Research Article

Felix Schlatter*, A. Piquerez, M. Habermacher, M. S. Ragettli, M. Röösli, M. Brink, C. Cajochen, N. Probst-Hensch, M. Foraster, and J.-M. Wunderli

Validation of large scale noise exposure modelling by long-term measurements

Abstract: Large scale noise exposure modelling is used in epidemiological research projects as well as for noise mapping and strategic action planning. Such calculations should always be accompanied by an assessment of uncertainty, on the one hand to check for systematic deviations and on the other hand to investigate the sources of uncertainty to address them in future studies.

Within the SiRENE (Short and Long Term Effects of Transportation Noise Exposure) project, a large scale nationwide assessment of Switzerland's road, railway, and aircraft noise exposure was conducted for the year 2011.

In the present follow-up study, we equipped 180 sleeping and/or living room windows with sound level meters for one week. The resulting dataset was used to validate noise exposure modelling within SiRENE. For the noise metric L_{DEN} the comparison revealed a difference of 1.6 ± 5 dB(A) when taking all measurements into account. After removing measurement sites with noise mitigation measures not considered in the modelling, the difference to the calculation was reduced to 0.5 ± 4 dB(A). As major sources of uncertainty, the position accuracy and topicality of infrastructure and building geometries, the traffic modelling as well as the acoustic source and propagation models were identified.

Keywords: Validation; Noise exposure model; Intermittency Ratio; Long-term measurements; Road traffic noise

*Corresponding Author: Felix Schlatter, J.-M. Wunderli: Empa, Laboratory for Acoustics/Noise control, Dubendorf, Switzerland, E-mail: felix.schlatter@empa.ch
A. Piquerez, M. Foraster: Swiss Tropical and Public Health Institute, Basel, Switzerland
M. Habermacher: n-Sphere AG, Zürich, Switzerland
M. Brink: Federal Office for the Environment, Bern, Switzerland
C. Cajochen: Centre for Chronobiology, Psychiatric Hospital of the University of Basel, Basel, Switzerland
M. S. Ragettli, M. Röösli, N. Probst-Hensch: Swiss Tropical and Public Health Institute, Basel, Switzerland; University of Basel, Basel, Switzerland

1 Introduction

Traffic noise is the dominant source of noise pollution and one of the most widespread sources of environmental stress and discomfort in daily life [5]. Traffic noise is an important public health factor and ranked together with air pollution as the major environmental health risk [14]. In Switzerland for example, currently every fifth inhabitant is heavily disturbed by road traffic noise [11].

Noise induced annoyance as well as somatic health effects are often investigated in epidemiological studies based on large scale noise exposure calculations [19, 33]. Epidemiological studies on noise induced health effects highly depend on accurate noise exposure data. The ascertained dose-effect relationships scale directly to the noise exposure input and thus, systematic deviations in the underlying noise exposure input data lead to biased dose-effect relationships. In addition, large non-systematic deviations have to be compensated with larger datasets. However, large scale emission models are also used for noise monitoring and for strategic action plans of noise mitigation measures [9, 11]. Thereby the accuracy of the exposure calculation is of great importance.

Calculations of large scale noise exposure are computationally expensive and the preparation of input data is laborious. A strategy to reduce calculation time and costs is to use undemanding noise models, to coarsen the resolution of the input data, and to reduce the number of receiver locations. In return, prediction uncertainty is increased [2, 18], which has to be fully compensated by the epidemiologists *i.e.*, with larger population samples, which again entails an increase of study costs [23]. Therefore it can be concluded that an optimized application of means can only be achieved when relying on detailed uncertainty assessment of the noise modelling already at a planning stage of epidemiological studies. Also in the data analysis an uncertainty assessment is indispensable as the modelling results are directly related to the uncertainties of the considered model parameters [29].

The uncertainty of large scale noise modelling can either be assessed by calculations or by comparison with mea-

surements [22, 24, 31]. However there are only few examples in literature, in which a detailed uncertainty evaluation was carried out as part of an epidemiological study. A recent and very detailed example of an uncertainty estimation was performed within the NORAH project [23, 26].. Thereby the uncertainty, given as a standard deviation, was estimated from 3 to 5 dB depending on the traffic type. From 2013 to 2016 a research project called SiRENE (Short and long-term effects of Transportation Noise Exposure) was performed in Switzerland with several results on noise induced health effects, annoyance, and sleep quality already published [6, 10, 12, 13, 15, 16, 28, 34]. In this paper, we present the results of a follow-up validation study of the exposure modelling within SiRENE based on long-term measurements. The validation results are used to assess the accuracy of the noise modelling within SiRENE and to check for unwanted biases in the modelling results. On that basis, conclusions on the primary sources of uncertainty and recommendations for further studies are derived.

2 Method

2.1 Exposure modelling within SiRENE

During the SiRENE study, a nationwide assessment of road, railway, and aircraft noise exposure was conducted for Switzerland for three different years (1991, 2001 and 2011) [20]. Noise was thereby modelled at façade points of all buildings in Switzerland, resulting in a total number of 54,300,00 receiver locations at 1,813,000 buildings. The noise exposure was modelled in an hourly resolution. On that basis, typical noise exposure metrics such as L_{Day}, L_{Night}, $L_{Evening}$, and L_{DEN} were derived. In addition to such noise exposure levels based on the Leq, the concept of the Intermittency Ratio (IR) was applied, a metric characterising the temporal variation of a noise exposure situation [36]. IR was evaluated for each façade point on an hourly basis as well as integrally for day and night-time.

In this subsection, a short overview of the modelling is given for each source type. Information on road traffic modelling is thereby given in more detailedness, as the validation measurements were primarily carried out at receiver locations with dominant road traffic noise.

2.1.1 Road noise modelling

Road noise modelling was performed using the sonROAD emission model [17] and the sound propagation model of StL-86 [2]. The sonRoad emission model predicts the noise emission as a function of vehicle type, driving speed, slope, and road surface properties. StL-86 accounts for geometrical divergence, air absorption, an idealised ground effect, as well as shielding by obstacles. The reason for using a significantly older (and thus simpler, non spectral) propagation model was primary due to computational efficiency. The decision to rely on StL-86 was supported by the fact that the propagation core of StL-86 proofed to be highly reliable in the past, showing no systematic over- or underestimations, and that the model is still recommended by the Swiss Federal Office for the Environment (FOEN) [32].

To calculate the noise emission with the sonROAD emission model, traffic information such as speed, traffic amount and composition is crucial. The Federal Department of the Environment, Transport, and Energy (DE-TEC) maintains a monitoring network and captures detailed annual traffic statistics. However, this monitoring system delivers traffic data for only a very small amount of streets, whereas the noise emission model needs traffic volumes and the driven speed of all roads. Therefore, the traffic model by Arendt [1] was used. This traffic model links population census data with available traffic information about street sections such as street width, number of lanes and road classification to predict the daily amount of traffic. For road sections with no traffic statistics, the driven speed was assigned based on the signalised speed, the vehicle category, and the road classification.

2.1.2 Aircraft noise

Aircraft noise was modelled for Switzerland's largest civil airports Basel, Zurich, Geneva as well as the major military airbase Payerne. Calculations were performed with FLULA2, a time-step model using point sources, which are moved along flight trajectories [8]. For Zurich airport, flight trajectories were considered using radar data for all years. For Geneva, radar data was only available for the year 2011 and 2001 (Footprints from 2000, Rescaled with actual number of movements). For the other years of calculation in Geneva and also for the other airports Basel and Payerne, idealised flight routes were used as input.

2.1.3 Railway Noise

Railway noise was calculated using the sonRAIL emission model [35] and the sound propagation model SEMIBEL [11]. The information about the railway system was obtained from the Federal Office of Transport (FOT). The Swiss Federal Railways delivered data about noise barriers as well as traffic statistics for the years of interest.

2.2 Measurements

Within the SiRENE-project, a socio-acoustic survey was performed to evaluate the degree of annoyance of the Swiss population to transportation noise [6]. This survey was conducted based on a stratified sample of 5592 participants. For our study, we re-contacted the same sample, which resulted in a total number of 102 volunteers, cooperating in our follow-up project. The volunteers were visited at home and a sound level meter was installed at the outer window surface of their living or sleeping room during approximately one week. We used sound level meters type Noise-Sentry RT, a class II measurement device with a measurement uncertainty of about 1 dB(A). The sound level meters were flush mounted to the outer face of the closed window and logged A-weighted 1-s-L_{eq}'s (Equivalent to 1s-L_{AE}). After a measurement period of seven days, the participants removed the sound level meters and sent them back.

To consider reflections of the window surface as well as the window ledge, the logged 1-s Leq values were corrected by -5 dB(A), to represent a measurement in the open window. The open window is the determinant receiver location for noise assessment as defined by the Swiss Noise Abatement Ordinance (NAO) [27]. The correction of -5 dB(A) was derived in [3, 4] based on measurements. The difference of -5 dB(A) instead of -6 dB(A) representing free field conditions is explained by additional reflections from the window frame and from the connected room.

2.3 Data preparation

The participants were instructed to journalise disturbing events other than traffic noise in a questionnaire. Based on this questionnaire, time periods with disturbing sounds were identified and removed. In some cases, participants mentioned that church bells are a dominant noise source. Accordingly, bell tolling around every full hour and addi-

tional events with dominant ringing were removed from the dataset.

To enable a comparison, the measured and corrected 1-s Leq's were energetically averaged over each weekday's hour according to Equation 4. Also on an hourly basis IR were calculated based on the short-term Leq's with 1s resolution. This resulted in 24 1h-Leq and 1h-IR values for every weekday of the 7-day measurement period. Figure 1 shows an example of the resulting diurnal pattern and the variation of levels over the week of measurement.
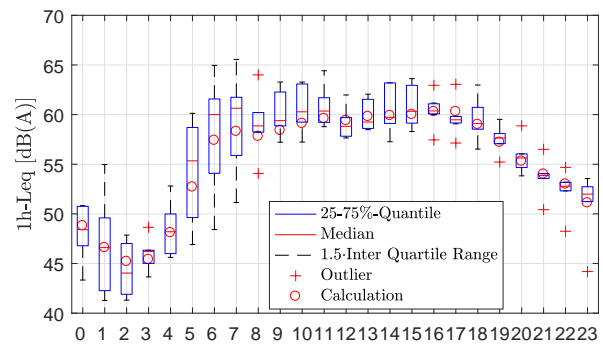

Figure 1: Example showing the diurnal pattern of 1h-Leq's over an entire week for one receiver .

In addition, to remove not journalised disturbing events other than traffic noise, an outlier detection based on Tukey's box plots was applied to the 24 1h-Leq's of the measurement period. 1h-Leq's that lay above the criterion according to Equation 1 were removed.

$$x_{high,h} > Q_{75\%,h} + 1.5 \cdot (Q_{75\%,h} - Q_{25\%,h}) \qquad (1)$$

with:

$Q_{25\%,h}$	25% Quantile of 1h-Leq's of hour h over all weekdays
$Q_{75\%,h}$	75% Quantile of 1h-Leq's of hour h over all weekdays
$x_{high,h}$	Outlier

In a last step, the outlier corrected 1h-Leq's were energetically averaged over all days of the measurement week to get a representative diurnal distribution of sound exposure. Within the SiRENE noise exposure modelling, hourly 1h-Leq levels and IRs were calculated for each façade point. In addition to the hourly values, characteristic noise levels L_{Day} (07 - 19h), $L_{Evening}$ (19 - 23h), L_{Night} (23 - 07h) as well as L_{DEN} were calculated based on the corrected and aggregated weekly 1h-Leq's. L_{DEN} was calculated accord-

ing Equation 2.

$$L_{DEN} = 10 \cdot \log_{10} \left(\frac{12}{24} \cdot 10^{\frac{Leq_{07-19}}{10}} \right.$$
$$\left. + \frac{4}{24} \cdot 10^{\frac{Leq_{19-23}+5}{10}} + \frac{8}{24} \cdot 10^{\frac{Leq_{23-07}+10}{10}} \right) \qquad (2)$$

2.3.1 Intermittency Ratio (IR)

The IR accounts for the acoustical energy produced by single pass-by events above a certain threshold, defined in relation to the total energetic dose during a given time-period. The concept and implications are documented in [36]. The IR is calculated as follows:

$$IR = \frac{10^{0.1 \cdot L_{eq,T,Events}}}{10^{0.1 \cdot L_{eq,T,tot}}} \cdot 100 \qquad (3)$$

Where $L_{eq,T,tot}$ equals the total equivalent continuous sound pressure level energetically averaged over one hour according Equation 4.

$$L_{eq,T,tot} = 10 \cdot \log_{10} \left(\frac{1}{T} \int_0^T 10^{0.1 \cdot L(t)} dt \right) \qquad (4)$$

where $L(t)$ is the measured and corrected continuous A-weighted sound pressure level at the receiver location. In this study, $L(t)$ represents the 1-s A-weighted $L_{A,eq}$-values. Single pass-bys only contribute to the $L_{eq,T,Events}$, if their level exceeds the threshold K, which is defined according to Equation 5 with C set to 3 dB(A).

$$K = L_{eq,T,tot} + C \qquad (5)$$

The acoustical energy of the events ($L_{eq,T,Events}$) is calculated with Equation 6 using the Heavy-side step function H.

$$L_{eq,T,Events} =$$
$$10 \cdot \log_{10} \left(\int_0^T H(L(t) - K) \cdot 10^{0.1 \cdot L(t)} dt \right) \qquad (6)$$

The average $\overline{IR}$ over M time periods k is calculated with Equation 7.

$$\overline{IR} = \frac{\sum\limits_{k=1}^{M} IR^k \cdot 10^{0.1 \cdot L_{eq,T,tot}^k}}{\sum\limits_{k=1}^{M} 10^{0.1 \cdot L_{eq,T,tot}^k}} \qquad (7)$$

Equation 7 is used to derive the averaged 1h-IR of the 7-day measurement period. Figure 2 shows the same dataset as already shown in Figure 1, this time for IR.
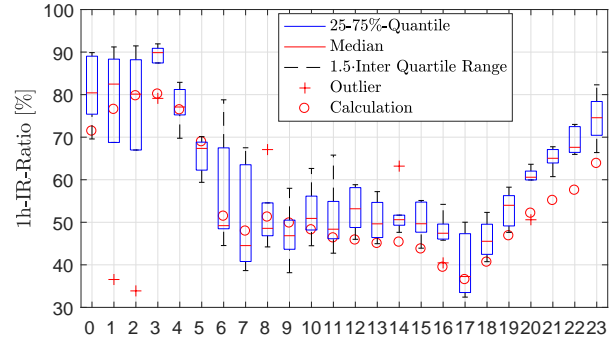

Figure 2: Example showing the diurnal pattern of 1h IRs over an entire week for the same receiver as 1.

2.3.2 Façade points

Within SiRENE, a maximum of three façade points per façade and floor with a minimal distance of 5 m was specified. The façade points were defined for each floor of a building. The number of floors thereby originated from the buildings and dwelling statistics from the Federal Statistical Office (FSO). To compare the calculated with the measured noise levels, each measurement site was manually assigned to a calculated façade point of the SiRENE exposure database, based on the address and the floor information of each participant.

2.3.3 Data cleansing

Firstly, some of the participants had moved to recently built houses and the corresponding measurements were thus not assignable to a façade point from the exposure database. Secondly, in some other cases, the measurements were done on windows leading to inner courtyards, which were not considered in the modelling.

Thirdly, a few measurements had to be removed due to faulty acoustic data.

Fourthly, some sound level meters were unfortunately placed on the window ledge. Those measurement locations are likely to exhibit a reduced exposure to traffic noise sources as the window ledge might shield direct sound propagation. Thus, these measurement locations were omitted in the analysis.

Finally, some of the participants partially closed shutters. During the visit and the measurement installation, the participants habits regarding the use of window shutters was asked for and journalised in the measurement log. Based on the measurement log, times with closed shutters were identified and classified into three categories: Closed

Table 1: Number of measurements and reasons for data cleansing.

	L_{DEN}	$L_{07\text{-}19}$	$L_{19\text{-}23}$	$L_{23\text{-}07}$
Total measurements	**180**	**180**	**180**	**180**
Missing data (No facade point defined, faulty acoustic data, missing measurement protocol …)	20	20	20	20
Noise Sentry flush mounted to window ledge	11	11	11	11
Closed shutters during nighttime	27	0	0	27
Closed shutters during daytime	1	1	1	0
Closed shutters all the time	8	8	8	8
Total removed measurements	67	40	40	66
Correct measurements (Dataset #1)	**113**	**140**	**140**	**114**
Not considered / newly built noise mitigation measures or road covers	10	12	12	10
Correct measurements after removing not considered noise mitigation measures (Dataset #2)	**103**	**128**	**128**	**104**
Strong building influences (winter gardens, loggias, side panels…)	4	5	5	4
Correct measurements after removing not considered noise mitigation measures and building influences (Dataset #3)	**99**	**123**	**123**	**100**

shutters during night-time (23-07), daytime (07-23) or all the time. As closed shutters cause additional shielding and do not represent the modelled situation, these time-periods were removed from the comparison (Dataset #1). The resulting number of measurements and the number of removed measurements are documented in Table 1.

In a second step, the measurement locations were checked for noise mitigation measures not considered in the modelling. It turned out, that the latter was the case at 12 locations. The noise mitigation measures consisted of noise barriers in 10 cases and road covers in two cases. Within this revision step, the measurement locations were removed to form a second comparison dataset, counting 103 (night-time) to 128 (daytime) measurements (Dataset #2). The range in the number of measurements corresponds to different removal of measurements due to closed shutters during the night.

Within SiRENE, the buildings were modelled as cubes. In reality, buildings exhibit different structure elements such as window ledges, side panels, loggias, winter gardens and so on. These structures may, in some cases, interfere substantially with direct sound propagation. Hence, in a third revision step, we removed 4-5 measurement locations, at which strong building influences were present. This was done by reviewing the photos of the measurement locations. With the help of the photos, the orientation of the measured façade point itself and the main noise source was checked and the location was removed, if direct sound propagation was likely to be shielded by structures of the receiving building. After removing these loca-

tions, the comparison was redone with the remaining 99 - 123 measurement locations (Dataset #3).

3 Results

As described in the methods part, the comparison is performed for the three datasets described in Table 1:

- – all data (Dataset #1)
- – all data after removing not considered noise mitigation measures (Dataset #2)
- – all data after removing not considered noise mitigation measures as well as building influences (Dataset #3)

3.1 Average noise levels and IRs

In Table 2, the average differences (μ) and the standard deviations of the differences (σ) are listed for the three datasets and four noise metrics. The datasets refer to the classification presented in Table 1.

Regarding L_{DEN}, the average difference decreases from a difference of 1.6 dB(A) in the dataset #1 to 0.5 dB(A) in Dataset #3. A comparable effect is also observable for the other noise metrics L_{Day}, $L_{Evening}$, and L_{Night}. Not only the average difference decreases from dataset #1 to #3, but also the standard deviation σ is reduced by about 1 dB(A). In Table 3, the resulting differences of the IR is shown for

Table 2: Comparison of the differences (calculation - measurement) of the different datasets for four noise metrics.

	L_{DEN}	L_{Day}	$L_{Evening}$	L_{Night}
		[dB(A)]		
$\mu_{Dataset\#1}$	1.6	2.3	1.7	1.5
$\mu_{Dataset\#2}$	0.8	1.6	0.9	0.5
$\mu_{Dataset\#3}$	0.5	1.5	0.7	0.2
$\sigma_{Dataset\#1}$	5.0	4.8	4.9	5.4
$\sigma_{Dataset\#2}$	4.2	4.2	4.0	4.5
$\sigma_{Dataset\#3}$	4.0	4.1	3.9	4.3

Figure 3: Average differences between calculated - measured 1h-Leq's

two time-periods, day and night (IR_{07-23} and IR_{23-07}) and all three datasets.

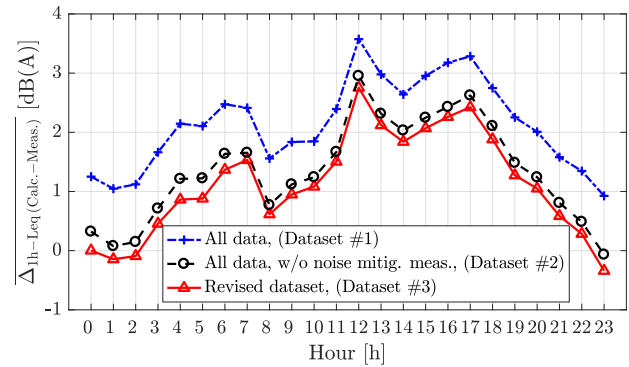**Table 3:** Comparison of calculated - measured IR for the three datasets and two different time-periods.

	IR_{Day}	IR_{Night}
	[%]	[%]
$\mu_{Dataset\#1}$	-7.3	-0.9
$\mu_{Dataset\#2}$	-6.5	-1.0
$\mu_{Dataset\#3}$	-6.7	-1.0
$\sigma_{Dataset\#1}$	16.8	18.0
$\sigma_{Dataset\#2}$	14.8	16.7
$\sigma_{Dataset\#3}$	14.7	17.0

Compared to the noise levels, the IRs hardly react to the revision steps in terms of the average difference. The IR, characterises more the eventfulness than the absolute noise levels, therefore only minor changes in terms of average difference are observed. The IR values are in fact more connected to the source model and scale with the distribution of traffic during a period of time. Therefore the removal of measurement locations has only a minor influence on the average IR.

3.2 Diurnal pattern

3.2.1 1h-Leqs

Figure 3 depicts the averaged differences between the calculation and the measurements for the three datasets. It is clearly visible that the removal of the not considered noise mitigation measures reduces the average differences. But even in the revised datasets the calculation overestimates the measurements on average during daytime, whereas the prediction is rather accurate during the night.

In Figure 4, the standard deviations (σ) of the differences (calculation - measurement) in 1h-Leq are shown. The standard deviation follows an opposite diurnal pattern, compared to the averaged differences presented in Figure 3. During the nighttime, the scattering and thus the standard deviation between calculation and measurement is larger.
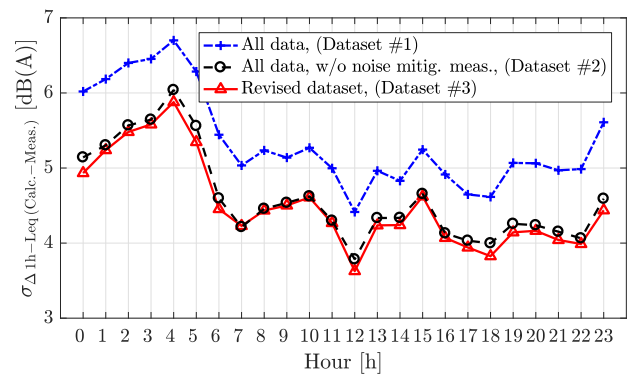

Figure 4: Diurnal standard deviation (σ) pattern of the difference (calculation - measurement) for the three different datasets.

Figure 5 displays absolute 1h-Leq-values of measured vs. calculated validation cases for different hours during the day. In this Figure, the removed measurement locations are displayed as red triangles and the data of the remaining dataset (dataset #3) are presented as blue circles.

The graph shows that the validation covered a wide range of noise levels. It illustrates also, that the calculation follows the 1:1 line over the whole range of levels, without indication of a systematic, level-dependent bias. However, the same pattern of overestimating the noise levels during daytime can be seen again as in Figure 3. The Figure also reflects the current practice of installing noise mitigation measures at highly exposed locations, as demanded by the

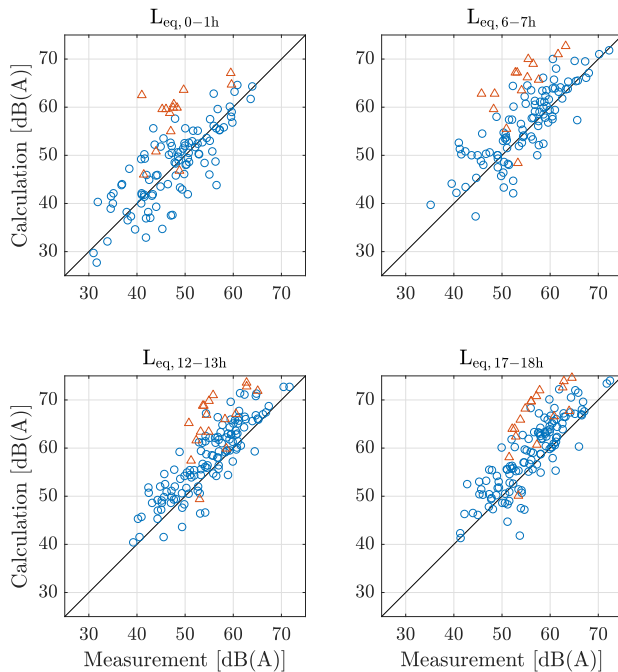Swiss noise abatement ordinance [27]. Therefore, the red triangles appear only at higher noise levels.

Figure 5: Scatterplot of measured vs. calculated 1h-Leq's for different hours. The blue circles represent the noise levels after the revision steps, whereas the red triangles depict the noise levels of removed measurement locations.

3.2.2 1h-IRs

Figure 6 depicts average differences between calculated and measured 1h-IR-ratios for the three datasets. It can be seen, that the calculation strongly overestimates the measured IRs during the night by 5 to 13 %, but only slightly underestimates it during the day. The three datasets show hardly any differences, *i.e.*, the exclusion of situations with not represented barrier effects in the modelling did not influence IR on average.

In Figure 7, the standard deviations of the differences of IR (calculated - measured) are depicted. In contrast to the 1h-Leq, the standard deviations follow a similar diurnal pattern compared to the average difference. A trend towards lower standard deviations from Dataset #1 to #3 is visible. This suggests that the rejection of measurement locations also improves data quality concerning the IR, even though the average deviation as shown in Figure 6 was not affected.
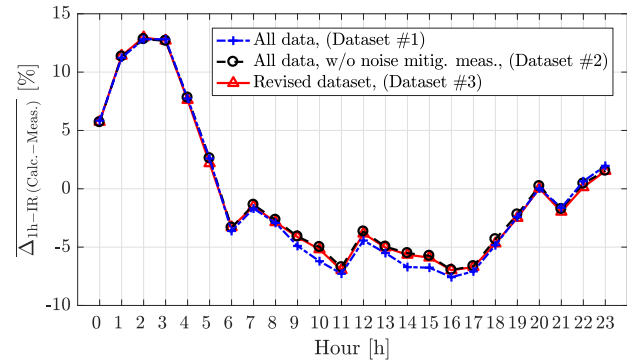

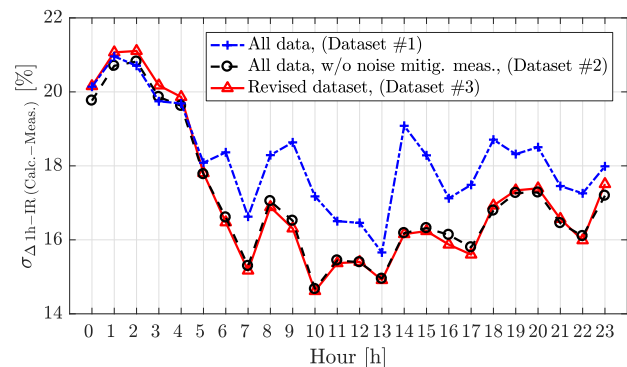**Figure 6:** Average differences (calculation - measurement) of the IR in [%].

Figure 7: Standard deviation (σ) of the difference (calculation - measurement) of the IR in [%].

Figure 8 depicts absolute measured versus calculated IRs at different hours of the day. Removed measurement locations due to not modelled noise mitigation measures and building influences on sound propagation are displayed as red triangles, whereas the blue circles represent the residual dataset #3. Similarly to Figure 5, a wide range of different IRs is covered.

Figure 8 demonstrates a typical characteristic of the IR: during the night-time, the noise exposure is much more event-based than during the day-time. As can be seen from the upper left part of the Figure, the main part of the points lay on the top right corner during the night hours. In addition, scattering is larger. Both can be explained by the fewer number of vehicles, which entails higher IRs.

3.3 Influence of municipality size

The volunteers of this study live in different regions of Switzerland. The different regions and thus measurement locations consequently exhibit different population densities. The population density directly affects the amount of traffic and thus noise levels. As a consequence, noise
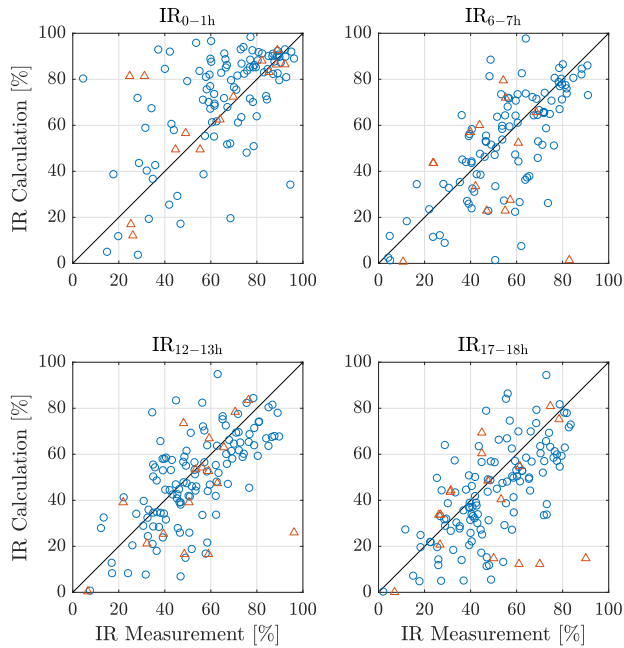
Figure 8: Scatterplot of measured vs. calculated IRs at different hours of the day. The blue circles represent the noise levels after the revision steps, whereas the red triangles depict the noise levels of removed measurement locations.

Figure 9: Hourly averaged differences in Leq for the dataset #3, divided according the community size
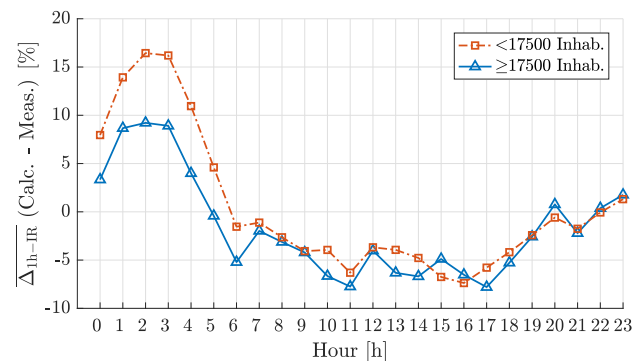
Figure 10: Hourly averaged differences in IR for the dataset #3, divided according the community size

levels should be larger in areas with a higher population density. To investigate this effect, the measurement locations were classified into two categories: rural and urban locations. The classification was performed according to the municipality size by applying a threshold of 17,500 inhabitants. This threshold resulted in an equal number of measurements in each category.

Figure 9 depicts the diurnal pattern of averaged differences in 1h-Leq's for the classified Dataset #3. The graph indicates, that the calculation in more urban areas tends to overestimate the measurements more during the daytime, than the calculations performed in rural areas. During nighttime, the opposite effect can be observed.

Figure 10 depicts the diurnal pattern of the averaged 1h-IR for the classified Dataset #3. In this Figure, the average difference in 1h-IR is larger during night-time at rural locations than at urban locations. During daytime, both differences remain almost at a constant level.

4 Discussion

In this study, a validation of a large scale noise exposure calculation was performed based on long-term measurements. The comparison was conducted with calcu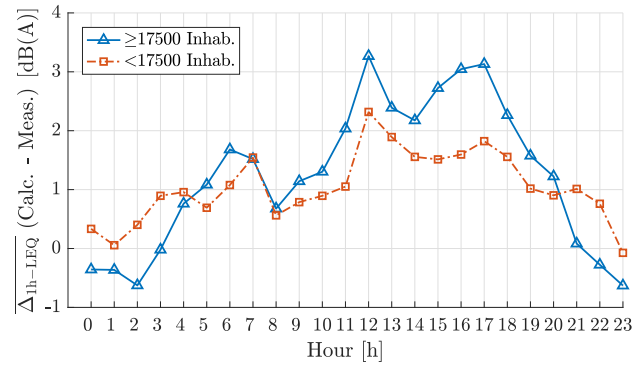lated exposure data of the SiRENE study and measurements performed at 180 receiver positions, each lasting one week. The data analysis was done in an hourly resolution for both, 1h-Leq as well as the 1h-IR.

The comparison of calculated and measured noise exposure levels of the total dataset (Dataset #1) revealed a general overestimation of 1.6 ± 5 dB(A) (L_{DEN}). However, at 12 measurement sites noise mitigation measures not considered in the modelling could be identified that explain most of the systematic overestimation. After the removal of these locations from the dataset, the overestimation was substantially reduced to 0.8 ± 4.2 dB(A) in dataset #2. By additionally checking the measurement locations for specific building structures that interfere with direct sound propagation, another 5 measurement locations were identified and removed in order to form a third dataset (Dataset #3). The comparison of this dataset revealed a deviation from the calculation of 0.5 ± 4.0 dB(A) only, regarding L_{DEN}. Hence it could be shown, that on average no or at least only a small systematic bias can be found, provided the modelling input is correct. As an additional noise metric, the IR, a quantity to evaluate the temporal variation of an exposure situation, was compared. In general, the comparison revealed a good agreement between calculation and

measurement, with a slight overestimation in the calculation of about 7 ± 17% during daytime and -1 ± 18% during night-time.

The resulting difference estimate of our study is of the same magnitude as the estimates from NORAH, which yielded a standard deviation of 3 to 5 dB(A) [23, 26]. This rather high level agreement is surprising as the method of determining uncertainty was fundamentally different (calculation vs. measurement) in the two studies.

4.1 Sources of systematic deviations and scattering

Several sources of systematic deviations and/or scattering were identified:

- Position accuracy and topicality of infrastructure and building geometries
- Traffic modelling
- Acoustic source and propagation modelling
- Noise measurements

4.1.1 Representation of infrastructure and buildings

The systematic deviation in the total dataset #1 can be explained with the quality of the input data for noise modelling: In total 10 to 12 measurement locations, a considerable amount of measurements, had to be excluded during the first revision step to form dataset #2. These measurements were removed because of not considered noise mitigation measures during the modelling. The large amount of such situations may be explained as the measurements were performed in 2016 and the modelling was done for 2011 with partly older input data. In the mean time however, numerous noise mitigation projects have been realised, especially at highly exposed measurement locations. The realisation of these noise mitigation measures is a benefit of Switzerland's Noise Abatement Ordinance [27], which determines the deadlines for noise mitigation projects for highways by 2015 and for other main roads by 2018. Consequently, the systematic overestimation identified in the original dataset #1 is not a weakness of the acoustic modelling but a problem of the topicality of the input data, respectively, the time difference between modelling and measurement.

To form dataset #3, the measurement locations were checked for structures like balconies, loggias, or side panels, that interfere substantially with direct sound propagation.

The installation of such structures in buildings is among others done to reduce the noise exposure of its inhabitants. Therefore they should be taken into account in the modelling.

4.1.2 Traffic modelling and acoustic emission modelling

In all datasets, a slight trend to overestimate the measured 1h-Leq values was found. One potential reason for the higher calculated than measured 1h-Leq values is the road traffic emission modelling. Within sonROAD, the driven velocities, the composition, and the amount of traffic are crucial. The calculation was done with time dependent traffic volumes, but with time independent traffic speed, based on signalled speed limits. The actual driven speed therefore was likely to be lower than assumed. It can often be noticed that in dense traffic the actually driven velocities are lower than the signalled speed limits. Dense traffic is much more probable for urban situations during daytime than for rural areas or night-time. This hypothesis is supported by the diurnal pattern and by the additional analysis done in Section 3.3, in which the influence of the municipality size was studied. In addition, the assumption of the source model of a constant travelling speed is not always correct. Particularly in urban areas, acceleration and breaking zones are very common, leading at least to an additional uncertainty of the modelling, as could also be shown in Section 3.3.

Further, the emission model of sonROAD does not account for temperature effects in the noise emission generation. It has been shown that the generation of rolling noise of tyre/surface is temperature dependent with gradients ranging from -0.03 to -0.09 dB/°C depending on the type of tyre and the driving speed [7, 25]. In Switzerland the day-night temperature difference is typically in the range of 10-15°C. Consequently a substantial part of the noticed day-night-difference might be attributed to the temperature dependency of the rolling noise.

Finally, road surface properties have a major impact on the road noise generation [30]. The road noise emission model sonROAD accounts for different road surfaces in terms of sound level corrections [17]. However, highly detailed, geo-referenced road surface properties were not available and therefore standard values were used for the calculation. This simplification in the modelling is likely to cause at least an increase of the standard deviation.

4.1.3 Noise measurement

We compared modelling calculations with measurements, which were conducted during spring and autumn of the year 2016. The comparison was done with calculations representing an annual average. As a consequence, the measurement represent only the noise emission within a specific moment of time at given temperatures and humidity. As we could not compare measurements distributed equally over the whole year, differences due to weather conditions as well as traffic distribution could not be evaluated.

Regarding the distribution of the 1h-IR, the calculated IRs tend to overestimate the measured values during the night, while during daytime even a small underestimation occurs. The overestimation of the calculated IRs during the night-time can mostly be explained by the measurements and cannot be attributed to the calculation method, mainly because of two reasons. Firstly the sensitivity of the class II measurement devices is limited and the noise floor is comparably high with about 31 dB(A). During quiet nights, the outside sound pressure levels are likely to fall below this threshold, resulting in a significantly reduced fluctuation of levels in the measurements, compared to reality - and compared to the modelling, which of course does not account for limitations of measurement devices. Secondly, in situations with very small traffic volumes, which are typical for core night hours, background noise is not dominated by traffic noise any more, but by other sources which are not represented in the modelling. The effect of such additional background noise is also visible in Leq$_{Night}$. IR however reacts much more sensitive to changes in background noise, which is why this influence is more prominent for the IR.

Furthermore also measurements themselves exhibit a considerable uncertainty. This has to be taken into account by comparing calculations with measurements. In particular, the placement of the sound level meters, the choice of devices and their calibration has to be evaluated with great care in order to retrieve reliable measurement data.

4.2 Limitations

The exposure modelling within SiRENE had been performed for the year 2011. This was in accordance with the most recent data sampling of the epidemiological dataset SAPALDIA, which also dated from that year and was used in the SiRENE study [10, 12, 13]. The validation measurements however were only performed in 2016. This leads to a difference of five years between measurements

and calculations and consequently changes in infrastructure (new roads, noise barriers and buildings) as well as changes in traffic data (traffic volume and speed) are not represented in the modelling. However, an update of the noise calculation was not foreseen in the scope of the project, and a rescaling of the calculated noise levels from 2011 seemed inappropriate.

In conclusion, we think, that the comparing measurements should have been performed in 2011. As a result, most of the differences due to missing noise mitigation in the model could have been avoided.

5 Conclusion and Outlook

With this validation study, we could show that large scale exposure modelling within the SiRENE study resulted on average in good agreement with measurements. Tthe exposure modelling does not exhibit a substantial systematic over- or underestimation, neither for situations with high nor low exposure levels. However, scattering between calculated and measured sound pressure levels as well as for the IRs are considerable. Consequently, the main goal of future studies would be to reduce scattering. Based on the present study, the following recommendations can be made for future investigations:

- The remaining differences and the scattering of the 1h-Leq are likely to be reduced by using improved traffic flow models. The model should especially include time dependent and locally varying traveling speeds for different vehicle classes as a function of traffic density and consider acceleration and breaking zones.
- With improved source data also source models should be used, which are able to take the above refinements into account. With the recently developed road traffic model from CNOSSOS [21], a model is available which considers unsteady travelling speeds as well as temperature influences on rolling noise.
- Great attention has to be given to the quality of the geo-referenced input-data. In addition, the input data has to be fully up to date and has to exhibit a high position accuracy.
- With higher accuracy of the input data, it is also recommended to use a detailed sound propagation model to consider, among others, reflections, which are likely to be important, especially in urban areas.
- A general weakness of both, input and sound propagation modelling, is that the influences of the re-

ceiving building itself are usually ignored. The calculation yields free field levels and does not account for interactions with the building of the inhabitant. However, it is current practice at residential buildings in highly exposed areas to optimize building structures to maximise noise protection. This again supports the use of detailed sound propagation models, which account for multiple barrier effects as well as reflections. Consequently, the modelling of sound propagation in this complexity requires also input data at a particularly high degree of detail and accuracy.

The usage of large scale noise models for feeding epidemiological models, performing noise monitoring, or the development of strategic action plans is a well established practice. In addition, a proper evaluation of the uncertainties or unwanted biases has to be performed either, by measurements or by calculation. Our evaluation strongly supports this attitude for other studies. With our study, we could show, that the model predictions are correct in the average, provided that accurate and freshest input data are used. A general problem is, that the environment is in a steady change and modelling permanently lags behind. As a consequence, large scale noise calculations are only reliable until substantial changes in infrastructure arise. As a consequence, inputs for epidemiological surveys have to be sampled within the valid timeframe of the noise model.

References

[1] Arendt. Verkehrsaufkommen flächendeckend ermitteln. http://arendt.ch/objects/websitedownload/53/download/VM_Flyer_d.pdf, 2009.

[2] BAFU. Computermodell zur Berechnung von Strassenlaerm, Teil 1, Bedienungsanleitung zum Computerprogramm StL-86. Technical report, Bundesamt für Umwelt, Schweiz, 1987.

[3] BAFU. Methode zur Ermittlung der Aussenlärm-Immissionen bei geschlossenem Fenster. Vollzug Umwelt, Mitteilung zur Lärmschutz-Verordnung (LSV) Nr. 7 (1995). Report, Bundesamt für Umwelt (BAFU), 1995. PR.

[4] BAFU. Methode zur Ermittlung der Aussenlärm-Immissionen bei geschlossenem Fenster. Vollzugshilfe zur Lärmschutzverordnung (LSV). Aktualisierung der Mitteilung zur LSV Nr. 7 von 1995. Bundesamt für Umwelt, Bern, Umwelt-Vollzug Nr. 1502:3 S, 2015. "https://www.bafu.admin.ch/dam/bafu/de/dokumente/laerm/uv-umwelt-vollzug/methode_zur_ermittlungderaussenlaerm-immissionenbeigeschlossenem.pdf.download.pdf/methode_zur_ermittlungderaussenlaerm-immissionenbeigeschlossenem.pdf".

[5] Mathias Basner, Wolfgang Babisch, Adrian Davis, Mark Brink, Charlotte Clark, Sabine Janssen, and Stephen Stansfeld. Audi-

tory and non-auditory effects of noise on health. The Lancet, 383(9925):1325–1332, 2014.

[6] Mark Brink, Reto Pieren, Maria Foraster, Danielle Vienneau, Ikenna Eze, E. Schaffner, and Jean Marc Wunderli. Do short-term temporal variations of noise exposure explaine of noise annoyance? In Proceedings of the Internoise 2016, Hamburg, pages 5720–5730, 2016.

[7] Erik Bühlmann, Ulf Sandberg, and Piotr Mioduszewski. Speed dependency of temperature effects on road traffic noise. In Proceedings of Inter-Noise 2015, San Francisco, 2015.

[8] Empa. FLULA2, a Method for the Calculation and Illustration of Aircraft Noise Exposure. Technical Program Documentation. Version 4. Technical report, Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory of Acoustics / Noise Control, April 2005 2010. https://www.empa.ch/documents/56129/103151/SaT_FLULA2_Dokument/62e3c7e1-e395-4975-9eba-fda3adf17962, Last Viewed: 27.10.2017.

[9] EU. Directive 2002/49/ec of the european parliament and of the council of 25 june 2002 relating to the assessment and management of environmental noise. Report, Official Journal of the European Communities, 2002.

[10] Ikenna C. Eze, Maria Foraster, Emmanuel Schaffner, Danielle Vienneau, Harris Héritier, Franziska Rudzik, Laurie Thiesse, Reto Pieren, Medea Imboden, Arnold von Eckardstein, Christian Schindler, Mark Brink, Christian Cajochen, Jean-Marc Wunderli, Martin Röösli, and Nicole Probst-Hensch. Long-term exposure to transportation noise and air pollution in relation to incident diabetes in the sapaldia study. International Journal of Epidemiology, 2017.

[11] FOEN. SonBase – The GIS Noise Database of Switzerland., 2009.

[12] M. Foraster, I. Eze, E. Schaffner, D. Vienneau, H. Heritier, S. Endes, F. Rudzik, L. Thiesse, R. Pieren, C. Schindler, A. Schmidt-Trucksäss, M. Brink, C. Cajochen, J. M. Wunderli, M. Röösli, and N. Probst-Hensch. Exposure to road, railway, and aircraft noise and arterial stiffness in the sapaldia study: Annual average noise levels and temporal noise characteristics. Environ Health Perspect, 125(9), 2017.

[13] Maria Foraster, Ikenna C. Eze, Danielle Vienneau, Mark Brink, Christian Cajochen, Seraina Caviezel, Harris Héritier, Emmanuel Schaffner, Christian Schindler, Miriam Wanner, Jean-Marc Wunderli, Martin Röösli, and Nicole Probst-Hensch. Long-term transportation noise annoyance is associated with subsequent lower levels of physical activity. Environment International, 91:341–349, 2016.

[14] O. Hanninen, A. B. Knol, M. Jantunen, T. A. Lim, A. Conrad, M. Rappolder, P. Carrer, A. C. Fanetti, R. Kim, J. Buekers, R. Torfs, I. Iavarone, T. Classen, C. Hornberg, and O. C. Mekel. Environmental burden of disease in europe: assessing nine risk factors in six countries. Environ Health Perspect, 122(5):439–446, 2014.

[15] Harris Héritier, Danielle Vienneau, Maria Foraster, Ikenna Collins Eze, Emmanuel Schaffner, Laurie Thiesse, Franziska Rudzik, Manuel Habermacher, Micha Köpfli, Reto Pieren, Mark Brink, Christian Cajochen, Jean Marc Wunderli, Nicole Probst-Hensch, and Martin Röösli. Transportation noise exposure and cardiovascular mortality: a nationwide cohort study from switzerland. European Journal of Epidemiology, pages 1–9, 2017.

[16] Harris Héritier, Danielle Vienneau, Patrizia Frei, Ikenna C. Eze, Mark Brink, Nicole Probst-Hensch, and Martin Röösli. The As-

sociation between Road Traffic Noise Exposure, Annoyance and Health-Related Quality of Life (HRQOL). International Journal of Environmental Research and Public Health, 11(12):12652–12667, 2014.

[17] K. Heutschi. SonRoad: New Swiss road traffic noise model. Acta Acustica United with Acustica, 90(3):548–554, 2004.

[18] C. Iglesias Merchan and L. Diaz-Balteiro. Noise pollution mapping approach and accuracy on landscape scales. Sci Total Environ, 449:115–25, 2013.

[19] L. Jarup, W. Babisch, D. Houthuijs, G. Pershagen, K. Katsouyanni, E. Cadum, M. L. Dudley, P. Savigny, I. Seiffert, W. Swart, O. Breugelmans, G. Bluhm, J. Selander, A. Haralabidis, K. Dimakopoujou, P. Sourtzi, M. Velonakis, F. Vigna-Taglianti, and HYENA Study Team. Hypertension and exposure to noise near airports: the hyena study. Environmental Health Perspectives, 116(3):329–333, 2008.

[20] I. Karipidis, D. Vienneau, M. Habermacher, M. Köpfli, M. Brink, N. Probst-Hensch, M. Roosli, and J. M. Wunderli. Reconstruction of historical noise exposure data for environmental epidemiology in switzerland within the sirene project. Noise Mapping Journal, 1:3–14, 2014.

[21] Stylianos Kephalopoulos, Marco Paviotti, and Fabienne Anfosso Ledee. Common noise assessment methods in Europe (CNOSSOS-EU). Common noise assessment methods in Europe (CNOSSOS-EU), pages 180–p, 2012.

[22] G. Licitra, L. Fredianelli, D. Petri, and M. A. Vigotti. Annoyance evaluation due to overall railway noise and vibration in Pisa urban areas. Sci Total Environ, 568:1315–25, 2016.

[23] M. Liepert, M. Muelbacher, U. Moehler, G. Thomann, and D. Schreckenberg. Uncertainty of calculated noise levels and its influence on exposure- response-relationship in the NORAH-project. Proceedings of ICBEN Conference, Zurich, Switzerland, 2017.

[24] R. Makarewicz and R. Golebiewski. Uncertainty of traffic noise prediction. ACTA ACUSTICA UNITED WITH ACUSTICA, 92:578–582, 2006.

[25] P. Mioduszewski, S. Taryma, and R. Wozniak. Temperature influence on tyre/road noise frequency spectra. In Proceedings of the Internoise 2016, Hamburg, pages 4924–4934, 2016.

[26] U. Möhler, M. Liepert, M. Mülbacher, M. Klatte, B. Vogelsang, and G. Thomann. The acoustic basis of the NORAH field studies. In Proceedings of the Internoise 2016, Hamburg, pages 7764–7767, 2016.

[27] NAO. Lärmschutz-Verordnung (LSV) vom 15. Dezember 1986 (Stand am 1. August 2010) (Noise Abatement Ordinance (NAO) from 15 December 1986 (Effective 1 August 2010)). SR 814.41, 1986.

[28] F. Rudzik, L. Thiesse, R. Pieren, J. M. Wunderli, M. Brink, N. Probst-Hensch, M. Röösli, and C. Cajochen. Effects of continuous and intermittent transportation noise on sleep fragmentation. In Proceedings of the Internoise 2016, Hamburg, pages 7726–7732, 2016.

[29] Alessandro Ruggiero. The importance of uncertainty in the noise pollution measurements. Environment Pollution and Climate Change, 01(01), 2017.

[30] Ulf Sandberg. Road traffic noise—the influence of the road surface and its characterization. Applied Acoustics, 21(2):97–118, 1987.

[31] Beat Schäffer, Stefan Plüss, and Georg Thomann. Estimating the model-specific uncertainty of aircraft noise calculations. Applied Acoustics, 84(0):58–72, 2014.

[32] G. Schguanin and T. Ziegler. Leitfaden Strassenlärm. Vollzugshilfe für die Sanierung. Stand: Dezember 2006. Umwelt-Vollzug Nr. 0637. Bundesamt für Umwelt, Bern. 46 S. Government document, 2006.

[33] Jenny Selander, Gösta Bluhm, Töres Theorell, Göran Pershagen, Wolfgang Babisch, Ingeburg Seiffert, Danny Houthuijs, Oscar Breugelmans, Federica Vigna-Taglianti, Maria Chiara Antoniotti, Emmanuel Velonakis, Elli Davou, Marie-Louise Dudley, and Lars Järup. Saliva cortisol and exposure to aircraft noise in six european countries. Environ Health Perspect, 117(11):1713–1717, 2009.

[34] L. Thiesse, F. Rudzik, R. Pieren, J-M. Wunderli, K. Spiegel, M. Brink, N. Probs-Hensch, Röösli M., and C. Cajochen. Short-term effects of nocturnal transportation noise on glucose metabolism. In Proceedings of the Internoise 2016, Hamburg, pages 7733–7738, 2016.

[35] T. Thron and M. Hecht. The sonRAIL emission model for railway noise in Switzerland. Acta Acustica United with Acustica, 96:873–883, 2010.

[36] Jean Marc Wunderli, Reto Pieren, Manuel Habermacher, Danielle Vienneau, Christian Cajochen, Nicole Probst-Hensch, Martin Roosli, and Mark Brink. Intermittency ratio: A metric reflecting short-term temporal variations of transportation noise exposure. J Expos Sci Environ Epidemiol, 2015.