Applicability of Fully Mobile Measurements for Noise Maps

Floris Copraij

June 2024

Applied Data Science master thesis
Department of Computational Sciences Utrecht University

1st supervisor: Jules Kerckhoffs
2nd supervisor: Marc Kreveld
In cooperation with: Max v.d. Wielen

Abstract
Noise maps provide insight into the distribution of environmental noise throughout an area. This makes them an essential tool in communication regarding- and planning measures against- noise pollution. Currently, most noise maps that are produced are based on mathematical models which create detailed analysis but are unable to capture all of the complexity inherent to noise. This study evaluates the feasibility of a hybrid approach which allows modelled noise maps to be extended with simple to collect fully mobile measurements.

A noise measurement setup was mounted to the roof of a car and measurements were collected in three different areas. Recording the noise measurements from a moving vehicle included interfering noise from both the collection vehicle and from wind hitting the microphone. To address this noise reduction, regression adjustments and data selection approaches were tested. The resulting models suffered from local inaccuracies and smaller value ranges when compared to reference models. Despite this, the models created were able to capture both general trends in the data and particularities of the areas measured, showing the potential of this hybrid approach given a larger collection of noise samples and more detailed model input.

1 Introduction

Noise maps are tools which provide insight into how noise is distributed throughout an area. They are a key tool in increasing the understanding and reducing the propagation of noise within an area. The reduction of noise is key as, although some may consider it a simple annoyance, regular exposure to the amounts of noise common in many cities can negatively affect one's mental and physical health (Basner et al., 2014).

Traditionally for the creation of maps, it was necessary to conduct a large amount of noise measurements. Due to the great spatial and temporal variety inherent to noise data, measurements must be collected at multiple locations at various different points in time (Barham et al., 2010, Banerjee et al., 2009), resulting in such procedures being expensive both in terms of time and resources. Because of this, it has become common to use modelling-based approaches for the creation of noise maps. Within Europe for example, most noise maps are based on CNOSSOS-EU (Kephalopoulos et al., 2012), a framework which provides guidance on how to calculate noise models based on road- and railway traffic, aircraft-, industrial-, and population-based noise. Although many modelling approaches are very comprehensive it is unfeasible to consider all possible noise sources and the unique characteristics of all the different areas where such a model is applied, limiting the effectiveness of this approach.

This study evaluates the viability of an alternative hybrid approach in which measurements taken from a moving vehicle are used to train a machine-learning model to create noise maps. Combining the mobile measurements with the general trends from a modelling approach could potentially create more detailed noise maps for any area measured. Performing measurements while driving has been successfully utilised in prior epidemiological work and allows for great spatial diversity amongst the measurements (Kerckhoffs et al., 2021). In this particular case, however, the noise generated while driving (by the car or passing wind) will likely interfere with the measurements and affect the resulting noise map. Hence the ultimate goal of this study is to propose and attempt various approaches for limiting this inference. In short, it attempts to answer the question *How can noise data obtained through car-based collection methods can be used to enhance noise maps?*

A more comprehensive overview of the effects of noise and existing techniques for collecting and modelling noise data is given in section 2. An overview of collected and external data used in this study is given in section 3. Section 4 and 5 then detail the methodology and results of this particular study. These results are then discussed in section 6 together with an overview of the challenges encountered at each step of the methodology. Conclusions are provided in section 7.

2 Literature

In this section prior works relevant to noise, noise measurements, and noise modelling are briefly discussed to provide additional context for the methods and results discussed in the remainder of this paper. First, the topic of noise is defined more clearly after which the models of noise monitoring and noise mapping are discussed in separate sub-sections.

2.1 Noise

The term noise is often utilised in various contexts. However, within this paper, noise is used to refer to the sounds a person hears in the background at a given point in time. Noise can refer to cars passing when standing next to a road, birds chirping in the trees, or children yelling far away. As a general rule, louder noises are more problematic than softer noises (Passchier-Vermeer & Passchier, 2000). However, not all noise is negative, in fact, a person enjoying a late-night ride in their car might consider the resulting noise whilst someone hearing the same noise while attempting to fall asleep may disagree. It follows that how noise is perceived can depend on the individual experiencing said noise, for example twice as many people consider aeroplane noise at 60dB to be an annoyance than they do train noise at the same intensity (Miedema & Vos, 1998).

Regardless of how any specific person experiences a particular noise, consistent exposure to loud noise has been found to have a wide range of negative implications. As one might expect, the primary effects of noise are annoyance and stress (Basner et al., 2014). However further effects due to this stress such as worse sleep (Muzet, 2007), increased odds of cardiovascular disorders (Nriagu, 2019), and impaired performance at school (Lercher et al., 2003) or work (Ryherd et al., 2008) are relatively common. Furthermore, evidence has also been found that environmental noise may cause tinnitus (Śliwińska-Kowalska & Zaborowski, 2017) or potentially also loss of hearing, although such evidence has mostly been found in the context of occupational noise ((Chen et al., 2020)).

2.1.1 Noise intensity

Within the context of noise studies, sound volume (or loudness, L) is measured in dBA instead of simple decibels (dB) as is typical. This is important as measuring loudness in dB considers all sound frequencies equally, which human hearing does not. By applying A-weighting to the frequencies of a sound signal to obtain the loudness in dBA, the signal is modified to better, though still imperfectly, emulate human perception (Pierre Jr et al., 2004).

In addition to considering the particularities of human hearing, it is also important to consider the particularities of the noise itself. Since even the A-weighted SPL is only able to represent the noise intensity at a particular point in time noise studies generally represent an area by some manner of weighted average. A useful and commonly used representation is the day-evening-night loudness (L_{DEN}), proposed in the original CNOSSOS framework (Kephalopoulos et al., 2012). The L_{DEN} is a weighted average of the noise intensity throughout the day. It is derived as a weighted average of each respective component (specifically L_{day}, $L_{evening}$, and L_{night}), all of which are representations of the average loudness at that given part of the day.

2.2 Noise Monitoring

Because of the temporal nature of noise, it is insufficient to measure it on a singular occasion. Instead, noise measurement campaigns are conducted, allowing researchers to collect multiple noise measurements spread out over time. Noise monitoring campaigns have been conducted in a variety of ways, most of which can be subdivided into three overarching categories.

First, static noise measurements in which measuring setups are installed at various locations to constantly monitor noise levels in that area (Alías and Alsina-Pagès, 2019, Bellucci and Cruciani, 2016). Though static noise measurements provide reliable, detailed noise data and often have great temporal variation, their key drawback is that they are expensive to conduct.

Second, participatory noise studies avoid the expensive setups required for static studies by collecting noise data from participants who record the noise using personal equipment such as their smartphones (Jezdović et al., 2021). Allowing individuals to conduct measurements without oversight allows for data which is likely to be spatially and temporally varied, however, due to the lack of oversight the quality of data is also varied.

Finally, there are mobile measurement methods which are the subject of this study.

2.2.1 Mobile Measurement Methods

Mobile noise measurements are simply the opposite of static measurements in that they are conducted using sensors that are regularly moved. This mobile approach allows for great spatial heterogeneity within the collected data, and routes can be defined ahead of time to ensure data collection includes specific locations and a diverse range of areas.

Most mobile measurement studies, such as Alsina-Pagès et al. (2016) either apply a stop-and-go approach in which measurements are taken when the vehicle (in this case busses) has come to a stop. Some more recent studies have looked at taking measurements while moving but only using bicycles or on foot (Quintero et al., 2019, Can et al., 2014). In both cases, measurements are rarely taken at speeds above 5 km/h to limit the interference of own-vehicle and wind noise as these interfere with the noise measurements. Wind noise in particular is especially difficult to remove due to its inconsistent intensity and frequencies. Although many approaches have been suggested for removing wind noise (Walker & Hedlin, 2009), removing it from samples which contain more wind noise than actual signal is not recommended (Ecotiere, 2012).

2.3 Noise Mapping

Noise maps represent the amount of noise at each point in an area. Noise maps are invaluable tools for communicating the state of the local noise environment to the population and for establishing policy or planning infrastructure changes to combat noise pollution.

Noise maps can be created in a number of ways. Perhaps the simplest one, as shown in Banerjee et al. (2009), is by applying ordinary kriging (or a similar Gaussian process regressor) on a set of noise measurements to create a contour map of the noise distribution within an area. Such maps can obviously be improved by increasing the amount of measurement locations or the quality of the measurements themselves. However, another way to improve the noise maps is by accounting for external covariance such as traffic information or building density, which provide insight into the origins and propagation of the measured sound. An example of such a noise map can be found in Bellucci et al. (2018), in which measurements from the DYNAMAP project (a large static noise monitoring project) are combined with information on traffic and weather conditions.

2.3.1 Model Based Mapping

Because of the time and resource costs of noise mapping efforts, noise maps are often based solely on noise models instead of measurements. Especially within the EU where it is mandated that noise maps for large urban areas are updated once every five years, municipalities often opt to create these noise maps based on models. Many such maps are based on the CNOSSOS-EU (Common NOise aSSessment methOdS) framework, which provides guidelines on the creation of model-based noise maps (Kephalopoulos et al., 2012).

The general approach of model-based noise maps is to consider firstly specific sources of noise, and then secondly how that noise is propagated throughout an area. For each source of noise, whether it be road traffic, aircraft noise, or industrial noise, a separate model has to be created which provides a generalised model of how loud the noise from such a source is likely to be. For example, CNOSSOS provides different classifications of both vehicle types and road surfaces to allow for differentiation between a light vehicle on asphalt roads compared to a heavy truck on a gravel road. This estimate is then combined with an estimate of how frequently such a noise occurs, such that between two otherwise equivalent roads, the busier road creates more noise compared to a calmer one. After modelling how noise is created, its propagation is then modelled by considering the behaviour of the noise including reflection of sound on building surfaces or other surfaces.

Although the models used in creating noise maps can be very complex and result in impressively detailed noise maps, they must by necessity make simplifications and are unable to consider each possible noise source. For example, the CNOSSOS framework considers noise created by persons to always come from a building, all the noise created by a car to come from a single point near the road surface, and it cannot account for various minor noise sources such as wildlife. These limitations, despite being inevitable, mean that model-based noise maps always retain a level of uncertainty and are unable to accurately represent the specific characteristics of each area.

3 Study area and data description

This section introduces both the study area, as well as each of the datasets used in this study. In the case of the latter, the process of manually collecting noise data is first outlined after which each of the additional datasets used is introduced.

3.1 Study Area

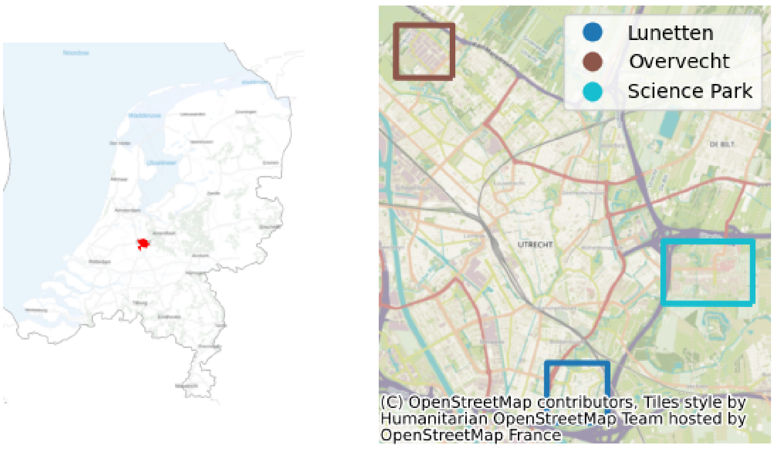


Figure 1: Overview of the study area

All noise sample collection was conducted in Utrecht, a city in the province of Utrecht, the Netherlands. However, the analysis is focused on three specific neighbourhoods, each representing a unique set of area characteristics and utilisation.

Utrecht is a city with roughly 377 thousand inhabitants spread out over 9377 km^2 of land surface area (CBS, 2024). It contains 110 neighbourhoods; including the city centre a mixed commercial-residential area, many both modern and older residential neighbourhoods, several industrial areas as well as the Utrecht University Campuses. Most noise in Utrecht is estimated to be caused by traffic, as every part of the city is connected via a network of major roads a small amount of which do have noise prevention measures such as sound barriers in place. The second major estimated major cause of noise is rail traffic. Utrecht is home to Utrecht Centraal, a major connecting station within the Netherlands, hence it contains a large amount of rail for either full-size passenger/freight trains or for light rail vehicles.

Due to the large area of Utrecht, this study focuses on three specific areas to reduce the amount of time necessary for data collection. In particular, it considers the following areas. *Lunetten*, a medium-sized residential area which neighbours both a highway and a rail track in the south-east of Utrecht. *Oud Zuilen*, a small industrial area comprised primarily of warehouses in the north-west of Utrecht. *Utrecht Science Park*, a University Campus containing various university lecture/office buildings, student accommodations, a University hospital as well as several pastures; it is adjacent to a major highway. Each of these areas is outlined in figure 1.

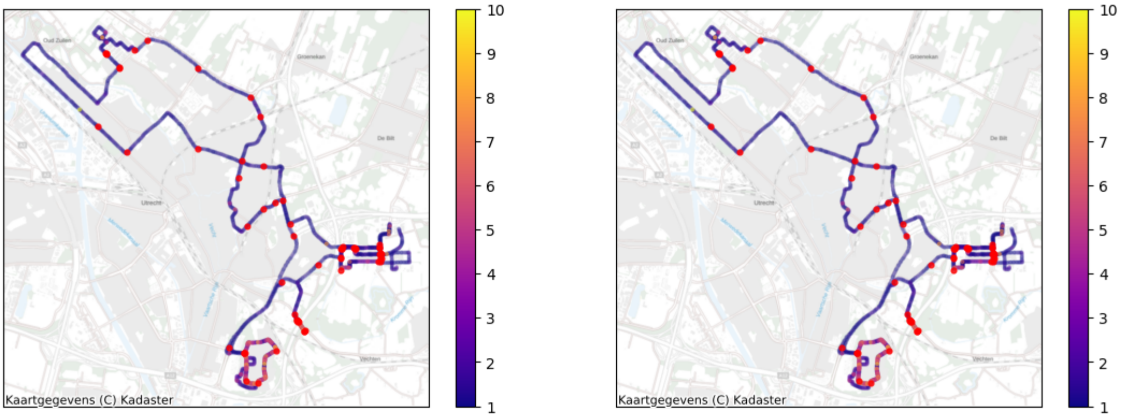
3.2 Data Collection

For the purposes of evaluating the validity of noise mapping via truly mobile recordings a sample of audio recordings taken while driving had to be collected. To this end, a car was equipped with an XLR microphone connected to a laptop and just under three hours of audio recordings were collected.

Figure 2: Microphone and interface on the roof of the collection vehicle

The microphone setup, shown in figure 2, consisted of an XLR microphone equipped with a windshield connected to a laptop via a USB audio interface. This interface was configured such that recordings were as loud as possible (to pick up smaller noises such as bird sounds) while ensuring recording levels did not peak above the maximum recording levels even when driving at high speeds. The setup was attached to the roof of an electric car which allowed for a reduction of the noise created by the collection vehicle. However, it should be noted that the collection vehicle also contained instruments for measuring air pollution. As such the noise-floor, the minimum loudness recorded was 60dBA as measured while the car was idling in an otherwise silent garage. This value will also be used to calibrate the noise measurements during audio-processing in *section 4.1*.

All recordings were conducted over a period of four hours on the same weekday morning. Hence the data and resulting noise maps are only representative of a small time frame (although we consider this sufficient for exploring the validity of this approach, suggestions for generalisability are given in the discussion, *section 6*). Recordings were collected at variable speeds, which affected the noise levels measured due to wind interference. The collection route was focused on visiting each of the three target areas, specific roads were revisited multiple times to increase the certainty of the noise measurement. An overview of the collection route is given in *figure 3*. The average speed at each location is also indicated, which was calculated based on the distance travelled in the 5 seconds surrounding each GPS measurement point (performed using a rolling window over GPS locations smoothed via a simple Kallman smoothing approach).



(a) Speed per location on collection route

(b) No. of observations per location on collection route (red circles indicate > 10 observations)

Figure 3: Overview of the collection route

3.3 Additional Noise Data

For reference purposes estimated noise values were included to compare against the measured/predicted values derived in this study. The reference data used belongs to the RIVM (Dutch National Institute for Public Health and the Environment)[1]. This dataset combines noise models for various sources including road, rail, and air traffic as well as industrial sources to create a fairly comprehensive L_{den} value represented using 10m x 10m raster grid. Although this model-derived data is likely imperfect due to the limitations mentioned in *section 2.3.1*, and lacking various possible noise sources it should provide a solid point of comparison compared to the collected noise levels.

3.4 Geographic Data

Most of the additional data used in this study was collected from OpenStreetMap OpenStreetMap contributors, 2017 (OSM), using the OSMnx package[2] for Python. Road segments were collected from the Utrecht area, relevant attributes were collected (road_id, max_speed, highway, bridge, and junction - the latter three defining the type of road (surface)). Road segment data was then added onto intersecting reference noise data nodes. Only nodes with road data were retained such that when creating a scatter plot of the reference noise levels a noise map can be seen. This resulted in 905 total nodes for the Lunetten area, 279 nodes for Oud Zuilen, and 1179 nodes for Utrecht Science Park.

After creating each of the nodes, supplemental data was added further describing their characteristics. Using the railway network data from OSM each node was assigned the distance to the nearest train track. Furthermore, a distance_to_recreation parameter was added based on the distance of each node to the nearest recreational amenity in OSM (for example a bar, restaurant, community_centre, or music_venue). Finally, for each node, it was indicated whether or not it interacted with one or more of the following land use area types: residential, commercial, or industrial.

3.5 Building Height Data

To account for noise reflection on building surfaces, building height was included in the data. Although it can be estimated using satellite imaging, in this case a prior analysis was used. Specifically, the 3D BAG dataset[3] was used to retrieve an estimated (70^{th} percentile) height for each building in the area. For each road node, the highest nearby building area was then used.

4 Methodology

In this section, different methods for processing of the collected audio recordings are considered and detailed. The resulting loudness values derived are then used to create several regression kriging

[1] https://data.overheid.nl/dataset/7133-geluid-in-nederland--lden-#panel-resources
[2] https://osmnx.readthedocs.io/en/stable/index.html
[3] https://3d.bk.tudelft.nl/projects/3dbag/

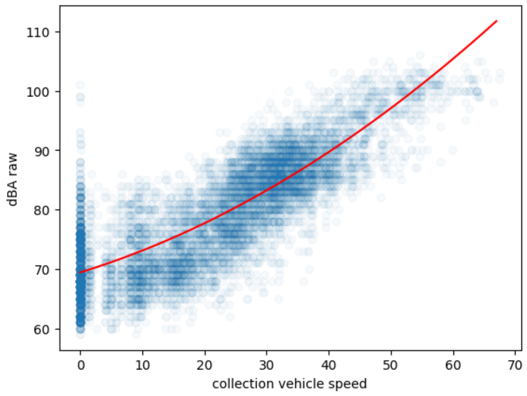


Figure 4: Regression line showing the correlation of the speed of the vehicle with the measured Noise levels

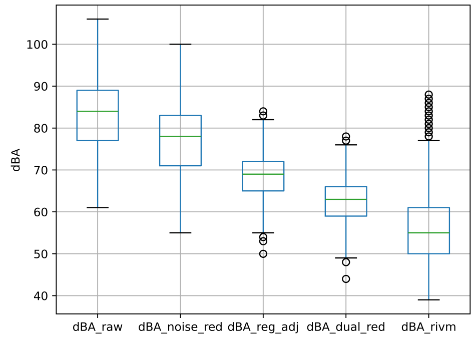


Figure 5: Overview of the value ranges of the noise intensity after different preprocessing procedures

models. First the model tuning and selection are discussed, after which analysis methods for the resulting models are described.

4.1 Audio Processing

To derive loudness values for the audio signal this signal was spit into 1-second long sections each of which was then reweighed using A weighting[4] to better represent human perception and then converted to a digital loudness value (in dB FS) via equation (1). This equation is made up of the standard formula for conversion from signal to loudness (20 log) but also includes smoothing of the original signal by means of root mean square smoothing (RMS). The digital signal is then converted into a final dBA SPL signal (A-weighted DB relative to sound pressure level) which represents the actual sound pressure level as experienced by humans. This is done by re-calibrating the values to those measured with a pre-calibrated dBA measurement device.

$$L = 20\log_{10}(\mathrm{RMS}(s)) \tag{1}$$

Though the above describes the procedure of direct conversion from a digital signal to SPL, this method fails to consider the presence of own-vehicle noise which likely interferes with the measured noise levels. To account for this two different methods were attempted. The resulting dB levels obtained through each method separately as well as a combined approach and the original values will be used and evaluated separately in later sections.

Firstly to deal with the interfering noise generated by the collection vehicle (and the other equipment in said vehicle) noise reduction was applied to the signal before conversion. Although several methods (such as wavelet thresholding or removal via Deep Learning) were considered, spectral gating was selected based on the requirement that the original signal must be left untouched as much as possible (which most of the methods mentioned/considered do not meet). Spectral gating works by creating a mask which, for each frequency, only allows for the signal to be included if it surpasses a certain threshold. Said threshold was set based on a recording containing only interfering noise. For the application of the noise reduction a noise clip was taken while the collection vehicle was idling in a garage, which was then used for noise reduction via the noisereduce package[5]. The noise-reduced signal was then once again converted into a separate set of dBA values.

The second measure pertains to reducing interfering noise generated by the wind hitting the microphone, the effect of which is multiplied at greater speeds. To analyse the effect of collection vehicle speed on the measured loudness (dBA raw, as in after direct conversion without noise reduction) a regression analysis was conducted as shown in *figure 4*. A clear correlation was found (R2: 0.694 without noise reduction and R2: 0.689 with), while deviation from the regression line remained roughly consistent at all driving speeds. Using this regression a new set of loudness values was created by adjusting each measured loudness value based on its speed using the regression

[4] https://github.com/endolith/waveform_analysis

[5] https://pypi.org/project/noisereduce/, parameters used:
nr.reduce_noise(y=signal, sr=sr, y_noise=noise_clip, stationary=False, prop_decrease=0.5)

Regression Model	paramater	value range	optimal value
Random Forest	n_estimators	[50, 100, 200]	50
	max_depth	[None, 10, 20, 30]	10
	min_samples_split	[2, 5, 10]	10
	min_samples_leaf	[1, 2, 4]	4
	max_features	[2, 4, 6, 8, 10]	2
Gradient Boosting	n_estimators	[50, 100, 200]	50
	learning_rate	[0.01, 0.1, 0.2, 0.3]	0.01
	max_depth	[3, 5, 7]	7
	min_samples_split	[2, 5, 10]	10
	min_samples_leaf	[1, 2, 4]	2
	max_features	[2, 4, 6, 8, 10]	2
Gaussian Process	kernel	[RBF, Matern, RationalQuadratic, WhiteKernel]	WhiteKernel(1)
	alpha	[1e-07, 0.0001, 0.01, 0.1]	0
	n_restarts_optimizer	[0, 1, 5]	0

Table 1: Model inputs

Model	RMSE base	R2 base	RMSE tuned	R2 tuned
LinearRegression()	6.5722	-0.7943	6.5320	-0.7723
RandomForestRegressor()	5.0658	-0.0660	4.7151	0.0764
GradientBoostingRegressor()	5.1355	-0.0955	4.7775	0.0518
GaussianProcessRegressor()	67.6580	-189.1536	4.9539	-0.0194

Table 2: Model scores before and after tuning

coefficient. This was done using both the raw dBA values, and the noise-reduced dBA values. The adjusted dBA ranges, which are much closer to the expected ranges, are shown in figure 5.

4.2 Model Selection

To generalise the measured values to a greater area a regression kriging model was used. The advantage of the regression kriging model is that it both considers the measurements present and the specific characteristics of the road and its surrounding area at each location. This allows it to generate noise maps for an entire area without access to measurements from each road, instead only requiring a subsection of each area and each road type.

A regression kriging model works by first learning a normal regression model, such as OLS or a decision tree, and then applying kriging to the residual errors. To select the most appropriate model an adapted hyperparameter tuning approach was adopted. Several baseline regression models were considered, specifically: Linear-, Gaussian Process-, Random Forest-, and Gradient Boosting Regressors. A range of parameter values was defined, both for each regression model, as well as for the kriging model itself. A k-fold hyperparameter tuning approach using grid search was then used which selected the optimal set of parameter values based on the resulting RMSE of the final kriged model, regardless of regression model performance.

The tuning method above was then performed using the regression-adjusted dBA data from all three areas, evaluated. The regression models, parameter ranges, optimal parameters, and final RMSE values for each regression model are shown in *table 2*. The best model found was EXPLAIN MODEL HERE.

4.3 Model Input and Analysis

Using the model parameters found in the previous section (*section 4.2*), various models were trained for the purposes of comparing the performance of each approach. During audio processing (*section 4.1*) four sets of noise loudness values were abstracted: raw dBA (signal converted directly to SPL), noise-reduced dBA, regression-adjusted dBA, and dual-reduced dBA (both noise-reduced and then regression-adjusted). Each set was then separated into two new sets: one containing measurements at all speeds, and one containing measurements at own-vehicle speeds of less than 20 km/h. Finally for each set two models were trained, one trained using data from each area, and one trained on data from only the Lunetten area as this area was covered most extensively during

model input			performance on heldout data (per area)							
			Lunetten		Overvecht		Science park		Overall	
dBA	speed	area	r2	rmse	r2	rmse	r2	rmse	r2	rmse
dBA_raw	any	all	0.550	6.397	-0.355	5.813	0.254	6.932	0.310	6.670
dBA_noise_red	any	all	0.546	6.419	-0.183	5.443	0.253	6.954	0.319	6.637
dBA_reg_adj	any	all	0.457	3.989	-0.240	4.296	0.079	4.807	0.168	4.597
dBA_dual_red	any	all	0.451	4.047	-0.329	4.462	0.095	4.774	0.169	4.609
dBA_raw	lt20	all	0.069	5.023	-0.052	4.222	-0.180	5.514	0.044	5.144
dBA_noise_red	lt20	all	0.060	5.084	-0.054	4.280	-0.200	5.494	0.033	5.163
dBA_reg_adj	lt20	all	0.096	4.257	-0.059	4.202	-0.116	4.910	0.078	4.571
dBA_dual_red	lt20	all	0.055	4.407	-0.065	4.276	-0.099	4.887	0.066	4.621
dBA_raw	any	Lun	0.543	6.447	-0.328	5.754	-0.091	8.381	0.080	7.699
dBA_noise_red	any	Lun	0.548	6.405	-0.377	5.872	-0.120	8.516	0.057	7.806
dBA_reg_adj	any	Lun	0.441	4.045	-0.206	4.237	-0.465	6.064	-0.191	5.498
dBA_dual_red	any	Lun	0.448	4.059	-0.192	4.225	-0.465	6.074	-0.186	5.506
dBA_raw	lt20	Lun	-0.036	5.300	-1.667	6.722	-1.486	8.004	-0.736	6.933
dBA_noise_red	lt20	Lun	-0.001	5.246	-1.558	6.668	-1.419	7.800	-0.676	6.798
dBA_reg_adj	lt20	Lun	-0.073	4.637	-0.818	5.507	-0.776	6.196	-0.368	5.569
dBA_dual_red	lt20	Lun	-0.027	4.595	-0.712	5.422	-0.782	6.223	-0.350	5.557

Table 3: Model scores (R^2* & $RMSE$) for each model per area

* R2 scores for fitted models can be negative indicating performance below the baseline (Cameron and Windmeijer, 1997)

data collection. In the end, the sixteen models shown in Table 3 were trained. With models being differentiated by audio pre-processing, measurement speed, and training-area approaches. A full overview of the noise maps generated by these is given in appendix B, future sections will show highlights where relevant.

To assess the quality of each noise map number of factors were considered. For each model, the prediction for each area is looked at, representing how well the model predicts the measured values. Then the model performance at generating a noise map is evaluated by comparing the generated noise map to the reference noise values through correlation and visual analyses.

5 Results

In this section, the performances of various modelling approaches described in *section 4* are described both in terms of model performance and qualitatively in comparison between modelling approaches as well as to existing noise models.

5.1 Model Performance

After training the Random Forest - Regression Kriging model (tuned in section 4.2) on different inputs, both overall performance and performance for each sub-area were tested. The results of these tests are given in table 3, several trends worth discussing can be identified in this table.

First, models trained based on 'reg_adj' and 'dual_red' dBA values outperform the other models, since those are unable to account for the wind interference noise. The application of noise reduction (from 'raw' to 'noise_red') appears to lower the dBA values at each location by a consistent amount. As such, patterns present in the data are not affected and model performance between the same models with and without noise reduction is comparable.

Second, models trained using only speeds below 20km/h have significantly worse R^2 scores, but improved RMSE values when compared to models without regression adjustments to account for wind interference. The latter is to be expected as limiting the maximum speed for measurements and the regression adjustment both intend to account for interfering wind noise. The former method however does this in a way that heavily reduces the available input data, which is the likely reason for the decreased R^2 scores.

Third, models trained using only data from the Lunetten area perform well in the Lunetten sub-area but perform poorly in other areas. Their performance is trivially worse within Lunetten when compared to models trained on all areas, likely due to the latter having access to more training data (or due to randomness in the random forest regressor).

Fourth, performance within each sub-area is fairly consistent but performance between sub-areas shows clear differences. The R^2 score for each model is best in the Lunetten area, and worst in Overvecht, the RMSE on the other hand is relatively consistent between areas. It should be noted that the RMSE values between areas are not perfectly comparable as possible value ranges in each sub-area differ. Similarly, the R^2 score is not perfectly reliable, as the kriging step in each model adjusts values based on the error present thereby tempering based on the R^2 score produced after regression.

Differences between models trained on measurements from any speed, and identical models trained on only measurements at lower speeds perform almost identically, both in terms of scoring as well as their resulting visualisations.

5.2 Regression Kriging Model Comparisons

Visualising the results of the regression kriging models shows several differences between each approach not revealed by simple statistical analysis. This subsection only includes visualisations which directly support present statements. An overview of all the model results visualisations is available in appendix B.

(a) using raw dBA values

(b) using noise reduced dBA values

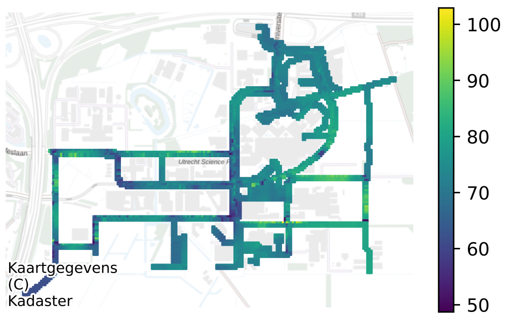
(c) using regression adjusted dBA values

Figure 6: Lunetten area noise maps generated based on different audio preprocessing dBA values

A general pattern can be observed in the differences between audio pre-processing methods. Values shown in models using noise reduction to reduce interference from the collection vehicle are consistently slightly lower compared to their non-reduced counterparts. An example of this can be seen in figure 6, which shows the results of modelling the raw dBA values compared to the noise-reduced values in the Lunetten sub-area. The patterns present in the (a) and (b) figures are remarkably consistent with the sole difference being the lower dBA values produced by the noise-reduced models. Similar behaviour can be observed in Appendix B when comparing the regression-adjusted values to the dual reduced values (as dual reduced values were noise reduced, then regression adjusted), in both the shown scenario as well as in other sub-areas.

Another general observation regarding audio processing is that although regression-adjusted graphs show similar trends compared to graphs from non-adjusted models, the differences between high and low noise values are much smaller. Figure 6 also shows a comparison between the raw noise values and the regression-adjusted values for the same area as before. Here it can be observed that although higher values are still predicted for the main road, the differences between these and predicted values on smaller roads are much smaller. Despite this, general trends such as the higher noise values near the train tracks in the northeast are still present.

An additional set of models was trained using only measurements taken when the speed of the collection vehicle was less than 20km/h, with the intention of limiting the amount of wind-based noise interference. Figure 7 shows the results of the noise-reduced models before and after limiting inputs to those collected at speeds below 20km/h, this time for the Science Park sub-area. Again it can be observed that overall trends are very comparable, with the main difference being that most noise values are much lower, similar to the effect of regression adjustment. Two key differences between the results of regression adjustment exist. Firstly the speed-limited model contains a small set of high noise values, which in the regression-adjusted model are less noticeable. Second, although not visible in the graphs, reducing the model input to only measurements taken at speeds below 20 km/h significantly reduces the amount of available data points (from 2349 down to 787 in total), as well as introducing bias in the collected data as certain locations are more likely to result in lower driving speeds.



(a) Based on any speed (b) Based on only speeds < 20km/h

Figure 7: Noise maps of the Science Park area using noise reduced dBA values

Figure 8: Regression Kriging Predictions for dBA_dual_reduced (trained on all areas)

Within model differences can be observed between the modelled sub-areas. Figure 8 shows the processed noise values for each sub-area (with both noise reduction and regression adjustment applied). Overall trends within the Lunetten area are the most intuitive with higher values on main roads and lower noise values on the smaller roads. The Overvecht area behaves similarly but with a large outlier in the southern part of the plot, as that area was relatively busy when the observation vehicle passed by. The model values for the Science Park area, although consistent with measurements are less intuitive, showing higher values on calm roads on the edge of the campus and lower values on the busy main roads at the centre. Similar behaviour also occurs with less audio processing as can be observed in Appendix B.

Figure 9: Regression Kriging Predictions for dBA_raw (trained on Lunnetten sub-area)

Finally, predictions for each area were also made using only training data from the Lunetten sub-area. Unsurprisingly based on the performance scores from the previous section, the performance of these models, an example shown in figure 9, is rather lacklustre. Without reference data points in the non-Lunetten sub-areas for the kriging algorithm, the results in those areas show only very broad trends with little difference between different roads or areas.

5.3 Reference Model

Although there exists no ground truth which can be used to compare either the measured or modelled values against. It is possible to compare the noise maps generated for this study to those from existing noise models. In particular, in the subsection, generated noise maps are compared against maps derived from the reference noise data described in section 3.3. These reference noise maps, shown in figure 10, should do a good job reflecting noise generated by modelled variables. However, they represent L_{DEN} values and do not include all possible noise sources, and as such are neither directly comparable nor perfect points of comparison.

Figure 10: Noise maps for each sub-area based on reference dBA values

Several clear differences between the values shown in the reference model can be observed. The contrast between noise levels for major roads and minor roads is much larger, especially compared to regression kriging models based on regression-adjusted dBA values. In general, values based on the reference model are mostly traffic-based, with parking lots and calm roads away from buildings having lower noise levels. Noise intensity in busier areas ranges around 65 to 75 dBA, which is similar to the values in the dual-reduced model shown earlier. Smaller streets with less traffic however have lower noise intensity values, compared to what can be observed in the generated models. Non-traffic-based noise sources, however, are less influential on the reference model. For example, the effect of the train tracks in the northeast of Lunetten is present but more local, and the effect of the hospital generator in the east of the Science Park area is not modelled.

model input			**reference model comparison**			
dBA	speed	area	corr (r)	Δmin	Δmax	rmsd
dBA_raw	any	all	0.145	-12.7	52.4	21.577
dBA_noise_red	any	all	0.150	-18.2	46.3	16.354
dBA_reg_adj	any	all	0.114	-23.3	34.8	12.375
dBA_dual_red	any	all	0.117	-29.5	29.0	8.621
dBA_raw	lt20	all	0.220	-10.0	43.1	16.083
dBA_noise_red	lt20	all	0.217	-16.0	37.0	11.176
dBA_reg_adj	lt20	all	0.264	-16.9	33.0	11.250
dBA_dual_red	lt20	all	0.246	-23.2	26.9	7.828
dBA_raw	any	Lun	0.369	-7.1	39.7	21.792
dBA_noise_red	any	Lun	0.357	-13.3	34.8	16.534
dBA_reg_adj	any	Lun	0.403	-18.5	28.4	12.613
dBA_dual_red	any	Lun	0.396	-24.2	22.7	8.382
dBA_raw	lt20	Lun	0.483	-2.5	33.4	17.519
dBA_noise_red	lt20	Lun	0.487	-8.5	27.2	12.150
dBA_reg_adj	lt20	Lun	0.457	-9.0	27.4	11.809
dBA_dual_red	lt20	Lun	0.455	-15.0	21.3	7.775

Table 4: Comparisons between each created model and the reference noise model

Table 4 shows an overview of how each model compares to the reference noise model. The regression kriging models, especially before audio processing, are much more likely to overestimate noise values (based on the reference model) than to underestimate them. After audio processing, a marked improvement in terms of root mean square difference (rmsd) can be observed showing audio processing brings the measured values closer to the expected values. Regardless the difference range (Δmax −Δmin) remains similar (roughly 60 at any speed with all areas). Although all models are positively correlated with the reference model, regression kriging models trained on all available data show weak positive correlation at best. Correlation is improved when using only measurements taken at lower speeds as well as when only measurements from the Lunetten area.

6 Discussion

Despite the limited sample of noise data available for this study, noise maps produced by the regression kriging models in the previous section performed reasonably well at recognising spatial noise patterns. The resulting noise maps show similar trends to those present in the reference noise maps while retaining specific patterns encountered during measurement. Although these results are promising regarding the viability of using noise measurements collected from a moving vehicle, there are still several challenges, some of which have been partially addressed within the methodology, which limit the effectiveness of this approach at this time. In this section, these challenges at each step of the process, namely noise data collection, audio processing, and noise modelling, are discussed together with possible solutions and open problems.

6.1 Challenges in Data Collection

The core problem in the creation of noise maps is the immense spatial and temporal diversity of noise. Collecting noise data from a moving vehicle makes it relatively simple and inexpensive to collect spatially diverse data. It is not however trivial, this becomes especially clear when looking at the noise maps produced of the Overvecht sub-area in the results section (5.2). Data in this area was only collected in a small area and on roads with similar sizes and surroundings, which was likely the reason why models tended to produce very homogeneous results for the entirety of the Overvecht area. This finding is consistent with (Quintero et al., 2019) and shows the importance of repeated measurements. Another example is the large spot of road area in the middle of the Science Park sub-area, which in the reference noise maps has significantly less noise compared to surrounding roads. This is likely a correct prediction from the reference noise map as those road segments represent a parking lot. The data available for this study however doesn't contain any noise samples collected from parking lots and therefore the model could not identify this. Based on this information, it is recommended that future attempts visit a diverse set of roads and road contexts in each area measured.

Data collection methods used for this study are severely lacking in temporal variation. Although it was not possible to address within this study, it is a key consideration for any larger-scale implementation. Simply including data from different points of time in the model will create a more representative model resulting in a noise map which should be a little more correlated with the L_{den} models. However such an approach risks over weighting certain hours of the day, particular days of the week, or even specific seasons depending on when measurements are conducted. Instead, it is likely wiser to create multiple regression kriging models each representing a specific segment of time and combining the results of these models to create a correctly weighted L_{den} noise representation. The simplest reliable variation of this approach would be to create three different models; one each for day, evening, and nighttime measurement. This approach mirrors how L_{den} represents a weighted average of day, evening, and nighttime noise levels, but it could also be expanded to consider different days of the week, holidays, or seasons.

6.2 Challenges in Audio Processing

When collecting noise data from a moving vehicle, two interfering sources of noise must be considered. Firstly, the collection vehicle itself creates noise which is picked up by the measurements but should not be included in the final noise map. Second, interference from the wind hitting the microphone the effect of which is amplified when the collection vehicle is driving at higher speeds.

This study has attempted to account for the interfering noise created by the collection vehicle through the application of noise reduction. Resulting noise models and maps indicate this approach has the intended effect of lowering the noise floor allowing for lower minimum measurements which are less loud than the collection car is even when idling. Despite having the intended effect, listening back to the audio files after noise reduction revealed that the collection vehicle noise was not completely removed and signals from other sources were also affected. As a result, generated noise maps were louder than the reference model in areas which are likely to be relatively silent. However, due to the reference noise values being in L_{DEN} and failing to account for environmental sources, correct values in such areas are likely higher than the 45dBA indicated in reference noise maps. To address this issue and bring the noise floor closer to the correct value, the application of more advanced noise reduction algorithms which affect the original signal as little as possible may be considered.

Interference from wind noise adds an additional layer of complexity to this problem because, based on section 2.2.1, the large amount of wind noise and its inconsistent intensity and frequencies mean it is likely impossible to remove it while retaining the original signal. Because of this wind noise was instead accounted for by lowering dBA levels of measurements during post-processing to account for the stronger wind noise experienced at higher speeds. Although this approach did again come close to having the intended effect it also introduced additional problems. Noise maps based on regression-adjusted measurements showed limited ranges of possible values. As most measurements near roads with high maximum speeds were recorded while driving at those high speeds, most such measurements were then lowered by the regression creating lower predicted noise levels than was likely correct. Two possible approaches to resolving this issue are to either measure the wind speed and regress measured noise levels based on this or to collect a sample of both static and mobile noise measurements at roads with different maximum speeds to attempt and find a correlation here as well.

To address both the collection-vehicle and wind interference, a stop-and-go approach similar to Alsina-Pagès et al., 2016, may be considered. Only taking measurements while the collection vehicle is standing still with the car turned off removes almost all possible sources of interference in the measurements. However, this does come at the cost of several disadvantages. Firstly, significantly less data is collected in general and the collected data cannot be as spatially diverse as data collected while moving. Additionally, there is an inherent risk in this approach as it creates a sampling bias in which only locations where it is feasible to stop the collection vehicle can be measured. Thus, a fully mobile approach is preferable so long as audio processing can be improved.

6.3 Challenges in Modelling

Section 5.2 and 5.3 thoroughly discussed the maps and models created in this study. An observation made in the latter section is the lacklustre performance of models trained only on data from a single sub-area. These models only show the general trends found in the regression model as there are no available measurements such that kriging can be applied over error residuals. As such these results can be used to gain insight into the effectiveness of the base regression model. A key property which can be observed in these maps is that there is only a small set of different predicted values. This can be attributed to the limited amount of data available to the regression model about each road segment. Improving the available model input beyond simple data points such as maximum speed should help improve the performance of the regression model, assuming sufficient data is available for each different road type. An important data point to include would be traffic information, which has repeatedly been used with great success in prior noise modelling, hence its inclusion in Kephalopoulos et al., 2012.

Considering the discussed performance of the regression model, the remaining models clearly show the strengths of the kriging approach. However it should be considered that given a low amount of available measurements, a singular measurement may be generalised over a large area but the kriging algorithm. Although this is not a problem for the noise maps in this study which represent loudness during a single measurement session when creating more general noise maps, such as L_{den}, this becomes an undesirable behaviour. Though this behaviour should never occur given sufficient temporal variation in the measurements at each measured location, it should still be considered whenever more general noise maps are created.

Finally, some observations regarding model performance compared to the reference models should be mentioned. The results section (5) indicated that models based on all measurements (not filtered on either speed or sub-area) produced more insightful models. Despite this, the correlation between reference noise models and filtered input models was significantly higher. The reason for this is that, although correlation with existing noise maps provides an important metric by which to evaluate whether or not the produced models capture similar types of trends, perfect correlation is undesirable as the intent is to also capture noise patterns which the reference models are unable to capture due to them not utilising real-world measurements. Thus while the correlation of non-filtered models is lower than desired, the inability of the filtered models to show observed deviations from the trend can be considered a larger issue.

7 Conclusion

Combining noise data collected via fully mobile noise collection with a regression kriging approach allows for the creation of noise maps which are both able to capture general trends similar to other noise modelling approaches and adjust for the particular distributions of noise observed in each sub-area. The resulting models are able to create such maps even using very limited data, however, the quality is dependent on measurements being taken multiple times for each area. Preferably these measurements would be taken at different times to increase temporal variety in measurements allowing for a fairer representation of the overall noise level in the measured area.

Collecting audio data from a moving vehicle is challenging due to interfering noise created by the collection vehicle and the wind noise generated. Within this study, it was shown it is possible to partially remove this interference through the application of noise reduction techniques and adjusting the resulting noise values based on the collection vehicle speed. However, the application of these techniques also affects the measured signal, resulting in imperfect measurements. Improving the reliability of the measurements themselves, for example, through adjusting based on wind speed or predicting wind interference based on a separate set of static noise measurements, is the key area of improving the viability of fully mobile noise measurement methods.

The application of regression kriging based on the measured noise data and a small set of external variables produced promising results. The resulting noise maps identified general trends for road types and represented particularities of each sub-area. Two main limitations were observed here. First, due to the challenges discussed regarding the data collection method, the range of predicted noise values was limited and variation between road types was lower than is likely correct. Second, the regression model was limited by the available external variables and would've likely performed better if it had had access to (simulated) traffic data and further data regarding the area surrounding each road segment.

Despite the mentioned shortcomings, the models created within this study were able to produce promising noise maps. Especially given the number of challenges inherent to measuring noise from the roof of a collection vehicle. Given a larger sample of measurements, and improved model inputs it is likely that the proposed methodology produces noise maps which correctly identify both general trends and reflect real-world measurements increasing certainty compared to noise maps based exclusively on models.

Acknowledgements

Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org

References

Alías, F., & Alsina-Pagès, R. M. (2019). Review of wireless acoustic sensor networks for environmental noise monitoring in smart cities. *Journal of sensors*, *2019*(1), 7634860.

Alsina-Pagès, R. M., Hernandez-Jayo, U., Alías, F., & Angulo, I. (2016). Design of a mobile low-cost sensor network using urban buses for real-time ubiquitous noise monitoring. *Sensors*, *17*(1), 57.

Banerjee, D., Chakraborty, S., Bhattacharyya, S., & Gangopadhyay, A. (2009). Appraisal and mapping the spatial-temporal distribution of urban road traffic noise. *International Journal of Environmental Science & Technology*, *6*, 325–335.

Barham, R., Chan, M., & Cand, M. (2010). Practical experience in noise mapping with a mems microphone based distributed noise measurement system. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, *2010*(6), 4725–4733.

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., & Stansfeld, S. (2014). Auditory and non-auditory effects of noise on health. *The lancet*, *383*(9925), 1325–1332.

Bellucci, P., & Cruciani, F. R. (2016). Implementing the dynamap system in the suburban area of rome. *Inter-Noise and Noise-Con Congress and Conference Proceedings*, *253*(3), 5518–5529.

Bellucci, P., Peruzzi, L., & Zambon, G. (2018). Life dynamap: Making dynamic noise maps a reality. *Proceedings of the Euronoise*, 1181–1188.

Cameron, A. C., & Windmeijer, F. A. (1997). An r-squared measure of goodness of fit for some common nonlinear regression models. *Journal of econometrics*, *77*(2), 329–342.

Can, A., Dekoninck, L., & Botteldooren, D. (2014). Measurement network for urban noise assessment: Comparison of mobile measurements and spatial interpolation approaches. *Applied acoustics*, *83*, 32–39.

CBS, S. N. (2024). Kerncijfers wijken en buurten 2023.

Chen, K.-H., Su, S.-B., & Chen, K.-T. (2020). An overview of occupational noise-induced hearing loss among workers: Epidemiology, pathogenesis, and preventive measures. *Environmental health and preventive medicine*, *25*(1), 65.

Ecotiere, D. (2012). Estimation of uncertainties due to the wind-induced noise in a screened microphone. *Acoustics 2012*.

Jezdović, I., Popović, S., Radenković, M., Labus, A., & Bogdanović, Z. (2021). A crowdsensing platform for real-time monitoring and analysis of noise pollution in smart cities. *Sustainable Computing: Informatics and Systems*, *31*, 100588.

Kephalopoulos, S., Paviotti, M., & Anfosso-Lédée, F. (2012). Common noise assessment methods in europe (cnossos-eu).

Kerckhoffs, J., Hoek, G., Gehring, U., & Vermeulen, R. (2021). Modelling nationwide spatial variation of ultrafine particles based on mobile monitoring. *Environment International*, *154*, 106569.

Lercher, P., Evans, G. W., & Meis, M. (2003). Ambient noise and cognitive processes among primary schoolchildren. *Environment and Behavior*, *35*(6), 725–735.

Miedema, H. M., & Vos, H. (1998). Exposure-response relationships for transportation noise. *The Journal of the Acoustical Society of America*, *104*(6), 3432–3445.

Muzet, A. (2007). Environmental noise, sleep and health. *Sleep medicine reviews*, *11*(2), 135–142.

Nriagu, J. O. (2019). *Encyclopedia of environmental health*. Elsevier.

OpenStreetMap contributors. (2017). Planet dump retrieved from https://planet.osm.org.

Passchier-Vermeer, W., & Passchier, W. F. (2000). Noise exposure and public health. *Environmental health perspectives*, *108*(suppl 1), 123–131.

Pierre Jr, R. L. S., Maguire, D. J., & Automotive, C. S. (2004). The impact of a-weighting sound pressure level measurements during the evaluation of noise exposure. *Conference NOISE-CON*, 12–14.

Quintero, G., Aumond, P., Can, A., Balastegui, A., & Romeu, J. (2019). Statistical requirements for noise mapping based on mobile measurements using bikes. *Applied Acoustics*, *156*, 271–278.

Ryherd, E. E., Waye, K. P., & Ljungkvist, L. (2008). Characterizing noise and perceived work environment in a neurological intensive care unit. *The Journal of the Acoustical Society of America*, *123*(2), 747–756.

Śliwińska-Kowalska, M., & Zaborowski, K. (2017). Who environmental noise guidelines for the european region: A systematic review on environmental noise and permanent hearing loss and tinnitus. *International journal of environmental research and public health*, *14*(10), 1139.

Walker, K. T., & Hedlin, M. A. (2009). A review of wind-noise reduction methodologies. *Infrasound monitoring for atmospheric studies*, 141–182.

A Code

All of the code used in this study is available through GitHub at https://github.com/Floris-C/noise-thesis

B Generated Noise Maps for each Model

This appendix contains the noise maps for each area from all of the models trained during this study. Sub-areas shown from left to right are Lunetten, Overvecht, and Science Park respectively. Colour bars are consistent per model but change between models (and therefore also between figures) to make each figure as legible as possible.

Figure 11: Regression Kriging Predictions for dBA_raw (trained on all areas)

Figure 12: Regression Kriging Predictions for dBA_noise_red (trained on all areas)

Figure 13: Regression Kriging Predictions for dBA_reg_adj (trained on all areas)

Figure 14: Regression Kriging Predictions for dBA_dual_red (trained on all areas)

Figure 15: Regression Kriging Predictions for dBA_raw_slt20 (trained on all areas)

Figure 16: Regression Kriging Predictions for dBA_noise_red_slt20 (trained on all areas)

Figure 17: Regression Kriging Predictions for dBA_reg_adj_slt20 (trained on all areas)

Figure 18: Regression Kriging Predictions for dBA_dual_red_slt20 (trained on all areas)

Figure 19: Regression Kriging Predictions for dBA_raw (trained on Lunnetten sub-area)

Figure 20: Regression Kriging Predictions for dBA_noise_red (trained on Lunnetten sub-area)

Figure 21: Regression Kriging Predictions for dBA_reg_adj (trained on Lunnetten sub-area)

Figure 22: Regression Kriging Predictions for dBA_dual_red (trained on Lunnetten sub-area)

Figure 23: Regression Kriging Predictions for dBA_raw_slt20 (trained on Lunnetten sub-area)

Figure 24: Regression Kriging Predictions for dBA_noise_red_slt20 (trained on Lunnetten sub-area)

Figure 25: Regression Kriging Predictions for dBA_reg_adj_slt20 (trained on Lunnetten sub-area)

Figure 26: Regression Kriging Predictions for dBA_dual_red_slt20 (trained on Lunnetten sub-area)