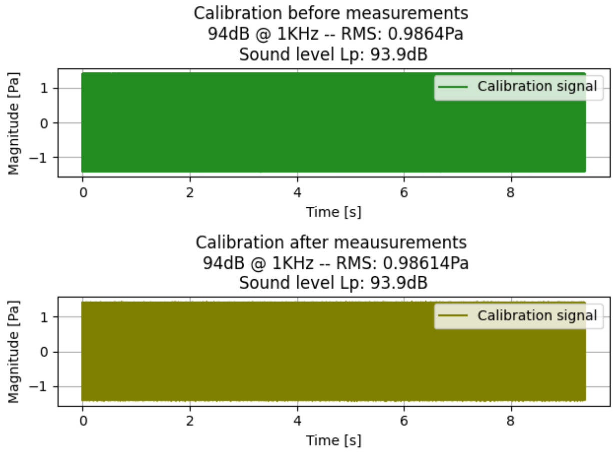

Figure 2: Calibration signals with the calculated L_p and rms value.

The calibration signal at 94 dB will have a rms value at $p_{rms} = 1.0023745$ Pa, while the calculated rms value from the calibration signals are

$$p_{rms} = \begin{cases} \text{Calibration signal} = 1.0024 \text{ Pa} = 94 \text{ dB} \\ \text{1st calibration} = 0.9864 \text{ Pa} = 93.9 \text{ dB} \\ \text{2nd calibration} = 0.9861 \text{ Pa} = 93.9 \text{ dB} \end{cases}$$

The difference between the calibration signals are so small that no changes has occurred on the recorder during the recording session.

3.2 30min Recording

The recorded sound signal was scaled with the scaling factor found in the previous chapter and filtered with the A-weighted filter as well as sorted into third-octave banks. The total sound pressure level for the entire 30 minutes was calculated to be

$$L_p = \begin{cases} \text{Non-weighted} = 71.1 \text{ dB} \\ \text{A-weighted} = 54.9 \text{ dB} \end{cases} \tag{7}$$

and the signal as a function of time can be seen in Figure 3.

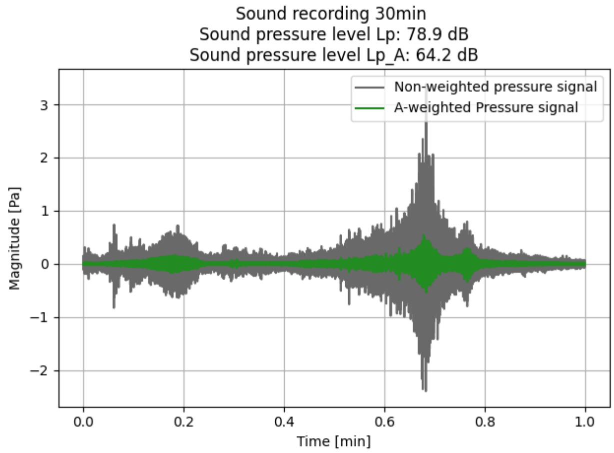

Figure 3: 30min recording signal with-and without A-weighting.

3.3 1 minute snippet

The entire 30-minute signal will show what frequencies are most present on a regular basis and mostly just capture the background noise. This is of no interest, so a short snippet with length of 1 minute was extracted where the passing of a train was present. The power spectrum for the 1-minute signal is seen in Figure 4 having both the Z- and A-weighted FFT and third-octave banks. From the power spectrum it's evident that the contribution of the lower frequencies dominates. It's hard to distinguish what contributions the train has on the soundscape from the shown power spectrum.

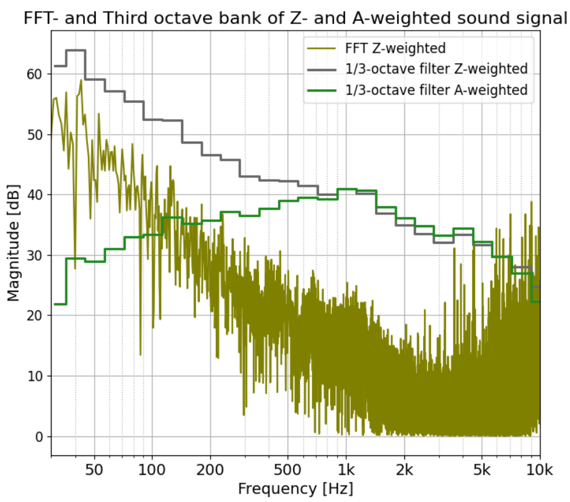

Figure 4: Power spectrum of 60 seconds from the recording signal with length of the transformed axis, N, equal to the sampling rate at 48000 Hz.

3.4 6 second snippet

To further see the power spectrum given by the train, the snippet was shortened down to 6 seconds where the train was close to the microphone. The power spectrum for a signal with length of 6 seconds can be seen in Figure 5 containing both the Z- and A-weighted FFT and third-octave banks from 31.5 Hz to 10k Hz.

It is clear from the power spectrum that the signal has distinct peaks at around 45 Hz, 90 Hz and 7 KHz. The high-pitched tone is produced by the electric motors as a high frequency screech, as well as the friction produced by the wheels and brakes.

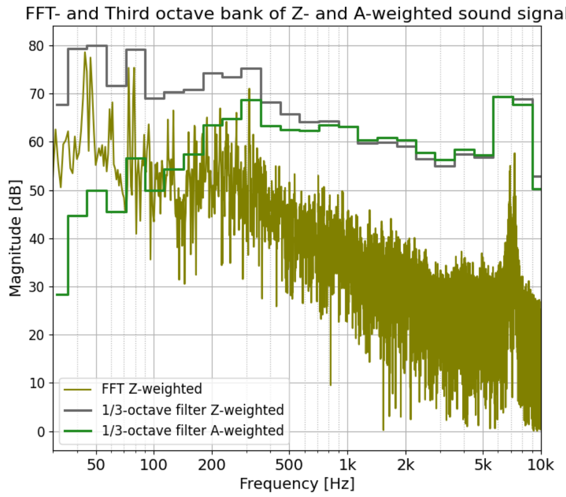

Figure 5: Power spectrum of 6 seconds from the recording signal with length of the transformed axis, N, equal to the sampling rate at 48000 Hz.

The equivalent continuous sound level (L_{eq}) for both a short and fast time constant, at respectively, 125 ms and 1 s is seen in Figure 6.

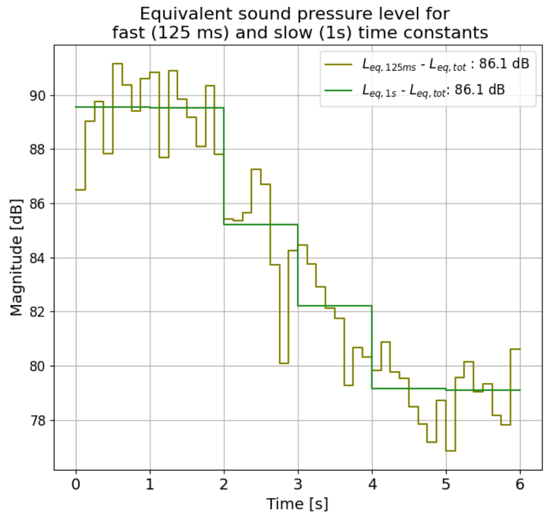

Figure 6: $L_{eq,125ms}$ and $L_{eq,1s}$ for the time snippet of 6 seconds.

4 Discussion

As can be seen in Figure 2 it is clear that the calibration signals show no significant differences, meaning no damage was done to the equipment and recording, and the scaling factor, A, received is valid. From the power spectrum in Figure **??**, it is clear that the lower frequencies have a significant contribution. Considering the contribution of wind and other natural sources as well as low frequencies always having a significant presence might be the reason for the huge presence of lower frequencies. By looking at the A-weighted third octave banks, it is clear that the contribution of the lower frequencies are less noticeable for humans, but rather the higher frequencies at around 1 KHz is the highest perceived frequency. By looking at Figure 5 from the short snippet with length of 6 seconds, it is clear that the overall contribution increase. This is because when standing close to the train, it becomes the dominant source. While at a stand still, and while moving, the train produce some distinct tones but has also an overall spread contribution of noise to the power spectrum due to pressure valves and other mechanical parts.

It is also evident that when the train begin to move, the screeching of from friction between the brakes, wheels and tracks produce an higher pitched tone at around 7 KHz and some distinct tones in the lower frequencies around 45 Hz and 90 Hz.

4.1 Problems and improvements

Problems faced during recording was that no trials were done before the 30-minute recording. By trying various locations, it would be easier to find the best place to record considering radiation pattern of the sound source. To further improve results from post processing, a filter to remove background noise could be implemented by comparing the 1-minute signal having the source with a 1-minute signal free from the source, assuming that the background noise has a constant power spectrum. The noise was mostly low frequent noise so a low-pass filter could also be applied.

Initially the data was post processed in the time domain, but further improved by filtering the data in the frequency domain. This was first visible for the third octave banks, where the A-weighted third-octave banks had a small difference in sound pressure level at 1 KHz, which should be equal for both Z- and A-weighted band levels. To reduce the amount of varying magnitude from the FFT, a moving average was applied with a window size of 2 samples and maintaining the total energy by converting first from dB to Pa and then back.

5 Conclusion

By analyzing the 1min signal it is easy to distinguish the screeching tones produces by the train wheels and engine with its fundamental at 2KHz and the related harmonics. Its contribution to the soundscape is evident, but the timewindow is rather small due to the few amounts of trains stopping by Marienborg station. The natural soundscape is filled with low frequency noise that is very evident from the power spectrum, but these frequencys have a rather low loudness for humans as can be seen by the A-weighted graph.

References

[1] ZOOM H5, ZOOMCORP. https://zoomcorp.com/en/us/handheld-recorders/handheld-recorders/h5/.

[2] Lecture notes 9, week34, Noise and Hearing. Guillaume Dutilleux. 2020.

[3] Jmrplens, GITHUB, PyOctaveBands.py. https://github.com/jmrplens/PyOctaveBand/blob/098eb7560627889fb0bad031280c98bcfe139d26/PyOctaveBand.py. 2021.

[4] Python-acoustics. https://github.com/python-acoustics/python-acoustics/blob/master/acoustics/weighting.py. 2019.

[5] Octave Band Frequencys, https://www.engineeringtoolbox.com/octave-bands-frequency-limits-d$_1$602.$html$.2010.

[6] Bureau of Indian Standards: Electroacoustics - Sound level meters BIS IS 15575-1. 2005.

[7] Window function, Wikipedia, https://en.wikipedia.org/wiki/Window$_f unction$.2021

[8] Head acoustics, application note, https://cdn.head-acoustics.com/fileadmin/data/global/Application-Notes/SVP/FFT-Octave-Analysis-Wavelet-Application-Note.pdf. 2018.