

ARTICLE OPEN



Modelling the urban acoustic environment using land use-based gradient boosting

Timo Haselhoff¹✉, Susanne Moebus¹, Mikel Jedrusiak², Bryce T. Lawrence³ and Frank Weichert²

© The Author(s) 2026

BACKGROUND: Long-standing research on the relationship between the urban acoustic environment (AE) and human health demonstrates the harmful effects of environmental noise. Meanwhile, an increasing number of smaller studies report health benefits for additional acoustic properties. However, studies on health-promoting AEs remain limited, largely due to the lack of methods for estimating high-resolution acoustic properties beyond conventional noise metrics.

OBJECTIVE: We investigate to what extent models based on land-use types (LUT) can predict urban AE properties, focusing on four acoustic indices (Articulation Index, Bioacoustic Index, Link Density and Sharpness). Additionally, we predict the LAeq, which enables us to compare the performance between our model, the strategic noise map of Bochum (SNM) and results from the literature.

METHODS: We use a dataset of 2,746 acoustic measurements from 785 locations in Bochum and 90 measurements from 22 locations in Essen to train and evaluate gradient boosting models. For model development, data is split into training/validation (668 locations in Bochum) and test sets (117 locations in Bochum, all locations in Essen). The models predict acoustic indices based on the area of 77 LUTs within 50 and 300 m buffers around each location.

RESULTS: Based on the root mean square error (RMSE), predictions for Link Density deviate on average by 0.17 and 0.21 from test-sets in Bochum and Essen. For LAeq, the RMSE is 4.8 dB(A) and 4.4 dB(A), respectively. The R^2 for Link Density is between 0.27 and 0.3, and for the LAeq between 0.52 and 0.46. The SNM performs worse in predicting LAeq for Bochum data (RMSE = 7.8 dB(A); $R^2 = -0.31$). Performances for other indices are mixed.

IMPACT: This study advances research on the urban acoustic environment by demonstrating that land use type-based models represent a promising approach to predict acoustic indices beyond conventional noise metrics. Using over 2,800 measurements from two German cities, the models for predicting the Link Density and the LAeq show moderate to good performance on two test datasets. Model predictions for the LAeq outperformed strategic noise maps in predicting total environmental noise. These findings open new pathways for large-scale, population-based health research by providing a promising, scalable, high-resolution method for characterising complex urban acoustic environments, supporting efforts to design healthier urban environments through higher acoustic quality.

SIGNIFICANCE: LUT-based models demonstrate their potential for predicting Link Density and LAeq, achieving moderate to strong performance across two independent test datasets. This can provide a scalable approach for investigating potentially health-relevant properties of the urban AE at high spatial resolution.

Journal of Exposure Science & Environmental Epidemiology; <https://doi.org/10.1038/s41370-026-00855-w>

INTRODUCTION

Considerable research efforts have been made to investigate the relationship between the urban acoustic environment (AE) and human health. The adverse health effects of environmental noise exposure (e.g., cardiovascular disease and mental health issues) are well documented [1]. However, the AE can be defined as “the sound from all sound sources as modified by the environment” [2], encompassing a broader range of acoustic properties than environmental noise alone (e.g., underlying sound sources, temporal variations or frequency characteristics). A growing body of research is now examining the potential health benefits of additional acoustic properties [3]. Laboratory studies, for instance,

have shown that calm and pleasant sounds are associated with greater reductions in heart rate [4], that natural sounds can promote stress recovery [5], and that biophonic AEs can positively influence functional brain connectivity [6]. Beyond laboratory settings, several field studies have also investigated the impact of natural sounds on various aspects of human health, reporting associations with reduced pain, lower stress levels, and improved cognitive performance [7].

However, there is a significant lack of comprehensive population-based studies examining the relationship between the AE and human health beyond environmental noise, which are crucial for examining the patterns, causes, and effects of health

¹Institute for Urban Public Health, University Hospital Essen, Essen, Germany. ²Department of Computer Science, Computergraphics, TU Dortmund University, Dortmund, Germany. ³Department of Spatial Planning, TU Dortmund University, Dortmund, Germany. ✉email: timo.haselhoff@uk-essen.de

Received: 23 July 2025 Revised: 9 February 2026 Accepted: 9 March 2026

Published online: 25 March 2026

outcomes in defined population groups within real-life settings [8] (pp. 555–583). Such studies require high spatial resolution data to assess exposure to the respective acoustic properties under investigation. For research on the relation between environmental noise exposure and human health, data are already available through strategic noise maps (SNMs), mandated by the Environmental Noise Directive (END) in the EU [9]. Here, SNMs with high spatial resolution are created for agglomerations of over 100,000 inhabitants, as well as for major roads, railways, and airports. The SNMs assess exposure to “unwanted or harmful outdoor sound created by human activities” [9]. While they are effective at quantifying noise emitted by “transport, road traffic, rail traffic, air traffic, and from sites of industrial activity” [9], there is currently no established method to estimate exposure to additional acoustic properties of the AE at a comparable spatial resolution. Consequently, the current unavailability of high spatial resolution data on the AE beyond environmental noise exposure is a major barrier to conducting population-based studies on the health impacts of the wider AE.

A promising approach to overcome this barrier of exposure assessment involves the use of land use-based models. These models can estimate environmental exposures at high spatial resolution by leveraging the statistical relationships between measured exposure data (e.g., air pollution) and land use types (LUTs) (e.g., highways or green space). Land use-based models were already successfully used for the estimation of outdoor air pollution [10]. The applicability of such models to acoustics is demonstrated by their effective use in estimating average noise levels, traffic noise, and in the extrapolation of SNMs [11–21]. Some of the major advantages of the application of LUT-based models are that they are cost-effective and computationally inexpensive, especially when compared to more complex alternatives like the creation of SNMs [17]. In addition, they are easily scalable as LUT information is often readily available and thus suited for data-poor regions where detailed exposure measurements are limited or unavailable (e.g., agglomerations with less than 100,000 inhabitants). Furthermore, as they are based on land use information, they can be directly integrated into urban planning processes. However, to the best of the knowledge of the authors, land use-based models have not yet been applied to estimate acoustic indices beyond noise.

In this work, we investigate the potential of LUT-based models to estimate acoustic properties beyond environmental noise exposure, as defined by the END. We focus on the acoustic properties captured by the Articulation Index [22], the Bioacoustic Index [23], the Link Density [24] and the maximum Sharpness [25]. These indices were chosen, as previous works have already shown their relations to LUTs as well as the human perception of pleasantness [26], health-related urban greenspace [27] and biophonic activity [23, 28], thus indicating potentially health-relevant acoustic properties. In addition, we integrate the A-weighted sound pressure level (LAeq) as a measure of total environmental noise. This also enables us to compare the performance to predict total environmental noise between our model, the SNM and results from the literature. To facilitate a scalable approach, we rely exclusively on LUTs as predictor variables because they are usually easily accessible and, in many cases, even legally mandated. Specifically, the research questions are:

1. To what extent can gradient boosting models based on LUTs predict properties of the urban AE at high spatial resolution?
2. How well do these models perform in estimating:
 - i. Acoustic properties beyond noise, when tested on the same-source data and external-source measurements?
 - ii. Total environmental noise levels, when tested on the same-source data, external-source measurements and in comparison to SNM predictions?

MATERIAL AND METHODS

To address our research questions, we train a gradient boosting model on LUT data provided by local authorities and acoustic measurements from the SALVE (AcouStic QuAlity and Health in Urban EnVironmEnts) project in Bochum [29]. We evaluate the performance of the model using 5-fold cross-validation, test data from SALVE, as well as acoustic measurements from the Be-MoVe (Participation-based transformation of active mobility for health-promoting urban and transport infrastructures) project conducted in Essen, Germany [30]. Additionally, we compare the model's performance in estimating total noise levels with the performance of the SNM of Bochum. Finally, we demonstrate the application of the model to the area of Bochum, Essen and the neighbouring city of Mülheim an der Ruhr.

Data on the acoustic environment

SALVE. To train and test our model, we rely on acoustic measurements from the SALVE project in Bochum. A comprehensive description of the SALVE study design is provided in Haselhoff et al. [29]. The sampling strategy was designed to capture the environmental contexts in close proximity to people's homes. The recording location selection was based on stratified LUT sampling as described in Haselhoff et al. [29]. At each sampling point, four audio recordings (5-min, 48 kHz, 24-bit depth) were made once each season between 03.2019 and 03.2020. Recordings took place between 09:00 and 17:00. The recording device was an NTi XL2 sound recorder with the M2230 omnidirectional microphone [31], mounted at a height of around 1.65 m. The device was calibrated to meet the standards IEC 61672:2013, IEC 61672:2003, IEC 61260:2014, IEC 61260:2003, IEC 60651 and IEC 60804. A total of 2746 audio recordings were gathered at 785 locations. The number of recording locations differs from 730, as we considered measurement location deviations of more than 5 m as significant.

Be-MoVe. To test our model, we draw data from the Be-MoVe project, which aimed to test co-created alternative forms of mobility and designs of public spaces in urban neighbourhoods. The study design is described in detail in Hornberg et al. [30]. Shortly, as part of the Be-MoVe project, soundwalks according to DIN ISO 12913 [2] were conducted to assess the acoustic environment at 22 locations in two districts in Essen. The districts are characterised by (i) apartment buildings, recreational areas and open space, and (ii) a diverse urban landscape, ranging from shopping streets and commercial centres to densely populated residential zones. For each soundwalk, at each listening station, an acoustic measurement (3-min, 48 kHz, 24-bit depth) was conducted, using the same device as within the SALVE project. Recordings were made between 10.05.2023 and 06.09.2023 between 18:00 and 19:30. Originally, the project was conducted from 2022 to 2023, but acoustic measurements from 2022 were made using an uncalibrated recording device. In total, we included 90 measurements across 22 listening stations.

Acoustic properties. For each recording, a number of acoustic indices are derived: The total noise level is assessed using the A-weighted sound pressure level (LAeq), which was directly provided by the recording devices. As a measure of intelligibility of human voice, we use the Articulation Index based on the root mean square (RMS) of the signal. The Articulation Index indicates how much background noise level can interfere with human speech and ranges from 0 (no speech understood) to 1 (all speech understood) [22]. As a measure of the energetic centre of gravity between frequencies, we use the maximum Sharpness value during each recording, according to DIN 45692 [25]. It is measured in Acum and ranges from zero to infinity, where higher values indicate sounds with more energy in higher frequency bands (i.e., whether a sound is perceived as sharp, shrill, bright or hissing). The Articulation Index, as well as the Sharpness are calculated using Artemis SUITE (version 15.1). Additionally, we use the Bioacoustic Index (BIO). The BIO ranges from zero to infinity, where higher values indicate greater differences between the quietest and the loudest 1 kHz frequency band between 2 and 8 kHz. In more rural areas, higher values shall indicate higher avian abundance, while in urban areas, higher values are found to be more indicative of road traffic and less green space [23, 28]. The BIO is calculated following the script used in Lawrence et al. [32]. Furthermore, we include the Link Density as a measure of acoustic dominance, i.e., a measure of how many factors (i.e., sound sources) contribute to the overall spectral dynamic. It ranges from zero to one, where higher values indicate higher acoustic dominance (e.g., by cars passing by). The Link Density is calculated following the procedure outlined in Haselhoff et al. [27].

As we want to predict the average acoustic properties at each location, we calculate the mean of all indices (using energetic averages for the LAeq) for each location from the measurements conducted at each sampling point, resulting in 785 measurement locations for the SALVE and 22 for Be-MoVe.

Training, validation, and test split. We partition the SALVE dataset into a training/validation (85%, $n = 668$) and a test (15%, $n = 117$) subset, orientating us at Almansi et al. [33] and the commonly used 70/30 split between training/validation and test data [34]. We call the latter test split SALVE_{Test}. The training/validation data are used for model development through 5-fold cross-validation. The SALVE_{Test} subset and the Be-MoVe dataset are used to evaluate the model's performance on unseen data.

Strategic noise maps. To evaluate our model's performance in predicting total environmental noise, we additionally compare it to the performance of the SNM in Bochum. For this, we received the SNM from the city of Bochum in 2022. To be comparable to our recordings, we use the L_{Day} in 1 dB(A) increments, i.e., the A-weighted long-term average sound level between 06:00 and 18:00, to match our recording times. The provided values range from 34.5 to 79.5 dB(A). Since there are multiple SNM depicting different noise sources (road traffic, railway traffic and industry), we combine them by overlaying the maps and energetically summing their respective LAeq values. As the SNM has a resolution of 5×5 m, we compare the LAeq point measurements from SALVE with corresponding polygon values from the SNM by extracting the predicted LAeq from the SNM at each measurement location.

Environmental data

The predictor variables are based on land use data provided by the Ruhr Regional Association (RVR) for the year 2019 [35]. In total, there are 146 LUT categories. We derive the LUTs around each recording location by calculating the proportion of each LUT category within buffers of (i) 50 m, to capture the immediate surroundings of each location, and (ii) 300 m radius, to capture potentially important acoustic impacts of LUT within a wider distance. Altogether, this results in 292 possible predictor variables per recording location.

Methods

Descriptive statistics. To provide an overview of the acoustic properties of our recordings, we show descriptive statistics, including the arithmetic mean, the minimum, the maximum and the standard deviation of the SALVE and the Be-MoVe datasets. In addition, since we have multiple recordings for each location, we report the within-location standard deviation to capture the inherent variability of respective acoustic properties. The reported within-location standard deviation is the arithmetic mean of each within-location standard deviation per recording location.

Feature reduction. To improve the efficiency of the gradient boosting models, we remove non-related LUT categories from the set of predictor variables by using Boruta feature selection. Briefly, Boruta feature selection is a method that identifies all relevant features by comparing the importance of the actual variables to that of randomly permuted "shadow features" of these variables, using a random forest classifier [36]. We repeat this for each acoustic index against all 292 LUT variables in our dataset and keep those LUT variables, which were identified by the algorithm as "important" or "tentative". Finally, we harmonize all identified LUT variables into a single predictor set for all acoustic indices (Appendix 1). We perform the calculation using the "Boruta" function from the Boruta package (version 8.0.0) in R (4.1.3), applying the default settings as recommended by the authors [36].

Gradient boosting. We train five separate gradient boosting regressor models to predict the respective acoustic indices using the LUT features identified by the Boruta algorithm [36]. Gradient boosting builds a predictive model by sequentially combining multiple weak learners. The gradient boosting framework consists of three key components: (i) a loss function to be minimized, (ii) a weak learner for generating predictions, and (iii) an additive model that incorporates new learners to reduce the residual errors of the ensemble. There are three main hyperparameters, which impact the performance of the model: the number of trees, the maximum depth of the trees, and the learning rate [37]. We optimize these parameters by performing a grid search combined with a 5-fold cross-validation. A 5-fold cross-validation is an evaluation technique used to assess how well a model

generalizes to unseen data. The dataset is randomly divided into five equal parts (folds). In each of five iterations, four folds are used to train the model, and the remaining fold is used for validation. This process repeats until each fold has been used once for validation. The final performance metric is the average of the five validation results. For the grid search, we iterate over the hyperparameters number of trees (10, 50, 100, 500, 1000, 5000), maximum tree depth (3, 4, 5, 6), and learning rate (0.0001, 0.001, 0.01, 0.1, 1.0). For each hyperparameter combination, a 5-fold cross-validation is used to evaluate the model performance. The optimal hyperparameter combination is selected based on the lowest root mean square error (RMSE). In the following, we call the model based on 668 locations SALVE_{Train}. Once the optimal hyperparameters are defined, and after the model is tested against the SALVE_{Test} data, we train the model on the entire SALVE dataset to maximize its learning from all available data and prepare it for the evaluation on the Be-MoVe data. This model is referred to as SALVE_{All}.

Furthermore, as we observed significant spatial autocorrelation in selected acoustic indices in previous publications [27], which can lead to artificial inflation of performance measures, we investigate the robustness of our results using leave-spatial-out-cross validation [17]. Recording locations were first grouped according to spatial proximity using a regular grid-based approach (300 and 1000 m). These spatial groupings were then used to define the folds in a grouped k-fold cross-validation scheme, ensuring that each spatial group was used exactly once as the test set across all folds.

Performance measures. To evaluate the performance of each model, we mainly rely on three performance measures: The root mean square error (RMSE), the mean absolute error (MAE) and the coefficient of determination (R^2). RMSE provides a measure of the average magnitude of the prediction error, with greater weight given to larger errors. MAE quantifies the average absolute difference between predicted and observed values, offering a more interpretable and less sensitive measure to outliers compared to the RMSE. To facilitate a comparison of model performance, the MAE and RMSE are further normalized by (i) the range (i.e., the difference between the maximum and minimum values) and (ii) the mean of the acoustic index within the respective target dataset. In addition, R^2 indicates the proportion of variance in the observed data explained by the model, with values closer to 1 signifying better predictive performance. We provide a visualisation to compare the model predictions against the true values using scatter plots (including Pearson's r) as well as histograms.

Shap. For a more detailed investigation of the model's performance, we assess the importance of each predictor variable using SHAP (SHapley Additive exPlanations). SHAP values provide a unified measure of feature importance by quantifying the contribution of each variable to individual predictions [38]. This approach allows for a consistent and interpretable explanation of the model output by attributing the prediction difference from the mean to each feature. We report SHAP values using a beeswarm plot for the five most important predictors for each model. SHAP values are derived by the application of the SALVE_{All} model on the entire SALVE dataset.

Sound maps

To demonstrate the scalable properties of the LUT-based model, we generate "sound maps" for the cities of Bochum, Essen, and Mülheim. If the models demonstrate sufficient performance, they could enable population-level exposure estimates to AE properties beyond conventional environmental noise metrics. To construct the sound maps, we overlay a grid of equally spaced sampling points (30 m apart) across the three cities. For each point, we calculate the LUT variables used in the model within buffers of 50 m and 300 m. We then apply the final prediction model to estimate acoustic properties at each location. For visualisation purposes only, we use Kriging [39] to interpolate values between the point estimates, to create a continuous surface of predicted values.

We calculate gradient boosting, hyperparameter tuning and performance measures using the sklearn package (0.42.2) and SHAP values using the shap package (0.46.0) in Python (3.9.7). The calculation of LUT areas in two buffer sizes, the sampling of the sound map points and the Kriging are performed using ArcGIS (3.2.0).

RESULTS

Data description

The LAeq for the SALVE dataset has a mean of 55.3 dB(A) with a range between 33.6 and 79.3 dB(A) (Table 1). The standard

deviation (STD) is higher between recording locations than within recording locations, showing a more stable sound pressure level at each location than between locations. In the Be-MoVe dataset, the STD is also higher between recording locations (5.1 dB(A)) than within them (2.5 dB(A)), but the values are overall lower here, underlining a less diverse AE in terms of LAeq than the one measured in the SALVE project. The Articulation Index has a mean of 0.846, a minimum of 0.207 and a maximum of 1 in the SALVE dataset, while also having a much higher between (0.136) than within STD (0.076). This tendency is even stronger in the Be-MoVe dataset, with a between STD of 0.146 and a within STD of 0.051. In contrast, the between STD for the BIO is smaller than the within STD in both datasets (SALVE: 0.985, 1.190; Be-MoVe: 0.824, 1.249), showing a higher variation within recording locations than between. For the Link Density, this behaviour is also observable for the SALVE dataset, but much less pronounced, with a between STD of 0.203 and a within STD of 0.210. Furthermore, the mean is higher in SALVE with 0.563 against 0.478 in Be-MoVe, and the Be-MoVe minimum (0.081) and maximum (0.927) fall within the range between the SALVE minimum (0.05) and maximum (0.984). The mean sharpness of the SALVE dataset is 2.029 Acum, ranging from 1.4 to 3.328 Acum. Similar to the Bioacoustic Index, a higher within STD (0.356 Acum) than between STD (0.305 Acum) can be observed, which is also the case for the Be-MoVe dataset (0.266 & 0.185 Acum resp.).

Modeling

Feature reduction. From the 292 initial predictor variables, 215 are identified as unimportant, resulting in a list of 77 LUT categories relevant for predicting acoustic properties in the urban environment (28 LUT categories for a 50 m buffer and 49 LUT categories for a 300 m buffer). A list of all relevant LUT categories can be found in Appendix 1.

Hyperparameter tuning. We apply a grid search in combination with a 5-fold cross-validation to tune the hyperparameters of the gradient boosting model on the test/validation data of the SALVE dataset. We find the optimal parameter for number of trees, maximum tree depth, and learning rate for the indices LAeq (1000, 3, 0.01), Articulation Index (5000, 3, 0.001), Bioacoustic Index (500, 3, 0.01), Link Density (500, 3, 0.01) and Sharpness (5000, 3, 0.001). The results for the performance measures from the final 5-fold cross-validation can be found in Table 2. For the LAeq, we find an MAE of 3.9 (STD = 0.4) and an RMSE of 5 (0.5). The higher RMSE indicates an impact of larger outliers between predictions and measurements. This pattern can be found for each acoustic index, with the RMSE being approximately 1.3 times greater than the MAE. Overall, the R^2 varies considerably across the models, ranging from a low of 0.133 (0.078) for Sharpness to a high of 0.522 (0.054) for the Articulation Index. The model for the BIO also shows a relatively low explanatory power (0.135 ± 0.104), while for the LAeq, it is located at the upper end (0.463 ± 0.058), and for Link Density, it falls in the mid-range (0.286 ± 0.056). These results indicate that the models show varying predictive quality [40]. In addition, the relatively low STDs in comparison to the means indicate a stability of model performance across different folds.

Model performance

To investigate the model performance on unseen data, we apply the SALVE_{Train} model on the SALVE_{Test} data and the SALVE_{All} model on the Be-MoVe data. Results can be found in Table 3 and in Appendix 2. For a more detailed investigation, we also provide scatter plots and histograms between true values and model predictions, as well as SHAP values to investigate the most predictive features (Fig. 1).

For the LAeq predictions of the SALVE_{Test} dataset, we find slightly lower MAE and RMSE as well as an increased R^2 in

Table 1. Descriptive Statistics for the SALVE and Be-MoVe dataset.

	SALVE					Be-MoVe				
	Mean	Min.	Max.	STD Between	STD Within	Mean	Min.	Max.	STD Between	STD Within
n	785	785	785	785	2746	22	22	22	22	90
LAeq	55.3	33.6	79.3	6.5	4.2	60.3	50.4	68.9	5.1	2.5
Articulation Index (RMS)	0.846	0.207	1.000	0.136	0.076	0.700	0.447	0.942	0.146	0.051
BIO	5.882	2.425	10.433	0.985	1.190	7.424	6.454	8.694	0.824	1.249
Link Density	0.563	0.05	0.984	0.203	0.210	0.478	0.081	0.927	0.234	0.167
Sharpness (Max)	2.029	1.400	3.328	0.305	0.356	1.901	1.585	2.385	0.185	0.266

Displayed are the arithmetic mean, the minimum (Min.), the maximum (Max.), the standard deviation between the average of each recording location (STD Between) and the average of the standard deviations of the measurements at the respective recording locations (STD Within). The unit of LAeq is dB(A). The unit of Sharpness is Acum.

Table 2. Results of the 5-fold cross-validation.

	MAE	RMSE	R^2	nMAE	nRMSE
LAeq	3.9 ± 0.4	5.0 ± 0.5	0.463 ± 0.058	0.115 ± 0.009	0.147 ± 0.014
Articulation Index (RMS)	0.072 ± 0.006	0.095 ± 0.006	0.522 ± 0.054	0.116 ± 0.020	0.155 ± 0.025
BIO	0.675 ± 0.063	0.903 ± 0.065	0.135 ± 0.104	0.118 ± 0.017	0.157 ± 0.020
Link Density	0.138 ± 0.006	0.171 ± 0.006	0.286 ± 0.056	0.158 ± 0.004	0.195 ± 0.006
Sharpness (Max)	0.212 ± 0.015	0.277 ± 0.019	0.133 ± 0.078	0.133 ± 0.019	0.174 ± 0.024

Displayed are the means \pm standard deviations of the respective performance measure.

MAE Mean Absolute Error, RMSE Root Mean Square Error, R^2 = Coefficient of Determination, nMAE MAE/Range, nRMSE RMSE/Range.

The unit of LAeq is dB(A). The unit of Sharpness is Acum.

Table 3. Performance measures for model application on test data.

	SALVE _{Test} (n = 117)					Be-MoVe (n = 22)				
	MAE	RMSE	R ²	nMAE	nRMSE	MAE	RMSE	R ²	nMAE	nRMSE
L _{Aeq}	3.8	4.8	0.518	0.134	0.168	3.5	4.4	0.458	0.167	0.210
Articulation Index (RMS)	0.056	0.078	0.609	0.114	0.160	0.123	0.156	-0.149	0.249	0.316
BIO	0.652	0.851	0.300	0.122	0.159	1.102	1.242	-3.692	0.491	0.555
Link Density	0.140	0.172	0.273	0.158	0.194	0.156	0.205	0.312	0.184	0.242
Sharpness (Max)	0.216	0.290	0.256	0.132	0.178	0.205	0.239	-0.673	0.257	0.299

For SALVE_{Test} the model trained on the SALVE_{Train} model is used. For Be-MoVe, the SALVE_{All} model is used.

MAE Mean Absolute Error, RMSE Root Mean Square Error, R²=Coefficient of Determination, nMAE MAE/Range, nRMSE RMSE/Range.

The unit of L_{Aeq} is dB(A). The unit of Sharpness is Acum.

comparison to the 5-fold CV results. This tendency is more pronounced for the model application on the Be-MoVe data, with even lower values for MAE and RMSE. However, all values fall within the results from 5-Fold cross-validation (\pm STD), underlining the robustness of the model for completely unseen L_{Aeq} data. For the Articulation Index, the BIO and the Sharpness, we find that the model's performance is comparable to or even better than that of the cross-validation when applied to the SALVE_{Test} data. However, the performance decreases strongly when applied to the Be-MoVe data. Here, MAE and RMSE double for the Articulation Index and increase by a factor of approx. 1.5 for the BIO. Although the values slightly decrease for Sharpness, the R² value drops below 0 for all three indices, indicating poor explanatory power in the model's prediction for the Be-MoVe data. This is underlined by investigating the scatter plots of the linear fit between predicted and measured values (Fig. 1). With a Pearson's *r* of 0.23 (95% CI [-0.21, 0.6]) and -0.18 ([-0.56, 0.26]), the BIO and the Sharpness show close to no substantial linear relation respectively, especially when compared to the model application on the SALVE_{Test} data (*r* = 0.57 [0.44, 0.68] and 0.54 [0.4, 0.66] resp.). In contrast, the Articulation Index still exhibits a linear but weaker relationship, with the *r* decreasing from 0.8 [0.72, 0.85] to 0.59 [0.22, 0.81]. Especially pronounced is the heteroscedastic pattern, which shows a very good fit for higher but an increasingly poor fit for lower Articulation Index values. The performance measures for Link Density remain relatively robust. Results on the SALVE_{Test} data are consistent with those from cross-validation, while the application to the Be-MoVe data shows a slight increase in MAE and RMSE, along with a modest increase in R².

Regarding the predictive importance of LUT categories, SHAP values indicate that "Main Streets" are consistently among the top predictors—ranking as the most important variable for all indices except Sharpness, where it ranks second. Similarly, "Highways" (for Articulation Index and Sharpness) as well as "Residential streets" (for BIO and Link Density) are among the top five most important predictors. Furthermore, different commercial areas play an important role for the AE in regard to Articulation Index and Link Density. The presence of all these LUT almost exclusively increases (BIO & Link Density) or decreases (Articulation Index & Sharpness) the respective indices. The opposite can be observed for the frequently important predictor "Deciduous Forest". Looking at the distribution by buffer sizes, the immediate surrounding seems to be slightly more important than the broader environment for predicting the acoustic indices, with twelve out of the twenty top five predictors relating to the 50 m Buffers. Here, the Link Density stands out, as all of its most important predictors belong to the 50 m buffer.

The results from leave-spatial-out-cross-validation (LSOCV) to account for the potential impact of spatial autocorrelation on the model performance can be found in Appendix 3. By applying LSOCV using spatial groupings based on a 300 and 1000 m regular grid, we find no artificial inflation of model performance, as all measures fall between the mean \pm the standard deviation reported in Table 2.

Model performance in comparison to the strategic noise Map

For an improved understanding of the results and an embedding in the overall context of predicting the environmental AE, we compare the performance of our model in predicting total environmental noise to the performance of the SNM of Bochum (Fig. 2). For all investigated measures, we see an overall improved performance of the LUT based models against the SNM (Fig. 2a). In direct comparison, the SNM MAE for predicting the L_{Aeq} is 1.84 dB(A) higher than that of the SALVE_{Train} model in predicting L_{Aeq} measures from SALVE, while the RMSE is 3 dB(A) higher. Considering the R², the SNM estimates show a poor performance in predicting total environmental noise, with a value of -0.307, indicating that the model performs worse than a naive model using the arithmetic mean of the test data. However, in contrast to the BIO and Sharpness models, the scatter plots (Fig. 2b) still reveal a substantial linear relation of *r* = 0.57 (95% CI [-0.52, 0.61]) between predicted and measured L_{Aeq} values. The biggest deviations between measured and predicted L_{Aeq} values are found for low predictions around 35 dB(A) and measured values between approx. 40 and 70 dB(A).

This tendency to an underestimation by the predictions is further emphasized by the histogram, which shows a right-skewed distribution of residuals. From there, we see a residual-range from -13.5 dB(A) to 35.1 dB(A). In contrast, residuals are more symmetrically distributed around 0 for the model predictions of the SALVE and Be-MoVe data, ranging from -11.3 to 16.7 dB(A) and -5.4 to 8.7 dB(A). Still, underestimates remain more prevalent also from these models.

DISCUSSION

The goal of this work is to investigate whether LUT-based models can predict properties of the urban AE and how they perform in doing so. As LUT data is often widely available, such models offer an efficient approach for estimating AE properties at high spatial resolution—information that is critical for large-scale studies on AE properties beyond noise modeled by SNMs.

For the LUT model applied to predict L_{Aeq}, all performance metrics indicate improved performance compared to the predictions from the SNM. We find improved (i.e., decreased) MAE (by approx. 2 dB(A)) and RMSE (3 dB(A)) for the model application on unseen data from two datasets. As an increase of 3 dB(A) corresponds to a doubling in energy, these represent substantial differences between the LUT-based model and the SNM. This is underlined by comparing the models' R², which even becomes negative for the SNM predictions. However, it should be noted that SNMs are not specifically designed to estimate total environmental noise, but rather noise from specific sound sources (major road, rail and air traffic as well as industry noise). Furthermore, estimates are made for a height of ~4 m and at a resolution of 5 x 5 m. Therefore, the SNM may be unable to capture acoustic differences at finer scales, unlike point estimates from the gradient boosting model, which, in theory, can achieve

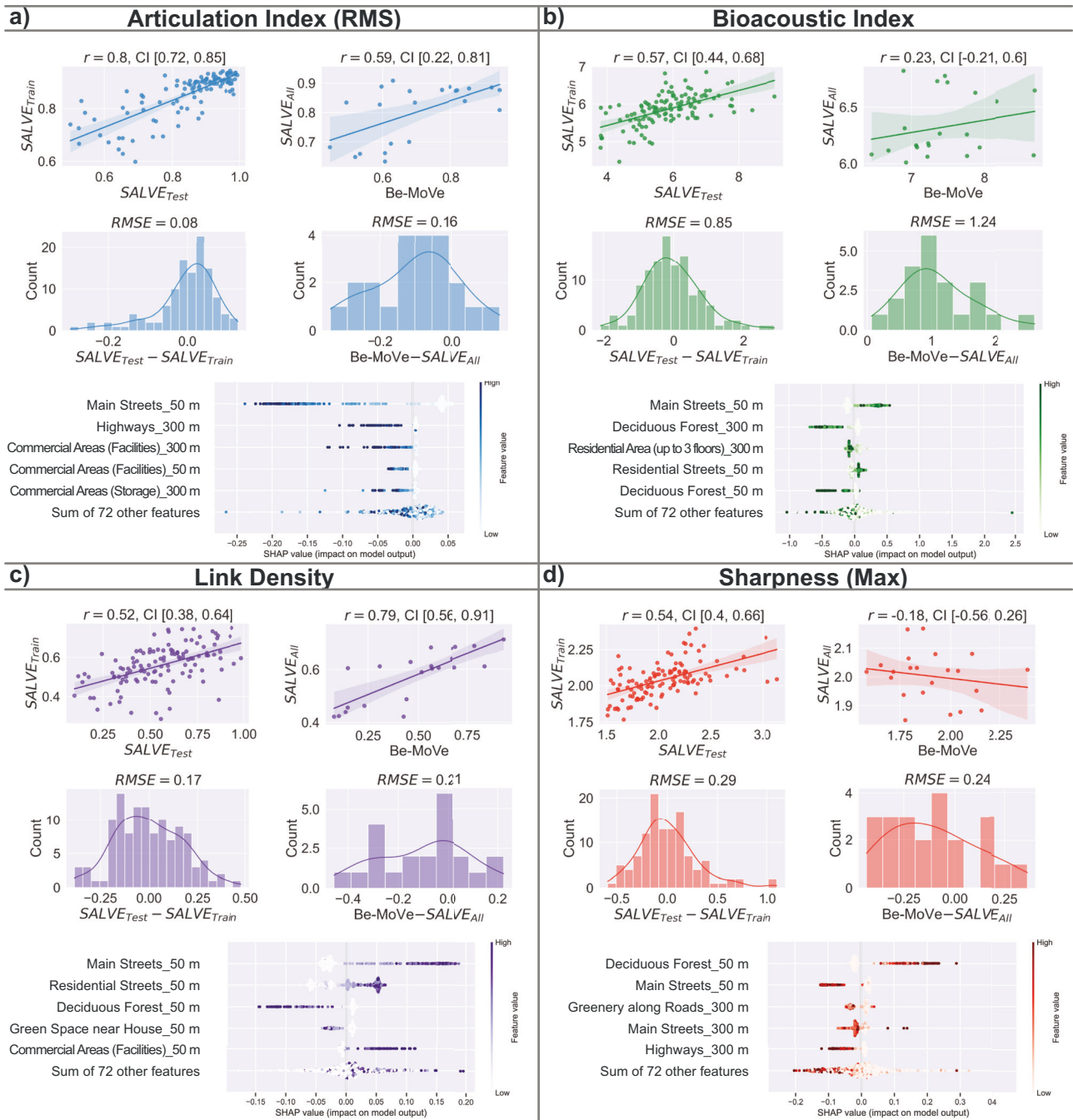


Fig. 1 Model performance for Articulation Index, Bioacoustic Index, Link Density & Sharpness. Model performance is visualized using scatter plots between test data and model predictions (incl. 95% confidence interval), histograms of the difference between test data and model predictions as well as SHAP values for the top five most important predictor variables in the gradient boosting model for Articulation Index (a), Bioacoustic Index (b), Link Density (c) and Sharpness (d); (RMSE = Root Mean Square Error, r = Pearson correlation). The unit of LAeq is dB(A). The unit of Sharpness is Acum.

arbitrarily high resolution. In addition, the exclusion of roads with fewer than six million vehicle passages a year complicates the use of SNMs for total noise assessment. Although SNM results are often treated as synonymous with total noise pollution, our results should not be viewed as a shortcoming of the SNM. Rather, the results highlight that there are substantial noise sources contributing to total environmental noise, which are not considered by SNMs. This finding is in line with several results from the literature, which also highlight the conceptual shortcoming of SNM when interpreted as total environmental noise [41, 42].

Our findings are largely in line with those of previous studies that modeled comparable noise concepts based on land use data. Liu et al. [14] report a MAE of 3.47 dB(A) and a RMSE of 4.44 dB(A) with an R^2 of 0.58 in predicting LAeq measurements from five Canadian cities. Aguilera et al. [11] report R^2 values between 0.66 and 0.87 and also provided comparisons to SNMs, with a Pearson's r^2 between 0.38 and 0.61. However, they predict road traffic noise, which represents only a part of the urban AE. Another study compared noise predictions from five different models to noise measurements from five cities in Bulgaria, focusing mainly on traffic lanes and industrial sites [43]. They find the best performing

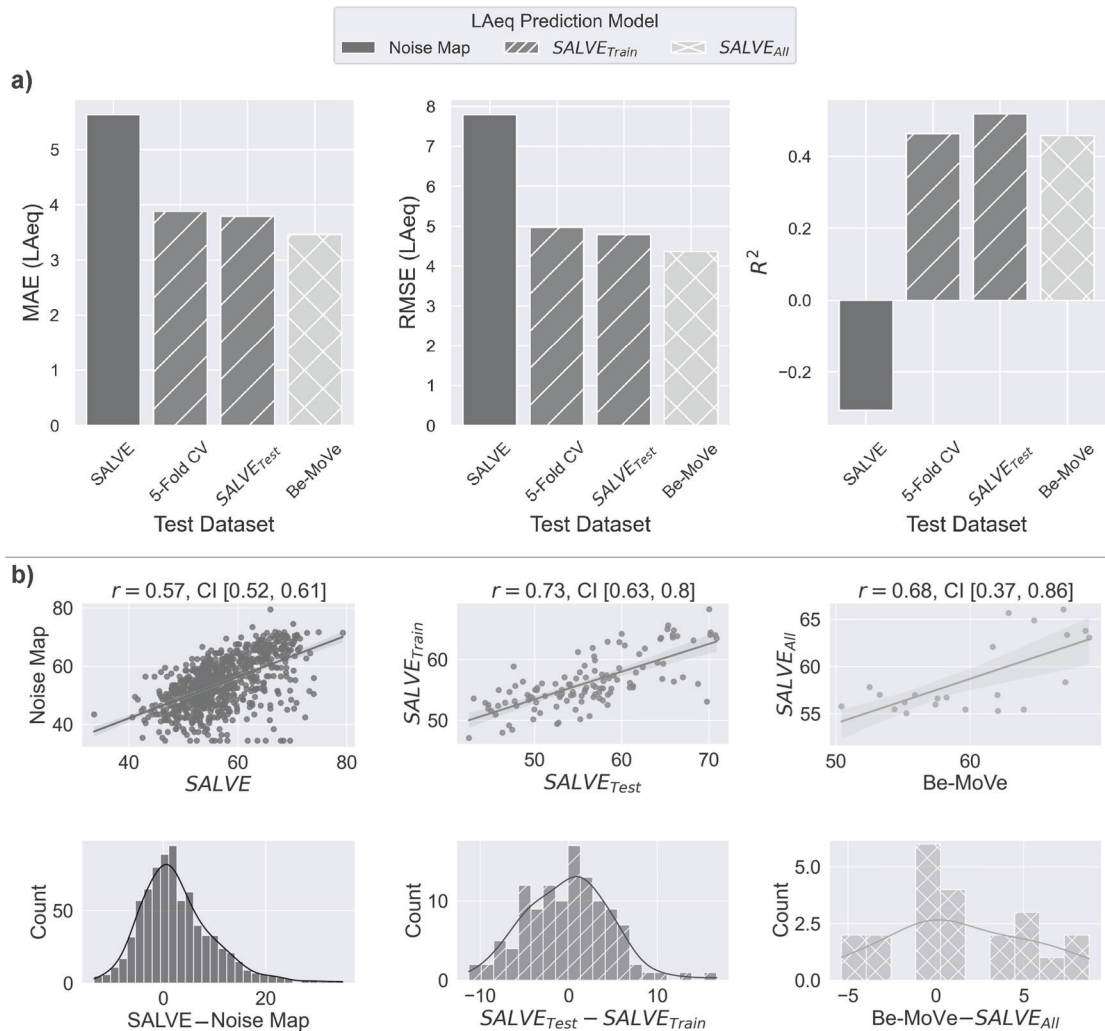


Fig. 2 Model performance for LAeq. **a** Model performance comparison in predicting total environmental noise, using the strategic noise map, the SALVE_{Train} model and the SALVE_{All} model against test data. “SALVE” corresponds to the total SALVE dataset of 785 recordings; (MAE=Mean Absolute Error, RMSE=Root Mean Square Error, R²=Coefficient of Determination). **b** Model performance visualized using scatter plots between test data (x-axis) and model predictions (incl. 95% Confidence Interval) (y-axis), as well as histograms of the difference between test data and model predictions. LAeq in dB(A) is given, if not indicated otherwise; (r = Pearson correlation, CI=Confidence interval).

model to be extreme gradient boosting, with an RMSE of 4.74 dB(A) and an R² of 0.68. In addition, there are several other studies that support our findings [13, 15, 16, 44]. Although these models include additional information to LUTs (e.g., traffic volume, meteorological data) and predict different forms of noise, their performance is close to what we find in our study. In comparison to the noise map performance here (RMSE = 7.8 dB(A); R² = -0.31), these results suggest that such models could be more effective in predicting noise exposure at a high spatial resolution.

In addition to predicting the LAeq, we also predicted the Articulation Index, the BIO, the Link Density and the maximum Sharpness. Here, results are mixed. While for all indices, a moderate to good performance on the SALVE_{Test} data can be observed, performance drops substantially when applied to the Be-MoVe data. This is especially true for the Articulation Index model, which showed the highest R² of all indices (0.609), but became negative for the Be-MoVe data (-0.149). Inspecting the scatter plots, we find that there is still a substantial linear relation between predicted and measured value ($r = 0.59$, 95% CI [0.22, 0.81]), though model performance clearly declines at Articulation Index values lower than ~0.8. As 68% of the Be-MoVe datasets' values are below this value, this might explain the low

performance for predictions in more dense and diverse built-up urban environments. Still, the model performs reasonably well in predicting Articulation Index values at locations with a high percentage of intelligible speech. No such patterns nor substantial linear relationships are found for the models predicting the BIO and the Sharpness. In both cases, the R² becomes negative for the Be-MoVe data. For the BIO, the MAE and the RMSE also strongly increase, while they stay approximately the same for the Sharpness. Overall, this indicates that the models incorporating only LUT information are not performing well in predicting these indices. One reason for that could be that the indices focus on frequency power-related information, rather than overall sound pressures. Since frequency power tends to vary more over time than across locations [45], a model based on seasonal averages may overlook important temporal dynamics. In addition, the suitability of the BIO and Sharpness indices for capturing meaningful information about the urban AE is still subject to debate [27]. In contrast, the model predicting the Link Density shows similar performance measures between the applications on both datasets. Although the R² values between 0.27 and 0.31 only show a moderate fit, a r of 0.52 (95% CI [0.38, 0.64]) for the SALVE_{Test} data and of 0.79 [0.56, 0.91] for the Be-MoVe dataset

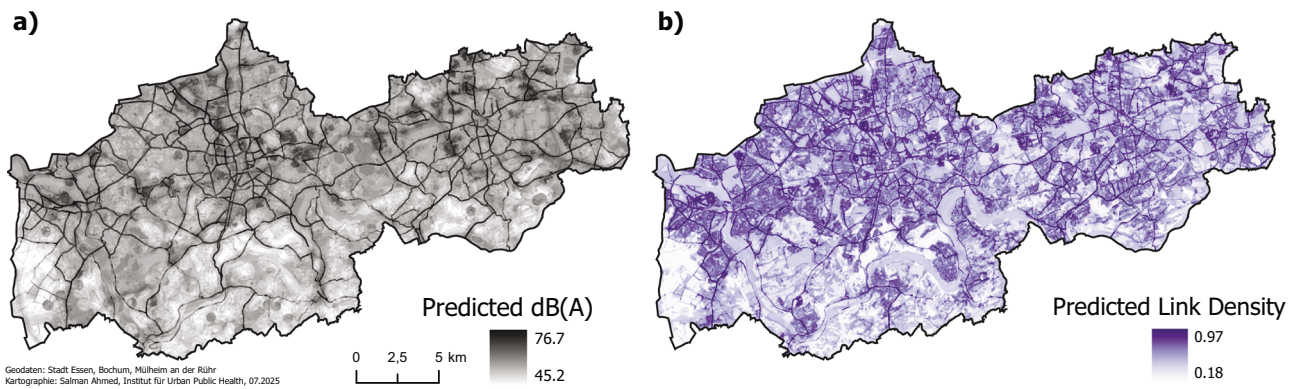


Fig. 3 Predicted sound maps for LAeq and Link Density. Application of the gradient boosting models to predict total environmental noise (a) and link density (b) for the area of the cities Bochum, Essen and Mülheim an der Ruhr.

indicate a robust performance across datasets. While similar concerns regarding frequency information also apply to Link Density, the current results are promising and are likely to improve with models that incorporate temporal information.

To the best of our knowledge, this is the first study to use LUT-based models for predicting acoustic indices that capture acoustic properties beyond noise at a citywide scale. Therefore, no comparable results for the specific indices used here are available. However, Clark et al. [46] used a LUT-based random forest approach to model the presence of different sound sources (e.g., traffic, animal) in Accra, Ghana. As model performance measures, the report r values range from 0.01 (Nature) to 0.72 (Animal and insects), with the majority of values being around 0.5.

Strengths and limitations

One of the major strengths of this study is the demonstration that models solely based on readily available LUT information represent a promising approach for predicting selected AE properties (here, LAeq and Link Density). Once the necessary LUT features are calculated and the model is built, this enables a fast application for predicting the respective AE properties on high spatial resolution. For example, Fig. 3 represents the application of the SALVE_{All} model to predict the LAeq and the Link Density for the research area of three cities (~450 km²). The calculation of LUT features took approximately 48 h, while the computation time for predicting each of the respective indices took less than 10 s on a standard office PC (using an 11th Gen Intel(R) Core(TM) i9-11900K at 3.50 GHz and 64 GB of RAM with no parallelization). In theory, these models can easily be applied to predict further acoustic properties of interest. Additional strengths are that we compare our results to the performance of the SNM using a comprehensive dataset of acoustic measurements across the city of Bochum from the SALVE project, as well as the model evaluation, using two independent datasets.

However, limitations of this study comprise the already mentioned neglected temporal information, which might be especially important to predict frequency-related AE properties. In addition, we only focus on the daytime AE, as measurements are only available between 09:00 and 19:30. As the night-time AE plays an important role for health-related issues like sleep [1], our models may not be suitable for accurately estimating exposure levels during this time frame. In addition, although two independent datasets were used for the evaluation of the model performance, the Be-MoVe dataset only comprises 22 recording locations. Another important limitation is that the model is only based on the relations between LUTs in the surroundings of acoustic measurements in Bochum; thus, potentially important LUTs (e.g., airports) are not considered in predicting AE properties. Therefore, additional studies of larger sample sizes and from different cities are needed to prove the scalability of LUT-based approaches—especially for the prediction of acoustic indices

beyond noise. Methodologically, we use a gradient boosting model, whose predictions are bound between the minimum and maximum values provided in the training data. While for indices like the Link Density or the Articulation Index, the empirical measures are close to the theoretical bandwidth, the model will fail with predictions for locations with higher or lower values for the other indices. Furthermore, although we performed sensitivity analysis using extreme gradient boosting [47] without observing performance improvements, other models may outperform gradient boosting in the future.

OUTLOOK

In this work, we demonstrate the predictive power of solely LUT-based models to estimate properties of the urban AE at high spatial resolution. As the model for the LAeq outperforms the SNM in predicting total environmental noise and similar performance was found throughout the literature, it represents a promising approach in estimating total environmental noise at high spatial resolution. Although the predictive power for the Link Density model was lower than that of the LAeq model, its results were also robust across datasets. Our results are particularly important for epidemiological studies, as the models offer fast, large-scale estimates, using readily available information. In the future, these models could be used to estimate exposure to AE properties, which can then be analysed in relation to human health, enabling the investigation of associations already observed in laboratory and field studies within population-based research. From a spatial planning perspective, processes that redefine or update land use mapping can leverage our model to account for the impacts on the acoustic environment. While planning measures aimed at altering existing land use types are often gradual and spatially constrained, structural transformations like the repurposing and qualification of former industrial brownfield sites create strategic opportunities for targeted redefinitions of land use. Furthermore, future work should investigate whether the inclusion of temporal information along with available environmental data (e.g., satellite imagery) will enhance the predictive power.

DATA AVAILABILITY

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. European Environment Agency. Environmental noise in Europe—2025. Luxembourg: Publications Office of the European Union, 2025.
2. DIN ISO 12913-1:2018-02. Acoustics—Soundscape—Part 1: definition and conceptual framework (ISO 12913-1:2014).

3. Schulte-Fortkamp B, Fiebig A, Sisneros JA, Popper AN, Fay RR. Soundscapes: humans and their acoustic environment. Springer International Publishing: Cham, Switzerland.
4. Medvedev O, Shepherd D, Hautus MJ. The restorative potential of soundscapes: a physiological investigation. *Appl Acoust.* 2015;96:20–6.
5. Annerstedt M, Jönsson P, Wallergård M, Johansson G, Karlson B, Grahn P, et al. Inducing physiological stress recovery with sounds of nature in a virtual reality forest—results from a pilot study. *Physiol Behav.* 2013;118:240–50.
6. Stobbe E, Forlim CG, Kühn S. Impact of exposure to natural versus urban soundscapes on brain functional connectivity, BOLD entropy and behavior. *Environ Res.* 2024;244:117788.
7. Buxton RT, Pearson AL, Allou C, Fristrup K, Wittemyer G. A synthesis of health benefits of natural sounds and their distribution in national parks. *Proc Natl Acad Sci USA.* 2021;118:e2013097118.
8. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Wolters Kluwer Health/Lippincott Williams & Wilkins; Philadelphia: 2008.
9. Directive E Directive 2002/49/EC of the European Parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise. *Off J Eur Communities.* 2002;189:12–25.
10. Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos Environ.* 2008;42:7561–78.
11. Aguilera I, Foraster M, Basagaña X, Corradi E, Deltell A, Morelli X, et al. Application of land use regression modelling to assess the spatial distribution of road traffic noise in three European cities. *J Expo Sci Environ Epidemiol.* 2015;25:97–105.
12. Alvares-Sanches T, Osborne P, White, P. The challenges of mapping quiet urban areas from mobile sound surveys. In: INTER-NOISE and NOISE-CON Congress and Conference Proceedings: Institute of Noise Control Engineering. 2024;270:5503–11.
13. Gharehchahi E, Hashemi H, Yunesian M, Samaei M, Azhdarpoor A, Oliaei M, et al. Geospatial analysis for environmental noise mapping: a land use regression approach in a metropolitan city. *Environ Res.* 2024;257:119375.
14. Liu Y, Goudreau S, Oiamo T, Rainham D, Hatzopoulou M, Chen H, et al. Comparison of land use regression and random forests models on estimating noise levels in five Canadian cities. *Environ Pollut.* 2020;256:113367.
15. Ragettli MS, Goudreau S, Plante C, Fournier M, Hatzopoulou M, Perron S, et al. Statistical modeling of the spatial variability of environmental noise levels in Montreal, Canada, using noise measurements and land use characteristics. *J Expo Sci Environ Epidemiol.* 2016;26:597–605.
16. Razavi-Termeh SV, Abolghasem S-N, Mohammadreza J-N, Choi S-M. Exploring multi-pollution variability in the urban environment: geospatial AI-driven modeling of air and noise. *Int J Digit Earth.* 2024;17:2378819.
17. Staab J, Schady A, Weigand M, Lakes T, Taubenböck H. Predicting traffic noise using land-use regression—a scalable approach. *J Expo Sci Environ Epidemiol.* 2022;32:232–43.
18. Wood LA, Yeager R, Guinn B, Taylor KC, Gaskins J, Loehr M, et al. Land-Use Regression Estimation of Cumulative Environmental Noise Exposure in Jefferson County, Kentucky. International Society for Environmental Epidemiology (ISEE) Conference Abstracts. 2021.
19. Xie D, Liu Y, Chen J. Mapping urban environmental noise: a land use regression method. *Environ Sci Technol.* 2011;45:7358–64.
20. Zheng G, Chen X, Huang K, Mölter A, Liu M, Zhou B, et al. Mapping environmental noise of Guangzhou based on land use regression models. *J Environ Manag.* 2025;373:123931.
21. Kumar S, Garg N. Integrated land use regression (iLUR) model with road traffic characteristics for environmental noise prediction and mapping in urban regions with heterogeneous traffic conditions. *Appl Acoust.* 2025;239:110830.
22. ANSI S3. 5-1997. Methods for the calculation of the speech intelligibility index. New York: American National Standards Institute; 1997.
23. Boelman NT, Asner GP, Hart PJ, Martin RE. Multi-trophic invasion resistance in Hawaii: bioacoustics, field surveys, and airborne remote sensing. *Ecol Appl.* 2007;17:2137–44.
24. Haselhoff T, Braun T, Fiebig A, Hornberg J, Lawrence BT, Marwan N, et al. Complex networks for analyzing the urban acoustic environment. *Ecol Inform.* 2023;78:102326.
25. DIN 45692-2009. Measurement Technique for the Simulation of the Auditory Sensation of Sharpness. Deutsches Institut für Normung; Berlin: 2009.
26. Lawrence BT, Hornberg J, Schröder K, Djeudeu D, Haselhoff T, Ahmed S, et al. Linking ecoacoustic indices to psychoacoustic perception of the urban acoustic environment. *Ecol Indic.* 2023;155:111023.
27. Haselhoff T, Schuck M, Lawrence BT, Fiebig A, Moebus S. Characterizing acoustic dimensions of health-related urban greenspace. *Ecol Indic.* 2024;166:112547.
28. Haselhoff T, Hornberg J, Fischer JL, Lawrence BT, Ahmed S, Gruehn D, et al. The acoustic environment before and during the SARS-CoV-2 lockdown in a major German city as measured by ecoacoustic indices. *J Acoust Soc Am.* 2022;152:1192–200.
29. Haselhoff T, Lawrence B, Hornberg J, Ahmed S, Sutcliffe R, Gruehn D, et al. The Acoustic quality and health in urban environments (SALVE) project: Study design, rationale and methodology. *Appl Acoust.* 2022;188:108538.
30. Hornberg J, Hemker F, Schröder K, Hinse M, Moebus S, Schröder J. Association between perceived sound type dominance and overall assessment of the acoustic environment using ISO 12913 soundwalks. *J Acoust Soc Am.* 2024;156:2827–37.
31. NTi Audio. Technische Daten Messmikrofone. [Available from: <https://www.nti-audio.com/Portals/0/data/de/Messmikrofone-Spezifikationen.pdf>, 2026].
32. Lawrence BT, Hornberg J, Haselhoff T, Sutcliffe R, Ahmed S, Moebus S, et al. A widened array of metrics (WAM) approach to characterize the urban acoustic environment; a case comparison of urban mixed-use and forest. *Appl Acoust.* 2022;185:108387.
33. Almansi KY, Ujang U, Azri S, Wickramathilaka N. Traffic noise prediction model using GIS and ensemble machine learning: a case study at Universiti Teknologi Malaysia (UTM) Campus. *Environ Sci Pollut Res.* 2024;31:60905–26.
34. Vrigazova B. The proportion for splitting data into training and test sets for the bootstrap in classification problems. *Bus Syst Res J.* 2021;12:228–42.
35. Regionalverband Ruhr. Flächennutzungskartierung. Daten für die Stadt- und Regionalplanung. [Available from: <https://www.rvr.ruhr/daten-digitales/geodaten/flaechennutzungskartierung/>, 2026].
36. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw.* 2010;36:1–13.
37. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat.* 2001;29:1189–232.
38. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* 30:2017.
39. Oliver MA, Webster R. Kriging: a method of interpolation for geographical information systems. *Int J Geogr Inf Syst.* 1990;4:313–32.
40. Gupta A, Stead TS, Ganti L. Determining a meaningful R-squared value in clinical medicine. *Acad Med Surg.* 2024. <https://doi.org/10.62186/001c.125154>.
41. Tang J-H, Lin B-C, Hwang J-S, Chen L-J, Wu B-S, Jian H-L, et al. Dynamic modeling for noise mapping in urban areas. *Environ Impact Assess Rev.* 2022;97:106864.
42. Wei W, Van Renterghem T, De Coensel B, Botteldooren D. Dynamic noise mapping: a map-based interpolation between noise measurements with high temporal resolution. *Appl Acoust.* 2016;101:127–40.
43. Helbich M, Hagenauer J, Burov A, Dzhambov AM. Traffic noise assessment in urban Bulgaria using explainable machine learning. *Sustain Cities Soc.* 2025;120:106169.
44. Huang Y-K. Improving noise exposure assessment for epidemiological studies. Doctoral dissertation, University of Illinois at Chicago: 2019.
45. Fiebig A, Templiner J, Haselhoff T, Moebus S. Psychoakustische Analysen von Umgebungsgeräuschen in einer Langzeitperspektive. DAGA 2023; Hamburg: Deutsche Gesellschaft für Akustik eV (DEGA); 2023.
46. Clark SN, Arku RE, Ezzati M, Bennett J, Nathvani R, Alli AS, et al. Moving beyond the noise: geospatial modelling of urban sound environments in a sub-Saharan African city. *Sci Rep.* 2025;15:21403.
47. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: Association for Computing Machinery; 2016. pp 785–94.

ACKNOWLEDGEMENTS

We would like to thank the “Amt für Stadtplanung und Wohnen” of the city of Bochum for providing access to the strategic noise map of Bochum. Their support and cooperation were instrumental in enabling the spatial analyses carried out in this study. We also acknowledge the support of the Open Access Publication Fund of the University of Duisburg-Essen.

AUTHOR CONTRIBUTIONS

Conceptualization: TH, SM, MJ, BL, and FW; Methodology: TH, MJ, and FW; Software: TH; Validation: TH; Formal analysis: TH; Investigation: TH; Resources: TH, SM, and FW; Data curation: TH, SM, MJ, and BL; Writing—original draft preparation: TH; Writing—review and editing: SM, MJ, BL and FW; Visualization: TH; Supervision: SM and FW; Project administration: SM and FW.

FUNDING

This work was supported by the HEAD-Genuit-Foundation [P-21/02- W]. The funding source had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICAL APPROVAL

The ethics committee of the Medical Faculty of the University of Duisburg-Essen evaluated the SALVE and Be-MoVe studies and raised no ethical issues.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41370-026-00855-w>.

Correspondence and requests for materials should be addressed to Timo Haselhoff.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026