DOI: 10.61782/fa.2025.0355

FORUM ACUSTICUM EURONOISE 2025

EVALUATING BACKGROUND NOISE CRITERIA FOR ESTIMATING OUTDOOR-TO-INDOOR LEVELS VIA A STATISTICAL LEARNING MODELLING FRAMEWORK

Michail Evangelos Terzakis[1*] **Cédric Van hoorickx**[1] **Maarten Hornikx**[1]
Faridah Naim[2] **Daniela Fecht**[3] **Anna L. Hansell**[4,5] **John Gulliver**[6]

[1] Building Acoustics Group, Department of the Built Environment, Eindhoven University of Technology, Eindhoven, The Netherlands
[2] School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Malaysia
[3] School of Public Health, Imperial College London, London, The United Kingdom
[4] The National Institute for Health Research (NIHR), Leicester Biomedical Research Centre (BRC), Leicester General Hospital, Leicester, The United Kingdom
[5] The National Institute of Health Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health, University of Leicester, Leicester, The United Kingdom
[6] Population Health Research Institute, City St George's, University of London, London, The United Kingdom

ABSTRACT

Urban noise exposure has shown significant negative associations with health effects, largely based on outdoor assessments. As most people spend the majority of time indoors, especially at night, assessments should include the attenuation of outdoor noise exposure to estimate indoor noise exposure for developing epidemiological evidence on health-related noise effects. In this study, a prediction modelling framework is defined for estimating the outdoor-to-indoor level differences for three time periods (day, evening, and night) via statistical learning models, employing synchronised and unsupervised outdoor and indoor noise exposure measurements from 49 locations in Greater London, UK. The study evaluates which background noise criteria expressed via percentile indicators (L_{A90}, L_{A95}, and L_{A99}) under two conditions are most effective. In the analysis, stronger correlations between the equivalent outdoor-to-indoor level differences and outdoor levels were obtained after applying the more stringent condition. Particularly, stronger night-time relationships (r = 0.66–0.76) were obtained compared to day-time (r = 0.55–0.63) and evening (r = 0.43–0.55) relationships. Regarding the modelling validation, the lowest RMSE was observed in the night-time models (RMSE = 2.0–2.8dB(A)), followed by the day-time (RMSE = 2.9–3.0dB(A)) and evening (RMSE = 2.5–4.0dB(A)) models. The study's findings support the value of random forest models in estimating outdoor-to-indoor level differences, executing errors mostly below the just perceptible level difference (3dB(A)).

Keywords: *outdoor-to-indoor, noise exposure, statistical learning, attenuation.*

1. INTRODUCTION

The impact of noise exposure on human health has shown significant negative associations with a range of non-auditory health effects effects, including annoyance, sleep disturbance, reduced quality of life, and cardio-metabolic outcomes [1, 2]. The consequences of chronic noise exposure in childhood may manifest in adverse outcomes in later life [2]. Furthermore, the impact of noise exposure,

**Corresponding author*: m.e.terzakis@tue.nl.

associated with the indirect accumulation of measurement of noise exposure in human organisms [3], underscores the importance of expressing noise exposure via objective noise indicators for exploring exposure-response relationships [2].

In urban noise exposure studies, outdoor noise exposure levels are usually employed as a basis for analyses in contrast to the utilisation of indoor levels [2]. This suggests the potential for a systematic risk of bias in the identification of exposure-response relationships. This is particularly salient in relation to vulnerable groups whose activities are predominantly confined to indoor environments. Hence, statistical prediction models have been developed with the purpose of estimating outdoor-to-indoor level differences based on outdoor levels as well as parameters related to outdoor environment, indoor environment, and building façade characteristics. This has been done via both supervised [4] and unsupervised [5] measurement approaches. These models allow for a more precise estimation of attenuation levels in comparison to a single-valued attenuation factor (e.g., [4–7]).

Focusing on the incorporation of unsupervised noise measurements in the estimation of outdoor-to-indoor level differences, a recent methodology has been proposed [5]. In this methodology, the L_{A99} percentile indicator has been applied for the estimation of background noise levels, while two defined background noise conditions (6dB(A) and 10dB(A)) verify the incorporation of the proper indoor (and their respective outdoor) noise levels.

The objective of this study was to establish attenuation levels for use in the adjustment of outdoor noise levels to facilitate the prediction of indoor noise exposure levels. First, the impact of indoor background noise levels on the relationship between the outdoor-to-indoor level differences and outdoor levels was investigated via experimental data from 49 dwellings in Greater London, UK. Second, statistical learning models (random forest), were developed for predicting the outdoor-to-indoor attenuation levels with respect to outdoor levels and predictors related to outdoor and indoor environment as well as façde characteristics.

2. METHODOLOGY

In this section, the principal methodologies are presented, corresponding to the dataset processing and the prediction modelling. For a comprehensive overview of the study design, including the detailed procedure of the noise exposure assessment, the reader is directed to [8].

2.1 Data Processing

Following the capturing of outdoor and indoor noise exposure levels [8], the A-weighted equivalent noise indicator was used for the assessment of outdoor ($L_{\mathrm{Aeq,Out}}$), indoor ($L_{\mathrm{Aeq,In}}$), and outdoor-to-indoor noise exposure levels ($L_{\mathrm{Aeq,Out\text{-}In}}$) for day-time (07:00-19:00), evening (19:00-23:00), and night-time (23:00−07:00) period based on an integration time of 1-minute and a time-period of 1-hour.

Since noise exposure measurements were conducted via an unsupervised approach, a data processing procedure was employed for capturing the representative outdoor-to-indoor levels. Particularly, noise exposure (i.e. full bandwidth) measurements, where indoor levels exceeded outdoor levels, were excluded. Then, the spectral characteristics of both outdoor and indoor exposure levels were evaluated in 1/3-octave bands (50Hz–5kHz). Consequently, measurements for which the indoor spectral levels were determined to exceed those of the outdoors were excluded. In order to reduce the impact of indoor background noise on the identification of the appropriate outdoor-to-indoor level differences, some percentile noise indicators were selected to express the background noise levels. These correspond to $L_{\mathrm{A99,In}}$, $L_{\mathrm{A95,In}}$, and $L_{\mathrm{A90,In}}$. Finally, two background noise conditions (6dB(A) and 10dB(A)) were defined associated to the minimum level difference between the equivalent indoor levels and the background noise levels (i.e., $L_{\mathrm{Aeq,In}} - L_{\mathrm{AN,In}}$). Differences lower than the two conditional levels indicate a significant effect of the indoor background noise, and therefore they were excluded.

2.2 Prediction Modeling

Initially, the relationship between outdoor levels and outdoor-to-indoor level differences per time-period (day, evening, and night), background noise indicator (L_{A90}, L_{A95}, and L_{A99}) and condition (6dB(A) and 10dB(A)) was verified via a simple regression model, using the R^2 metric and Pearson correlation coefficient (r). For explaining the percentage of unexplained variance in the estimation of the outdoor-to-indoor level difference, random forest (RF) models were developed. Numerical predictors associated to the outdoor noise levels ($L_{\mathrm{Aeq,Out}}$), the room volume (RV), the window size (WS) and the window numbers (WN), as well as to the categorical variables associated to the window type (WT), the room type (RT), the building type (BT), and the site type (ST) were applied. In all the models, the datasets were split into 70% for training and 30% for testing. The random forest models were ap-

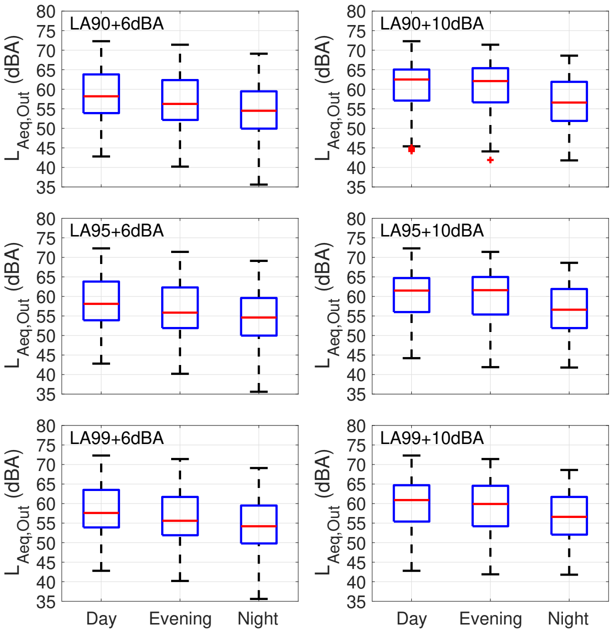

Figure 1. Outdoor levels per time period, background noise indicator and condition.

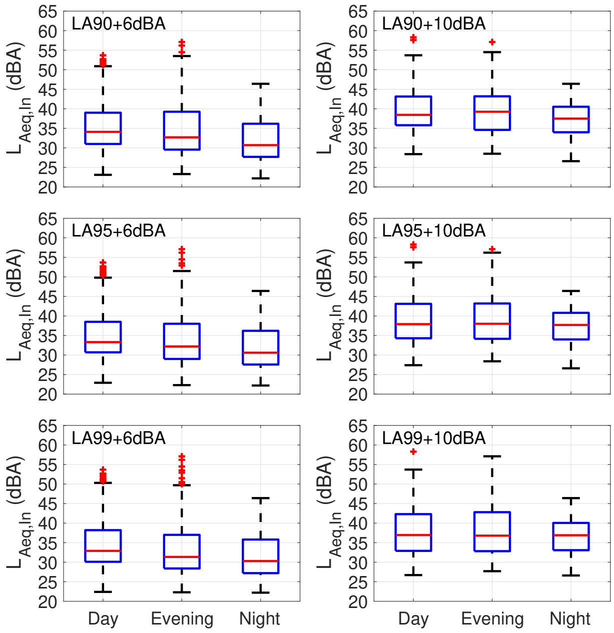

Figure 2. Indoor levels per time period, background noise indicator and condition.

plied via the *ranger* package [9] in RStudio software. All the models were tuned via three (hyper)parameters: the number of trees (n_{tr}), the number of variables (n_{v}) used at each split (m_{tr}), and the complexity of trees (c_{t}). The number of trees was chosen to range from 70 to 1200 trees with a step of 10 trees. The number of split-variables used to tune the models ranged from 2 to 8 variables with a step of 1 variable. Regarding the complexity of tree parameter (tree nodes), the number of nodes ranged from 1 node to 15 nodes with a step of 1 node. Moreover, the (normalized) importance of each predictor per model was evaluated. Finally, the error between predicted and observed $L_{\text{Aeq,Out-In}}$ levels was verified via the root-mean-squared error (RMSE) metric [10].

3. RESULTS

3.1 Assessment of Noise Exposure Levels

First, the outdoor, indoor, and outdoor-to-indoor noise exposure levels were assessed per time period, background noise indicator and condition, and the results are presented in Fig. 1, Fig. 2 and Fig. 3, respectively.

As seen from Fig. 1 and Fig. 2, the median outdoor and indoor exposure levels based on the 6dB(A) background noise condition are lower than the corresponding

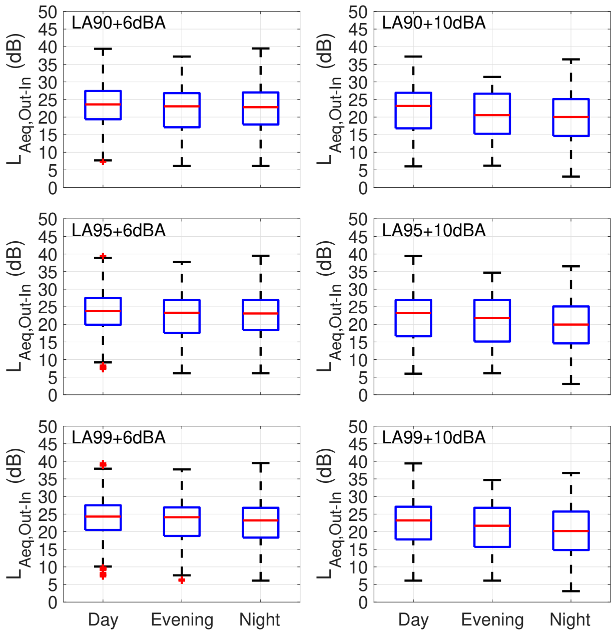

Figure 3. Outdoor-to-indoor levels per time period, background noise indicator and condition.

levels based on the 10dB(A) background noise condition for all time periods and background noise indicators. Fur-

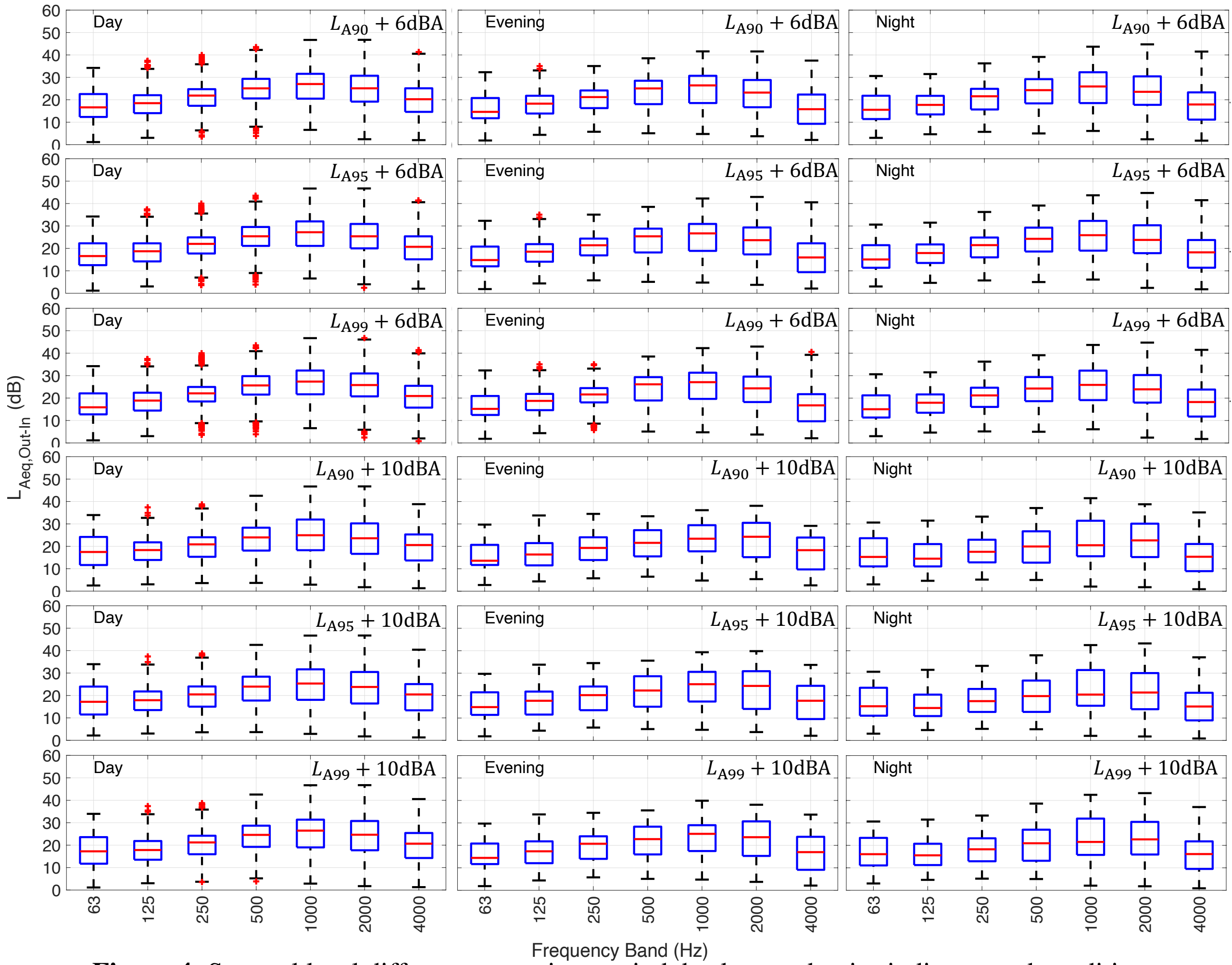

Figure 4. Spectral level differences per time period, background noise indicator and condition.

thermore, the median day-time outdoor and indoor levels are generally higher than evening and night-time levels across the background noise conditions and indicators.

Opposite to Fig. 1 and Fig. 2, the median outdoor-to-indoor levels in Fig. 3 presented a downward trend by increasing the background noise condition across the background noise indicators and exposure periods. The decline in outdoor-to-indoor levels in the transition to stringent background noise conditions is associated with the higher transmission of energy affected less by the indoor background noise. Consequently, the stringent background noise condition is responsible for the presence of broader IQR (interquartile range) across the time periods and background noise indicators.

Focusing on the median average spectral outdoor-to-indoor levels per 1/1-octave bands in Fig. 4, it can be seen that levels based on the 10dB(A) condition present less outliers compared to the levels based on the 6dB(A) condition. Furthermore, the stringent background condition presents broader IQR at most frequency bands across the time periods and background noise indicators.

3.2 Relationship between $L_{\mathrm{Aeq,Out}}$ and $L_{\mathrm{Aeq,Out\text{-}In}}$

For exploring the relationship between outdoor and outdoor-to-indoor levels, simple regression models per background noise indicator and condition were executed. In Fig. 5, the scatter plots between the $L_{\mathrm{Aeq,Out}}$ and $L_{\mathrm{Aeq,Out\text{-}In}}$ are presented, while in Tab. 1, the R^2 and Pearson correlation coefficient (r) are summarized.

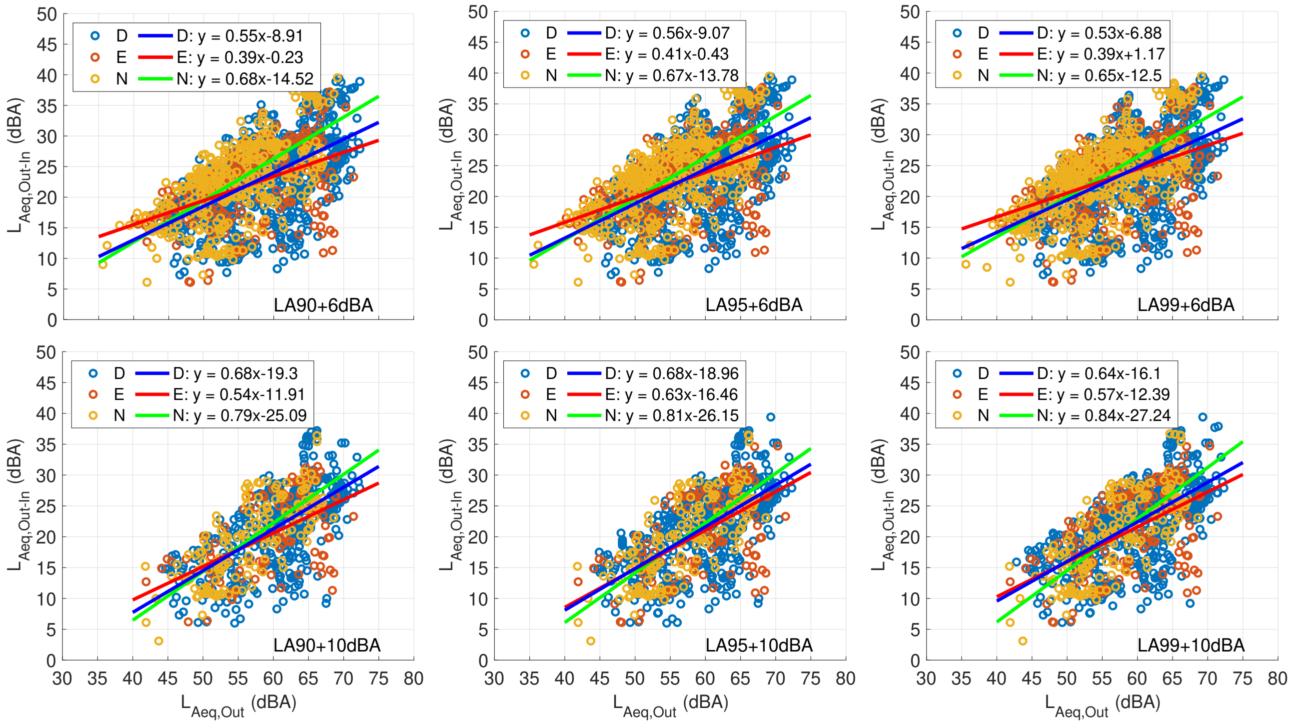

Figure 5. Relationship between outdoor and outdoor-to-indoor levels per time period and setting.

Table 1. Diagnostics per explored relationship.

Model	BGN: 6dB(A)		BGN: 10dB(A)	
	R^2	r	R^2	r
$L_{\text{Aeq,LA90,D}}$	0.34	0.58	0.40	0.63
$L_{\text{Aeq,LA95,D}}$	0.34	0.58	0.39	0.62
$L_{\text{Aeq,LA99,D}}$	0.31	0.55	0.39	0.62
$L_{\text{Aeq,LA90,E}}$	0.20	0.44	0.27	0.52
$L_{\text{Aeq,LA95,E}}$	0.20	0.45	0.30	0.55
$L_{\text{Aeq,LA99,E}}$	0.19	0.43	0.27	0.52
$L_{\text{Aeq,LA90,N}}$	0.44	0.66	0.57	0.76
$L_{\text{Aeq,LA95,N}}$	0.44	0.66	0.57	0.75
$L_{\text{Aeq,LA99,N}}$	0.44	0.66	0.56	0.75

Green cells: Moderate, Blue cells: Strong

As seen from Fig. 5 and focusing on Tab. 1, an increase in background noise from 6dB(A) to 10dB(A) resulted in stronger correlations, indicating the inclusion of less affected indoor background noise levels. In particular, the correlation coefficient exhibited an increase from moderate to strong for day-time exposure. For evening exposure, the correlation coefficient increased, yet it maintained a moderate correlation degree. Regarding night-time exposure, the correlation coefficient exhibited an increase, and therefore it continued to demonstrate a strong degree. As the indoor background levels exhibited less significant variations at night-time stronger correlations were obtained in comparison to the day-time and evening time periods. By comparing r and R^2 with respect to the type of background noise indicator, it can be seen that there were no significant differences among them across the time periods and background noise conditions. Hence, it can be suggested that for a single response relationship, the background noise indicator is less important than the background condition.

3.3 Prediction Modelling

Following the tuned RF models, the importance of each predictor in the estimated response ($L_{\text{Aeq,Out-In}}$) was identified via the normalized importance and the results are presented in Fig. 6.

As seen from Fig. 6, the $L_{\text{Aeq,Out}}$ emerges as the most important predictor across all models, explaining the majority of the variance in the estimation of the outdoor-to-indoor levels. For the 6dB(A)-based models, the RV, WS, and BT predictors mostly accounted for more than 40% of the variance across noise conditions and time periods. In contrast, the ST and WT predictors for the day-time and

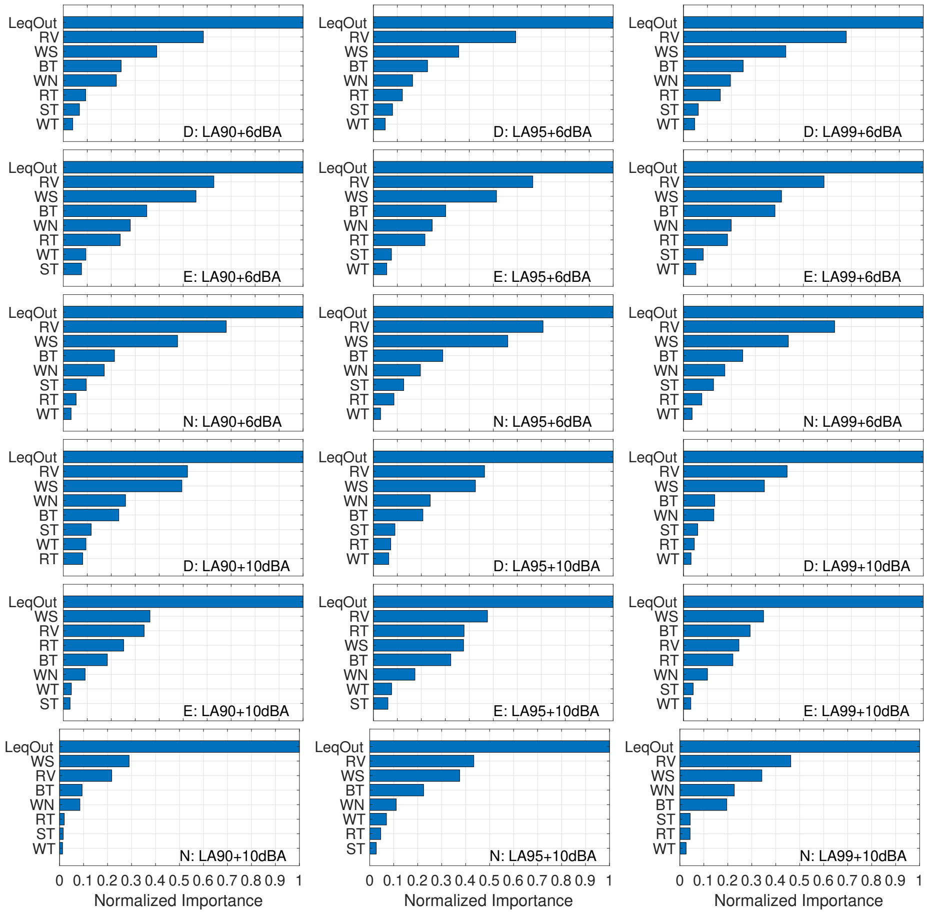

Figure 6. Importance of predictors in models per time period, background noise indicator, and condition.

evening models as well as the RT and WT predictors for the night-time models explained less than the 10% of variance. In the 10dB(A) models, the RV and WS predictors accounted for most of 40% of variance for the day-time models. For evening models, the WS and RV predictors in the L_{A90} model, the WS, RV, BT, and RT predictors in the L_{A95} model, and the WS in the L_{A99} model accounted for more than 30% of variance. For night-time models, the RV and WS predictors primarily accounted for a greater of 30% variance across the background noise indicators. Finally, in all the time period models, the ST and WT predictors explained less than the 10% of variance.

For validating the performance of the developed models, the test datasets were employed and the observed outdoor-to-indoor levels were compared with the estimated levels. The performance of models based on the training and test datasets, is presented in Tab. 2.

Table 2. Modelling performance.

Model	BGN: 6dB(A)				BGN: 10dB(A)			
	Train Dataset		Test Dataset		Train Dataset		Test Dataset	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
$L_{\text{Aeq,LA90,D}}$	3.09	0.76	2.91	0.77	3.12	0.77	2.98	0.79
$L_{\text{Aeq,LA95,D}}$	2.77	0.80	2.97	0.77	3.25	0.77	2.64	0.83
$L_{\text{Aeq,LA99,D}}$	2.75	0.79	2.91	0.78	2.85	0.80	2.99	0.80
$L_{\text{Aeq,LA90,E}}$	3.41	0.71	2.53	0.83	4.08	0.59	4.00	0.61
$L_{\text{Aeq,LA95,E}}$	3.12	0.76	2.74	0.81	3.89	0.77	3.65	0.77
$L_{\text{Aeq,LA99,E}}$	3.36	0.71	2.94	0.76	3.90	0.66	3.42	0.76
$L_{\text{Aeq,LA90,N}}$	2.12	0.90	2.38	0.89	2.99	0.79	2.00	0.90
$L_{\text{Aeq,LA95,N}}$	1.93	0.92	2.28	0.89	2.87	0.82	2.22	0.88
$L_{\text{Aeq,LA99,N}}$	2.26	0.88	2.01	0.92	2.28	0.88	2.83	0.84

Green cells: RMSE≤3dB(A), Blue cells: RMSE>3dB(A)

As illustrated in Tab. 2 and with particular attention to the test dataset, night-time models exhibited the lowest RMSE levels, followed by day-time models across the background noise indicators and conditions. In the case of the evening models, a greater degree of variation was exhibited, with RMSE levels exceeding 3dB(A) for the more stringent background condition across all background noise indicators. In consideration of the less stringent background condition, RMSE levels lower than 3dB(A) were obtained in the evening models and across the background noise indicators.

3.4 Single-Valued Attenuation Factors

Focusing on the single-valued attenuation factors, the estimated factors were compared to the observed ones and the results are presented in Fig. 7.

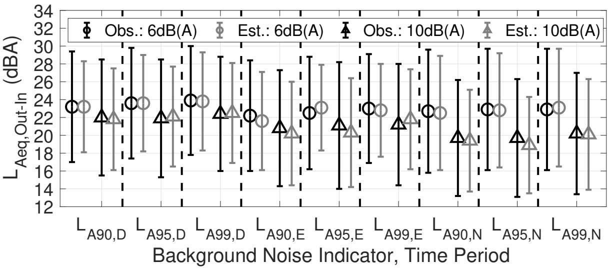

Figure 7. Single-valued attenuation factors per time period, background noise indicator and condition.

As Fig. 7 illustrates, the outdoor-to-indoor attenuation factors for the 6dB-based models are higher than those for 10dB-based models. These differences are more pronounced in the case of night-time noise exposure than of day-time and evening exposure, presenting mean differences up to 4dB(A). Regarding the background noise indicator, mean differences up to 2dB(A) were detected for evening and day-time periods.

4. DISCUSSION

4.1 Noise Assessment

The incorporation of methodologies for the identification of appropriate noise exposure levels, associated with the transmission of noise exposure from outdoors to indoors, could be characterised as important, especially when noise exposure measurements are conducted via an unsupervised approach. The assessment of the relationship between the outdoor and outdoor-to-indoor levels revealed a stronger correlation by increasing the background noise condition from 6dB(A) to 10dB(A) across the time periods and background noise indicators. Consequently, a decrease in the median/mean outdoor-to-indoor levels and an increase in their interquartile ranges across the time periods and background noise indicators was obtained. Consequently, this is associated with the further minimization of noise levels affected by background levels, and therefore with the inclusion of noise levels related to higher transmission of energy. Our results demonstrate analogous trends to those previously depicted in [5]. Focusing on the type of background noise indicator for mitigating indoor background noise levels, it was observed that there were no significant differences in the outdoor-to-indoor levels between the background noise indicators used.

Various outdoor-to-indoor attenuation factors have been reported for closed windows, ranging from 25dB(A) to 32dB(A) [4, 6, 7, 11]. These factors are contingent upon the characteristics of the measurement locations and the measurement methodology employed. In our study, analogous outcomes were obtained, particularly in the context of the 6dB(A) background noise condition, wherein the attenuation factor approximates 23dB(A) across the majority of the background noise indicators and time-periods. However, in the case of the 10dB(A) background noise condition, the outdoor-to-indoor level differences exhibited more pronounced fluctuations in terms of the time period, as opposed to the background noise indicators. Specifically, outdoor-to-indoor level differences of approximately 22dB(A) for day-time, 21dB(A) for evening, and 20dB(A) for night-time were obtained. This indicates the significant impact of the variations in indoor background noise levels and characteristics across the different time periods. Day-time outdoor and indoor levels reach higher levels in relation to the evening and night-time ones from various indoor and outdoor sound sources. Night-time exposure presents less level variations in both environments as well as lower indoor background noise levels, indicating a more representative period for exploring the noise propagation from outdoors to indoors.

4.2 Prediction Modeling

Regarding the performance of the prediction models, almost similar diagnostics were obtained across the different models. In order to gain a deeper understanding of the performance of the prediction models, the RMSE is interpreted by the ability of healthy ear to detect changes in broadband noise [12]. In this context, noise level differences of 3dB(A) and 5dB(A) are associated with just and clearly perceptible differences, respectively [12]. The majority of the day-time and night-time models are subjected to RMSEs that are up to the just perceptible level difference (3dB(A)), in contrast to the evening-time models where the RMSEs are up to the clearly perceptible level difference (5dB(A)). Regarding the background noise indicators, their performance presented similar diagnostic characteristics with respect to the time period. It is important to note that the number of predictors employed in this study was limited. Incorporating a greater number of measurements and predictors might lead to a more comprehensive dataset based on detailed built environment characteristics.

5. CONCLUSIONS

In this study, a methodology to derive outdoor-to-indoor noise attenuation levels for day, evening, and nigh time exposure was developed, utilising unsupervised outdoor and indoor noise exposure measurements with information related to building characteristics in random forest (regression) modelling. This method minimizes the presence of dominant indoor noise sources and background noise levels, considering the equivalent spectral characteristics of noise measurements. Across the various used background noise indicators, similar results were obtained. Finally, the attenuation factors obtained in the present study have been generated from UK data, reflecting localized noise sources and the characteristics of the built environment.

6. ACKNOWLEDGMENTS

This research was funded by the European Union's Horizon 2020 research and innovation program under grant agreement No. 874724, Equal-Life Project, part of the European Human Exposome network.

7. REFERENCES

[1] I. van Kamp, S. Simon, H. Notley, C. Baliatsas, and E. van Kempen, "Evidence Relating to Environmental Noise Exposure and Annoyance, Sleep Disturbance, Cardio-Vascular and Metabolic Health Outcomes in the Context of IGCB (N): A Scoping Review of New Evidence," *International Journal of Environmental Research and Public Health*, vol. 17, p. 3016, 4 2020.

[2] M. E. Terzakis, M. Dohmen, I. van Kamp, and M. Hornikx, "Noise Indicators Relating to Non-Auditory Health Effects in Children—A Systematic Literature Review," *International Journal of Environmental Research and Public Health*, vol. 19, p. 15633, 11 2022.

[3] W. Babisch, H. Ising, and Gallacher JE, "Health status as a potential effect modifier of the relation between noise annoyance and incidence of ischaemic heart disease," *Occupational and Environmental Medicine*, vol. 60, 10 2003.

[4] B. Locher, A. Piquerez, M. Habermacher, M. Ragettli, M. Röösli, M. Brink, C. Cajochen, D. Vienneau, M. Foraster, U. Müller, and J. Wunderli, "Differences between Outdoor and Indoor Sound Levels for Open, Tilted, and Closed Windows," *International Journal of Environmental Research and Public Health*, vol. 15, 1 2018.

[5] M. Terzakis, C. Van hoorickx, M. Hornikx, F. Naim, A. Hansell, and J. Gulliver, "Estimation of Low Frequency Outdoor-to-indoor Noise Attenuation Factors Based on Night-time Noise Exposure," in *30th International Congress on Sound and Vibration (ICSV)*, (Amsterdam, The Netherlands), pp. 1–8, 2024.

[6] WHO, "Environmental noise guidelines for the European region," *World Health Organization*, 2018.

[7] S. Pirrera, E. De Valck, and R. Cluydts, "Nocturnal road traffic noise assessment and sleep research: The usefulness of different timeframes and in- and outdoor noise measurements," *Applied Acoustics*, vol. 72, pp. 677–683, 9 2011.

[8] F. Naim, D. Fecht, A. Hansell, and J. Gulliver, "Assessment of residential exposure to aircraft, road traffic and railway noise in London: Relationship of indoor and outdoor noise," in *Euronoise, 2018*, (Crete, Greece), pp. 2905–2912, 2018.

[9] M. N. Wright and A. Ziegler, "ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R," *Journal of Statistical Software*, vol. 77, no. 1, 2017.

[10] E. Bartz and T. Beielstein, *Hyperparameter tuning and optimization applications. In: Hyperparameter tuning for machine and deep learning with R: a practical guide.* Springer, 2023.

[11] F. Scamoni and C. Scrosati, "The facade sound insulation and its classification," in *Forum Acusticum, Kraków, Poland*, 2014.

[12] E. Murphy and E. A. King, "Principles of Environmental Noise," in *Environmental Noise Pollution*, pp. 9–49, Elsevier, 2014.