*buildings*

MDPI

*Article*

# Evaluation of Psychoacoustic Machine Learning Assessment Method for Predicting Occupant Well-Being

**Kuen Wai Ma [1], Cheuk Ming Mak [2,*], Fu-Lai Chung [3] and Hai Ming Wong [4]**

[1] Department of Mechanical Engineering, National Taipei University of Technology, Taipei 10608, Taiwan; kwma@ntut.edu.tw
[2] Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[3] Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China; korris.chung@polyu.edu.hk
[4] Faculty of Dentistry, The University of Hong Kong, Pok Fu Lam, Hong Kong Island, Hong Kong, China; wonghmg@hku.hk
* Correspondence: cheuk-ming.mak@polyu.edu.hk; Tel.: +852-2766-5856

**Abstract**

In modern buildings, the air-conditioned indoor environment is vital for occupant productivity and well-being, yet fan noise and airflow turbulence can significantly compromise these benefits. Human–environmental interactions are complex processes that traditional energy-based acoustic metrics are often insufficient to model. Therefore, this study aims to advance the multidimensional sound quality assessment framework for building acoustics. Three methods, the conventional regression approach (CRA), general prediction model (GPM), and psychoacoustic machine learning (PML) assessment methods, were evaluated for predicting three perceptual dimensions (*Evaluation*, *Potency*, *Activity*; *EPA*) and negative noise impacts on occupant well-being (*O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*). Based on 432 multidimensional sound quality assessments across four general types of air-conditioned built environments, the PML achieved the best goodness-of-fit for the *EPA*-score perdition (adjusted $R^2$ = 0.61) compared to CRA (0.32) and GPM (0.15) and effectively predicted all negative noise impacts (adjusted $R^2$ = 0.53–0.61). The PML assessment method offers a smart and reliable solution for sound quality and well-being prediction through psychoacoustic heatmaps encoding time-varying psychoacoustic features in 227 × 227 pixels from 30 s soundtracks of the built environment for sustainable building design.

**Keywords:** air-conditioned built environment; building acoustics; multidimensional sound quality assessment; noise impact prediction; occupant well-being; perceptual dimensions of sounds; psychoacoustic heatmap; psychoacoustic machine learning; sustainable design

Check for updates

Academic Editor: Ángel Fermín Ramos Ridao

Received: 7 February 2026
Revised: 1 March 2026
Accepted: 3 March 2026
Published: 5 March 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

The field of building acoustics [1,2] encompasses the assessment and control of noise within built environments, with particular emphasis on how environmental sound characteristics affect the well-being of building occupants [3]. Modern buildings increasingly rely on air-conditioning systems as integral components for maintaining satisfactory indoor environmental conditions [4]. These systems inevitably generate noise through multiple mechanisms, including mechanical fan operation and aerodynamic interactions where turbulent airflow encounters duct discontinuities [5,6]. Consequently, systematic acoustic

assessment has emerged as a critical requirement for mitigating these issues [7]. Conventional acoustic assessment approaches [8,9] for air-conditioned built environments utilize established acoustic metrics including A-weighted equivalent continuous sound pressure level ($L_{Aeq}$), noise criteria (*NC*), noise rating (*NR*), and room criteria (*RC*). While these single-value metrics offer straightforward acceptability assessments of indoor noise conditions, they possess inherent limitations in characterizing the acoustic properties and complex psychological dimensions of occupant experience that extend beyond simple magnitude sensitivity [10]. Historically, noise assessment, prevention, prediction methodologies, and management frameworks have prioritized objective acoustic characterization, largely because objective metrics present fewer measurement and prediction challenges compared to subjective human responses [11]. Contemporary understanding emphasizes that human–environmental interactions constitute the foundation for advancing acoustic assessment in building acoustics. This recognition establishes multidimensional sound quality assessment [12] as fundamental for effectively analyzing human–environmental relationships and supporting evidence-based building design.

Public recognition of the health consequences [13] of daily noise exposure has increased following the update of the World Health Organization (WHO) noise guidelines in 2018 [14]. Daily noise exposure is linked to discomfort as well as stress biomarkers, disease progression, and mortality [15]. Undesired acoustic environments are recognized to cause stress [16], annoyance [17], and discomfort [18], all of which exert negative impacts on occupant well-being [19]. Human–environmental interactions with acoustic environments include sequential processes from sound generation and transmission to perception, ultimately concluding in diverse noise impact outcomes. Although ISO 15,666 [20] and ISO 12,913 standards [21–23] address subjective assessment of perceptual responses, these assessments exhibit substantial variability among building acoustics studies [24,25]. The semantic differential method [26], which employs questions constructed from bipolar semantic pairs representing opposing meanings [27,28], is a commonly adopted psychological scale for the quantitative evaluation of objective phenomena. A comprehensive systematic review [29] analyzing applications of the semantic differential method for subjective assessment discovered the fundamental structure of sound perception in terms of the *Evaluation*, *Potency*, and *Activity* (*EPA*) framework. The *Evaluation* (*E*-), *Potency* (*P*-), and *Activity* (*A*-) dimensions respectively represent general judgment, sensation of sound energy content, and sensation of the temporal and spectral content of sounds. Hence, acoustic metric selection for sound quality assessment should incorporate measurements addressing energy, spectral, and temporal characteristics of sound.

In objective acoustic characterization, psychoacoustic metrics comprise a collection of objective measures designed to quantify actual sound sensations, grounded in the psychoacoustic scale introduced in 1961 [30]. Total loudness (*N*) is the most widely recognized psychoacoustic metric, quantifying loudness perception with calculation procedures standardized in ISO 532-1 [31]. The *N* measurement provides valuable supplementation to conventional acoustic assessment methods for assessing indoor acoustic comfort [32] compared to *NC* and *NR*. In addition, sharpness (*S*), roughness (*R*), and fluctuation strength (*FS*) are usually accepted psychoacoustic metrics for evaluating the spectral properties of sounds. These metrics respectively quantify sharpness perception via energy skewness calculations, roughness perception through rapid amplitude modulation measurement, and fluctuation strength perception via slower amplitude modulation assessment [33]. Applying statistical sound level metrics such as $L_{A90}$ and $L_{A10}$, rather than time-equivalent levels, enables superior differentiation between ambient noise levels and event noise levels through consideration of temporal characteristics. These psychoacoustic metrics and statis-

tical noise levels were also found to be significantly correlated with perceptual assessments across the *E*-, *P*-, and *A*-dimensions [34,35].

Contrasting with objective acoustic characterization, developing reliable, valid, and practical psychometric instruments for assessing multidimensional perceptual impacts remains an ongoing research priority. Various subjective assessment approaches, together with their theoretical factor models, have been proposed across different investigations [36]. However, analytical inconsistencies among these studies limit cross-study comparability. The Psychoacoustics Perception Scale (PPS) was therefore constructed based on the *EPA* framework to systematically evaluate all essential perceptions within the fundamental human perceptual dimensions of sound [12]. Identification of the *EPA* framework offers a promising approach for addressing challenges posed by extensive factor variability in multidimensional environmental sound quality assessment. Moreover, the *EPA* framework establishes a comparative structure across different studies and provides a basis for predicting negative noise impacts associated with occupant well-being. Recent research [37] has successfully applied multidimensional sound quality assessments to outdoor soundscapes by combining a comprehensive set of acoustic and psychoacoustic metrics with the PPS. This study evaluated the occupant well-being-related noise impacts (e.g., *O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*), which the WHO links to the disease burden of environmental noise. The results confirmed the value of the *EPA* framework in multidimensional sound quality assessments and facilitated the development of general prediction models (GPMs) that predictively model perceptual influences on occupant well-being.

Other than the conventional regression approach (CRA), the development of artificial neural networks (ANNs) with feature learning capabilities represents a significant advancement in machine learning (ML) algorithms for predictive modeling of perceptual influences on occupant well-being. While ANNs inspired by biological neural networks [38] have revolutionized fields like image classification [39], particularly through non-saturating ReLU activation functions [40], their application in psychoacoustics remains limited. A systematic search of the ScienceDirect database revealed that most "psychoacoustic machine learning" (PML) studies focused on objective signal classification tasks such as heartbeat analysis [41], vehicle diagnostics [42], and sound categorization [43–45], rather than on subjective human impact. Only a few studies have applied PML to predict subjective noise impacts [46–48], highlighting a gap in intelligent, holistic methodologies for assessing air-conditioned built environments.

To address the limitations of CRA assessment methods in capturing temporal variations and spectral complexities, a prior study [49] developed the psychoacoustic heatmap machine learning model (PHMLM). This innovative method incorporated the EPA framework to transform dynamic psychoacoustic metrics into detailed visual formats for forecasting noise effects. Unlike classical Mel-spectrograms [50], which display energy in constant frequency bands without accounting for human auditory sensations such as loudness and sharpness, the PHMLM utilized a preprocessing technique that reduced input data size by approximately 5000-fold. By encoding multidimensional acoustic information into 227 × 227-pixel psychoacoustic heatmaps with 0.002 s temporal resolution, the PHMLM efficiently captured energy, spectral, and temporal content for deep neural network analysis of multidimensional sound quality. This prior work [49] demonstrated that PML assessment methods could achieve predictive consistency with established regression methodologies (including GPM and CRA), providing a foundation for the comparative evaluation presented in this study.

The primary objective of this research is to evaluate different assessment methods (CRA, GPM, and PML) for predicting sound quality and subsequent negative noise impacts

on occupant well-being in air-conditioned built environments using multidimensional objective acoustic characteristics. Detailed model development procedures have not been repeated in Section 2, as the historical progression of different methods used to analyze the complex interaction between the acoustic environment and occupant perception has been comprehensively presented in prior published studies (CRA [12,23], GPM [37], and PML [49]). Compared to these prior studies, this work delivers a comprehensive performance evaluation, cross-method comparison, and hybrid integration analysis through balanced sampling of air-conditioned environments, independent subjective perceptual response measurements beyond the original development datasets, and cross-domain generalization testing (outdoor GPM with indoor PML). These advances identify the relative strengths of each method and optimal integration strategies for building acoustic design. This significance of this research lies in delivering a smart, reliable, efficient, and user-friendly assessment method that can inform sustainable building design and enhance indoor acoustic environments by predicting occupant well-being responses to air-conditioning noise within built environments.

## 2. Materials and Methods

### *2.1. Multidimensional Sound Quality Assessment Framework*

A holistic approach to sound quality assessment [12,37,49] is essential for capturing the complex interaction between the acoustic environment and occupant perception. As illustrated in Figure 1a, this study adopts a comprehensive framework comprising four key stages I-IV for objective and subjective data collection:

- Stage I. Objective acoustic characterization: This stage involves measuring the physical properties of sound using acoustic and psychoacoustic metrics (see Figure 1b) that capture energy content (e.g., $L_{Aeq}$, *RC*, *N*), spectral content (e.g., *S*, *R*, and *FS*), and temporal content (e.g., $L_{A10}$–$L_{A90}$ and $N_5$–$N_{95}$);
- Stage II. Fundamental perceptual responses: This stage measures the fundamental subjective perceptual responses to sound using the PPS, which consists of nine questions (*E1*: *Quiet–Noisy*, *E2*: *Relaxed–Tense*, *E3*: *Pleasant–Unpleasant*, *P1*: *Quiet–Loud*, *P2*: *Light–Heavy*, *P3*: *Weak–Strong*, *A1*: *Deep–Metallic*, *A2*: *Low–High*, and *A3*: *Dull–Sharp*) on a 7-level semantic differential scale (see Figure 1c);
- Stage III. *EPA* framework: This stage categorizes subjective responses into three fundamental dimensions (*E*, *P*, and *A*), serving as the core link in the human–environmental interaction process between physical sound stimuli and occupant perception (see Figure 1c);
- Stage IV. Occupant well-being-related noise impacts: The final stage evaluates the consequently negative impacts on occupant well-being (see Figure 1c), specifically *Discomfort* (*O1*), *Annoying* (*O2*), *Stressful* (*O3*), and *Unacceptable* (*O4*).

This multidimensional structure provides a standardized ground-truth for comparative analysis, ensuring that the CRA [12,23], GPM [37], and PML [49] assessment methods are evaluated against a consistent and holistic set of inputs from objective acoustic characterization to predict sound quality and subsequent negative noise impacts on occupant well-being (arrows in Figure 1a).

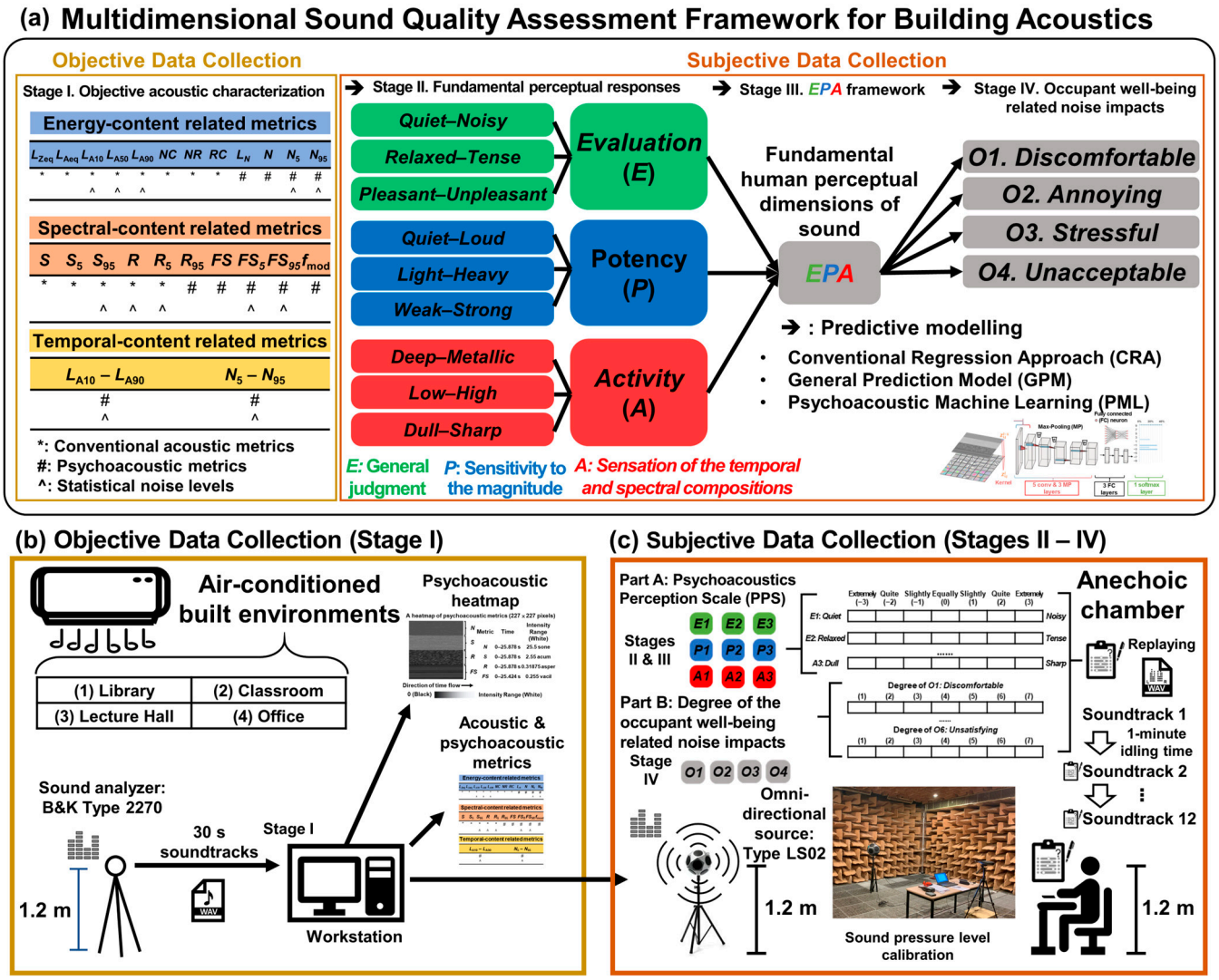



**Figure 1.** The multidimensional sound quality assessment framework and experimental workflow for the comparative evaluation of conventional regression approach (CRA), general prediction model (GPM), and psychoacoustic machine learning (PML) assessment methods in predicting occupant well-being in air-conditioned built environments. (**a**) The theoretical framework comprising four key stages (I–IV) linking objective acoustic characteristics to occupant well-being; (**b**) Experimental setup of objective data collection in Stage I, illustrating objective data collection in four air-conditioned built environments; (**c**) Experimental setup of subjective data collection in Stages II-IV, detailing the subjective jury listening tests and questionnaire structure (Part A: Psychoacoustic Perception Scale and Part B: occupant well-being-related noise impacts).

#### 2.1.1. Objective Data Collection in Air-Conditioned Built Environments

To capture a representative sample of indoor acoustic conditions, sound recordings were conducted in four distinct air-conditioned built environments: (1) Library, (2) Classroom, (3) Lecture Hall, and (4) Office (see Figure 1b). These four air-conditioned built environments (Library, Classroom, Lecture Hall, and Office) represent essential daily spaces where sustained HVAC noise exposure critically impacts occupant concentration, learning, communication, and productivity. Libraries demand quiet for individual study and reading, classrooms support interactive teaching and group discussions, lecture halls accommodate large-group instruction and presentations, and offices facilitate collaborative work and focused tasks, each presenting distinct psychoacoustic challenges from HVAC systems. Their selection reflects prevalent built environments in educational institutions

and workplaces where air-conditioning noise management directly influences occupant well-being and performance. These space types encompass the majority of indoor scenarios encountered daily by students, faculty, and professionals. Detailed acoustic characteristics of these environments were analyzed and compared in Section 3 to better understand the differences between environments. Soundtracks with an equal number from different air-conditioned built environments were randomly selected in this study from the prior PML development study [49]. Although identical to prior study recordings, independent listening tests were conducted to collect subjective perceptual responses, eliminating data leakage between the training phase of the prior study and the multidimensional sound quality assessments in this study. All 30 s soundtracks were recorded using a high-precision two-channel handheld analyzer (Type 2270; Brüel & Kjær, Nærum, Denmark) equipped with binaural recording microphones. The recording position was standardized at a height of 1.2 m above the floor to simulate the typical ear height of a seated occupant. A total of 432 30 s soundtracks were extracted from the recordings for both objective and subjective analysis of noise in air-conditioned built environments. Conventional acoustic metrics were recorded directly from the sound analyzer, while psychoacoustic metrics were computed using MATLAB R2024a (MathWorks, Natick, USA). The detailed calculation equations and unit definitions based on ISO 532-1 [31] and Zwicker's book [30] are provided in Appendix A.

#### 2.1.2. Subjective Perceptual Response Measurement

Jury listening tests were conducted to collect subjective responses to the recorded soundtracks (see Figure 1b). The tests were conducted in an anechoic chamber at the Hong Kong Polytechnic University to eliminate the influence of background noise and room reflections. To minimize the influence of confounding environmental factors, the anechoic chamber maintained a controlled temperature of $25 \pm 2$ °C with adequate lighting conditions. The 30 s soundtracks were replayed using an omni-directional dodecahedron loudspeaker (Type LS02; Acoem, Limonest, France) positioned 1.2 m from the listener, matching the original recording height. To ensure accurate reproduction, the sound pressure level (SPL) was calibrated using a 1 kHz reference tone at 60 dB measured at the listener's ear position. Before the listening tests, written informed consent was obtained from all participants, and demographic details (gender and age) were documented. Individuals with self-reported hearing impairments or chronic noise exposure were excluded to ensure data validity. This criterion minimized potential biases from undetected auditory threshold shifts that could confound perceptual judgments.

Participants evaluated a total of 12 soundtracks using a self-administered questionnaire structured in two parts to capture the subjective responses defined in the assessment framework (see Figure 1c). All 12 soundtracks per participant were randomly assigned. Individual participant sets were not environment-balanced (due to randomization), but collectively ensured the same number of assessments per environment type. A 1 min idling period separated consecutive assessments. Part A employed the PPS to quantify sound quality perceptions across the three fundamental *E*-, *P*-, and *A*- dimensions. Part B assessed occupant well-being-related noise impact, where participants rated the specific negative outcomes (*O1*–*O4*) on a 7-point Likert scale. This dual-part instrument ensured that both the intermediate perceptual attributes and the consequential well-being impacts were systematically recorded for every acoustic stimulus.

### *2.2. Assessment Methods for Predictive Modeling of Subjective Responses*

Three assessment methods (CRA, GPM, and PML) were applied and compared in this study to evaluate the feasibility of predicting occupant well-being.

2.2.1. Conventional Regression Approach (CRA)

The CRA represents the traditional methodology in environmental noise studies [23] involving subjective data. The objective acoustic and psychoacoustic metrics are directly used as independent variables in linear regression models (see Equation (1)) to predict subjective responses such as *EPA*-scores in the *EPA* framework.

$$y = B_1x_1 + B_2x_2 + B_3x_3 + \cdots B_nx_n + C, \quad (1)$$

where y is the dependent variable, $B_1$–$B_n$ are the regression coefficients (unstandardized coefficients) for the independent variables ($x_1$–$x_n$) as predictors, and $C$ is the y-intercept.

2.2.2. General Prediction Model (GPM)

The fundamental distinction between the GPM and the CRA lies in their reliance on experimental data. While the CRA fits a regression model to the specific participant responses observed in this study, the GPM operates as a standalone predictive tool. In this context, the GPM is the "pure-objective" approach developed for occupant well-being prediction [37]. It derives predictions directly from acoustic and psychoacoustic metrics using established formulations, without further adjustment to the subjective data collected in this specific experiment.

By bypassing the need for dataset-specific correction, the GPM evaluates the generalizability and direct applicability of using established objective equations to predict subjective *EPA*-score ($GPM_{EPA}$, Equation (2)) and occupant well-being-related noise impacts ($GPA_{O1–4}$, Equations (3)–(6)). Although the GPM was originally formulated for broader applications in outdoor soundscape conditions, this study specifically applied the GPM to test its ability in indoor air-conditioned built environments.

$$GPM_{EPA} = 1.03 \times (L_{A10} - L_{A90}) + 0.68 \times L_{A50} - 0.28 \times N - 30.4 \times R - 39.7, \quad (2)$$

$$GPM_{O1:\ discomfortable} = 0.17 \times GPM_{EPA} - 5.51 \times FS + 3.7 \times R_5 + 0.04 \times L_{Zeq} - 0.03 \times L_{A90} + 2.06, \quad (3)$$

$$GPM_{O2:\ Annoying} = 0.18 \times GPM_{EPA} - 5.86 \times FS - 0.04 \times f_{Mod} + 4.77, \quad (4)$$

$$GPM_{O3:\ Stressful} = 0.18 \times GPM_{EPA} + 3.86, \quad (5)$$

$$GPM_{O4:\ Unacceptable} = 0.17 \times GPM_{EPA} - 5.58 \times S + 8.81 \times R_5 + 3.77. \quad (6)$$

2.2.3. Psychoacoustic Machine Learning (PML) Assessment Method

The third and most advanced approach evaluated in this study is the PML assessment method, which utilizes the PHMLM [49], originally developed for *EPA*-score prediction. Unlike the equation-based GPM, the PML is a data-driven framework designed to capture complex, non-linear relationships between acoustic features and subjective perception.

For each 30 s soundtrack containing air-conditioning sounds, a corresponding psychoacoustic heatmap was generated as a 227 × 227-pixel grayscale image (see Figure 1b). These dimensions align with the input requirements of the AlexNet architecture [39] (see Table 1). Detailed psychoacoustic heatmap generation procedures, including matrix selection and resolution rationale, can be found in the study of PML development [49]. The heatmap encodes temporal variations in the four most commonly applied psychoacoustic metrics in environmental noise studies (*N*, *R*, *S*, and *FS*), calculated with a time resolution of 0.002 s. This resolution utilizes 96 samples at a 48 kHz sample rate, following ISO 532-1 standards. The specific covering durations and maximum intensities mapped to the pixel value of 255 (white) were set as follows:

- Rows 1–57: *N* covering 25.878 s, with a maximum intensity of 25.5 sone (0.1 sone/pixel);
- Rows 58–114: *S* covering 25.878 s, with a maximum intensity of 2.55 acum (0.01 acum/pixel);

- Rows 115–171: *R* covering 25.878 s, with a maximum intensity of 0.31875 asper (0.00125 asper/pixel);
- Rows 172–227: *FS* covering 25.424 s, with a maximum intensity of 0.255 vacil (0.001 vacil/pixel).

**Table 1.** Architecture of a psychoacoustic heatmap machine learning model.

| Layer | Configuration/Dimensions | Activation/Function |
|---|---|---|
| Input | 227 × 227 × 1 (grayscale heatmap) | Input layer for psychoacoustic heatmaps |
| Convolution 1 | 96 filters, 11 × 11 kernel, stride 4 | ReLU (Feature extraction—transferred) |
| Max Pooling 1 | 3 × 3 max pooling, stride 2 | Down-sampling |
| Convolution 2 | 256 filters, 5 × 5 kernel, padding 2 | ReLU (Feature extraction—transferred) |
| Max Pooling 2 | 3 × 3 max pooling, stride 2 | Down-sampling |
| Convolution 3 | 384 filters, 3 × 3 kernel, padding 1 | ReLU (Feature extraction—transferred) |
| Convolution 4 | 384 filters, 3 × 3 kernel, padding 1 | ReLU (Feature extraction—transferred) |
| Convolution 5 | 256 filters, 3 × 3 kernel, padding 1 | ReLU (Feature extraction—transferred) |
| Max Pooling 3 | 3 × 3 max pooling, stride 2 | Down-sampling |
| Fully Connected 6 | 4096 neurons | ReLU, Dropout (0.5) (High-level feature processing) |
| Fully Connected 7 | 4096 neurons | ReLU, Dropout (0.5) (High-level feature processing) |
| Fully Connected 8 | 19 neurons | Softmax (classification for −9 to +9 scale) |
| Output | 19-level probability distribution | *E*-/*P*-/*A*-score prediction |

Notes. ReLU = Rectified Linear Unit activation function; Dropout = regularization technique to prevent overfitting; *E* = *Evaluation*; *P* = *Potency*; *A* = *Activity*. The final predicted *EPA*-score is derived by combining the separate outputs of three parallel models independently trained for *E*-, *P*, and *A*-scores.

For this comparative study, the PML model was deployed using the pre-trained PHMLM-EPA established in prior work [49]. As the training phase (including hyperparameter optimization and stochastic gradient descent with momentum (SGDM) solver configuration) was completed and validated in the prior study, this research utilizes the fully trained network directly to generate predictions for the current dataset. This ensures that the comparison between PML, GPM, and CRA focuses strictly on predictive performance.

### *2.3. Statistical Analysis*

All statistical analysis were executed using IBM SPSS Statistics 26 (IBM Corp., Armonk, NY, USA) at a significance level of 0.05 [51]. As a preliminary step, the Shapiro–Wilk test [52] was applied to assess the normality of the dataset. Given that the acoustic metrics and subjective responses demonstrated non-normal distributions ($p$-values < 0.05), non-parametric statistical methods were adopted for the two-tailed statistical tests in this study. To characterize the acoustic variations within the study, the Kruskal–Wallis test was employed to evaluate statistically significant differences in acoustic characteristics and subjective responses across the different types of air-conditioned built environments. Upon the detection of significant variances ($p$-values < 0.05), Dunn's post hoc test with Bonferroni adjustments was performed to identify specific pairwise differences between different built environments. The internal consistency of the semantic differential scales was validated using Cronbach's α (> 0.60) [53], justifying the reliability of combining individual items into composite *E*-, *P*-, and *A*-scores.

The comparative evaluation of the predictive modeling approaches (CRA, GPM, and PML) focused on two key performance indicators: Spearman's rank correlation coefficient ($\rho$) and the adjusted coefficient of determination (adjusted $R^2$). Spearman's $\rho$ was applied to test the strength and direction of the monotonic relationships (from −1 to 1) between the model-predicted values and the ground-truth experimental data, both for *EPA*-score and occupant well-being-related noise impacts (O1–O4).

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}, \tag{7}$$

where $n$ is the total sample size and $d_i$ is the difference between the two ranks of each sample.

The adjusted $R^2$ served as a metric for goodness-of-fit, quantifying the proportion of variance in the subjective responses that could be explained by each model while penalizing model complexity. A higher adjusted $R^2$ value indicates superior explanatory power, allowing comparison between the simpler regression-based approaches (CRA and GPM) and the more complex PML assessment methods.

$$Adjusted\ R^2 = 1 - \left(1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}\right) \times \frac{(n-1)}{(n-p-1)}, \tag{8}$$

where $n$ is the total sample size, $p$ is the number of predictors, and $\hat{y}_i$ and $\overline{y}$ are the predicted values and the mean of y, respectively.

## 3. Results

### *3.1. Descriptive Statistics of Multidimensional Sound Quality Assessment*

A total of 432 jury listening tests were conducted for analysis, derived from 36 participants (69.4% male, 30.6% female; mean age 24.1 ± 2.2 years) evaluating 12 soundtracks each. The number of sound quality assessment in this study was balanced across four different types of air-conditioned built environments: (1) Library, (2) Classroom, (3) Lecture Hall, and (4) Office, with 108 assessments per type.

#### 3.1.1. Multidimensional Objective Acoustic Characterization

The descriptive statistics for the objective acoustic and psychoacoustic metrics across the four environment types are summarized in Tables 2 and 3. Shapiro–Wilk normality tests indicated that none of the acoustic or psychoacoustic metrics followed a normal distribution ($p$-values < 0.05), requiring the use of non-parametric statistical analysis. The results of the Kruskal–Wallis tests showed significant differences ($p$-values < 0.001) in acoustic energy, spectral, and temporal content among the environment types.

**Table 2.** Descriptive statistics of the acoustic characteristics of the soundtracks ($n$ = 432) recorded from air-conditioned built environments in terms of nine conventional acoustic metrics.

| Metric | Unit | 1. Library ($n$ = 108) | 2. Classroom ($n$ = 108) | 3. Lecture Hall ($n$ = 108) | 4. Office ($n$ = 108) | Total ($n$ = 432) | Kruskal–Wallis Test | Post Hoc Tests |
|---|---|---|---|---|---|---|---|---|
| | | Median (IQR) | Median (IQR) | Median (IQR) | Median (IQR) | Median (IQR) | $p$-Value | Order |
| $L_{Zeq}$ | dB | 66.1 (8.6) | 72.4 (5.8) | 72.7 (8.2) | 64.4 (9.5) | 68.4 (8.3) | <0.001 | 4 < 1 < 2 = 3 |
| $L_{Aeq}$ | dBA | 45.1 (7.5) | 48.6 (8.6) | 49.4 (8.7) | 50.0 (7.0) | 48.4 (8.6) | <0.001 | 1 < 2 = 3 = 4 |
| $L_{A10}$ | dBA | 45.5 (7.5) | 49.2 (9.2) | 50.0 (8.1) | 50.3 (7.3) | 48.9 (8.6) | <0.001 | 1 < 2 = 3 = 4 |
| $L_{A50}$ | dBA | 45.1 (7.5) | 48.5 (8.5) | 49.3 (9.0) | 50.0 (7.0) | 48.4 (8.6) | <0.001 | 1 < 2 = 3 = 4 |
| $L_{A90}$ | dBA | 44.7 (7.4) | 47.9 (8.1) | 48.9 (9.3) | 49.6 (7.5) | 48.0 (8.3) | <0.001 | 1 < 2 = 3 = 4 |
| $L_{A10}$–$L_{A90}$ | dBA | 0.75 (0.2) | 0.81 (0.4) | 1.1 (0.6) | 0.70 (0.2) | 0.79 (0.4) | <0.001 | 1 = 4 < 2 = 3 |
| *NC* | NC | 43 (9) | 46 (9) | 47 (8) | 46 (8) | 46 (8) | <0.001 | 1 < 2 < 3 = 4 |
| *NR* | NR | 43 (8) | 46 (9) | 48 (7) | 47 (7) | 46 (9) | <0.001 | 1 < 2 < 3 = 4 |
| *RC* | RC | 42 (8) | 45 (8) | 47 (8) | 47 (8) | 45 (8) | <0.001 | 1 < 2 = 3 = 4 |

Notes. $n$ = number of soundtracks, $L_Z$ = unweighted sound pressure level, $L_A$ = A-weighted sound pressure level, *NC* = noise criterion, *NR* = noise rating, *RC* = room criteria, $_{eq}$ = time-equivalent, $_{5/10/50/90/95}$ = percentiles of 5/10/50/90/95%, IQR = interquartile range, post hoc tests = Dunn's tests of an acoustic metric in multiple comparisons between different environmental types.

**Table 3.** Descriptive statistics of the acoustic characteristics of the soundtracks ($n$ = 432) recorded from air-conditioned built environments in terms of fifteen psychoacoustic metrics.

| Metric | Unit | 1. Library ($n$ = 108) | 2. Classroom ($n$ = 108) | 3. Lecture Hall ($n$ = 108) | 4. Office ($n$ = 108) | Total ($n$ = 432) | Kruskal–Wallis Test | Post Hoc Tests |
|---|---|---|---|---|---|---|---|---|
| | | Median (IQR) | Median (IQR) | Median (IQR) | Median (IQR) | Median (IQR) | $p$-Value | Order |
| $L_N$ | phon | 64.5 (8.7) | 67.4 (8.7) | 69.1 (9.6) | 83.1 (8.1) | 70.2 (13.5) | <0.001 | 1 < 2 < 3 < 4 |
| $N$ | sone | 5.5 (3.5) | 6.7 (3.8) | 7.5 (5.0) | 19.9 (10.8) | 8.1 (7.6) | <0.001 | 1 < 2 < 3 < 4 |
| $N_5$ | sone | 5.7 (3.6) | 6.9 (4.0) | 7.9 (4.8) | 20.7 (11.0) | 8.4 (7.6) | <0.001 | 1 < 2 < 3 < 4 |
| $N_{95}$ | sone | 5.1 (3.3) | 6.2 (3.6) | 6.7 (4.9) | 18.1 (9.9) | 7.5 (7.0) | <0.001 | 1 < 2 = 3 < 4 |
| $S$ | acum | 1.2 (0.1) | 0.87 (0.2) | 1.1 (0.2) | 1.2 (0.1) | 1.1 (0.3) | <0.001 | 2 < 3 < 1 < 4 |
| $S_5$ | acum | 1.3 (0.1) | 0.95 (0.3) | 1.2 (0.3) | 1.3 (0.1) | 1.2 (0.3) | <0.001 | 2 < 3 < 1 < 4 |
| $S_{95}$ | acum | 1.1 (0.1) | 0.81 (0.2) | 0.90 (0.3) | 1.2 (0.1) | 1.0 (0.3) | <0.001 | 2 < 3 < 1 < 4 |
| $R$ | asper | 0.07 (0.02) | 0.07 (0.03) | 0.08 (0.02) | 0.12 (0.05) | 0.08 (0.03) | <0.001 | 1 = 2 = 3 < 4 |
| $R_5$ | asper | 0.11 (0.03) | 0.11 (0.04) | 0.12 (0.03) | 0.19 (0.10) | 0.12 (0.05) | <0.001 | 1 = 2 < 3 < 4 |
| $R_{95}$ | asper | 0.04 (0.01) | 0.04 (0.02) | 0.05 (0.01) | 0.07 (0.03) | 0.05 (0.02) | <0.001 | 1 = 2 = 3 < 4 |
| $FS$ | vacil | 0.01 (0.04) | 0.03 (0.02) | 0.02 (0.03) | 0.06 (0.05) | 0.03 (0.04) | <0.001 | 1 = 3 < 2 < 4 |
| $FS_5$ | vacil | 0.02 (0.05) | 0.04 (0.03) | 0.03 (0.03) | 0.08 (0.06) | 0.04 (0.05) | <0.001 | 1 < 3 < 2 < 4 |
| $FS_{95}$ | vacil | 0.005 (0.03) | 0.02 (0.02) | 0.007 (0.02) | 0.05 (0.04) | 0.02 (0.03) | <0.001 | 2 < 1 = 3 < 4 |
| $f_{Mod}$ | Hertz | 135 (48) | 133 (55) | 141 (35) | 89 (70) | 131 (48) | <0.001 | 4 < 1 = 2 = 3 |
| $N_5$–$N_{95}$ | sone | 0.64 (0.3) | 0.78 (0.4) | 1.1 (0.4) | 2.0 (1.2) | 0.96 (0.7) | <0.001 | 1 < 2 < 3 < 4 |

Notes. $n$ = number of soundtracks, $_{5/10/50/90/95}$ = percentiles of 5/10/50/90/95%, IQR = interquartile range, $N$ = total loudness, $S$ = sharpness, $R$ = roughness, $FS$ = fluctuation strength, $f_{Mod}$ = dominant modulation frequency determined in the calculation of $R$, post hoc tests = Dunn's tests of an acoustic metric in multiple comparisons between different environmental types.

The built environment of l. Library consistently exhibited the lowest noise levels across most conventional metrics, including $L_{Aeq}$ (Median = 45.1 dBA), $L_{A10}$ (45.5 dBA), and $RC$ ratings (42). Post hoc Dunn's tests confirmed that the Library environment was significantly quieter than the other three types. The built environments of 2. Classroom and 3. Lecture Hall demonstrated significantly higher temporal variability ($L_{A10}$–$L_{A90}$ = 1.1 dBA) compared to the other environment types. The results suggested that although the built environment of 4. Office is louder than the Library, the noise profile is more constant than that of the university Classroom and Lecture Hall.

Similar to the conventional acoustic metrics, the psychoacoustic characteristics exhibited significant variations across the four environment types ($p$-values < 0.001). The built environment of 4. Office was characterized by the highest intensities across most psychoacoustic metrics. It recorded the highest median $N$ (19.9 sone), $R$ (0.12 asper), and $FS$ (0.06 vacil). Post hoc Dunn's tests confirmed that the Office environment was significantly louder and rougher than all other types. In addition, the Office environment illustrated the highest temporal variability in loudness fluctuation ($N_5$–$N_{95}$ = 2.0 sone). In contrast, the built environment of 2. Classroom, despite being relatively quiet ($N$ = 6.7 sone), was distinguished by its spectral characteristics. It exhibited the lowest $S$ (0.87 acum), suggesting a duller sound profile compared to other built environments.

#### 3.1.2. Multidimensional Subjective Perceptual Responses

The subjective assessments quantified via the PPS in the *EPA* framework are presented in Figure 2. Significant differences were observed for all perceptual dimensions across the four environment types ($p$-values < 0.05). The built environment of 1. Library was consistently rated as the most favorable environment, achieving the lowest *EPA*-score ($-7.0 \pm 4.8$). In the design of the PPS, lower values indicate better sound quality (e.g., $-3$ = "Extremely Pleasant" vs. +3 = "Extremely Unpleasant"). Therefore, the Library environment was perceived as the "Quietest" ($E1/P1$), most "Relaxed" ($E2$), and most "Pleasant" ($E3$). Moreover, the built environment of 2. Classroom showed the lowest

A-score (−2.6 ± 4.0), a result that aligns with its lowest objective *S* value (0.87 acum) in Table 2.

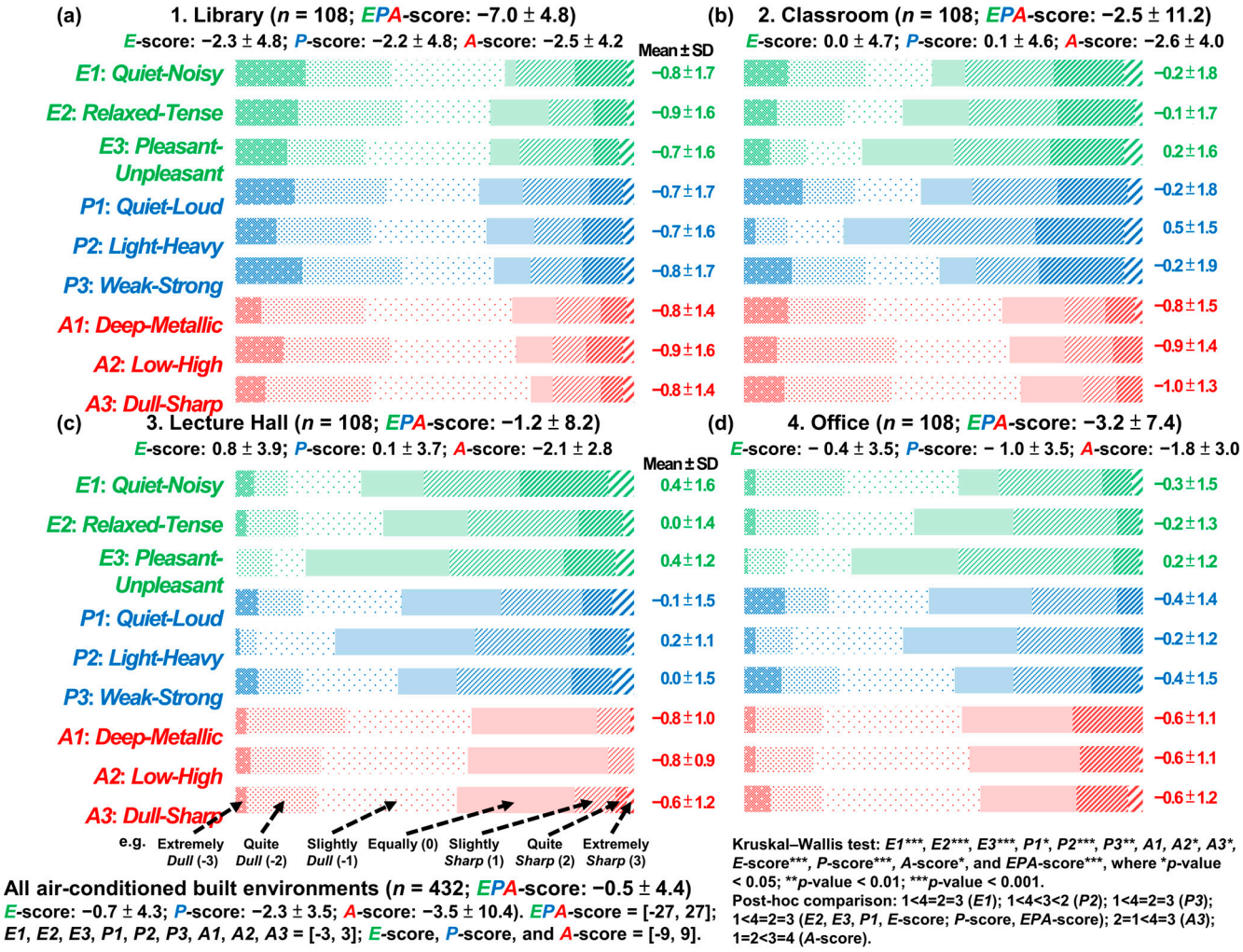



**Figure 2.** A horizontal stacked bar chart of the perceptual data (*n* = 432) captured using the Psychoacoustics Perception Scale (PPS) in *Evaluation*, *Potency*, and *Activity* (*EPA*) dimensions for four types of air-conditioned built environments: (**a**) 1. Library, (**b**) 2. Classroom, (**c**) 3. Lecture Hall, and (**d**) 4. Office.

### 3.1.3. Reliability of Sound Quality Assessment

Cronbach's α analysis was conducted to verify the internal consistency of the subjective responses measured by the PPS in quantifying sound quality perceptions across the three fundamental *E*-, *P*-, and *A*-dimensions. As shown in Table 4, the Cronbach's α for all perceptual *E*-, *P*-, *A*-, and *EPA*-dimensions exceeded 0.80 across all built environments, indicating good internal consistency in the collected data [53].

**Table 4.** Cronbach's α coefficients for subjective responses across four types of air-conditioned built environments.

| Environment Type | *Evaluation* (*E*) | *Potency* (*P*) | *Activity* (*A*) | *EPA* |
|---|---|---|---|---|
| 1. Library (*n* = 108) | 0.969 | 0.961 | 0.943 | 0.977 |
| 2. Classroom (*n* = 108) | 0.929 | 0.877 | 0.945 | 0.916 |
| 3. Lecture Hall (*n* = 108) | 0.915 | 0.875 | 0.880 | 0.872 |
| 4. Office (*n* = 108) | 0.888 | 0.829 | 0.863 | 0.827 |
| Total (*n* = 432) | 0.934 | 0.897 | 0.918 | 0.923 |

### *3.2. Performance Evaluation of Assessment Methods*

#### 3.2.1. Prediction of Individual *E*-, *P*-, and *A*-Scores Using the PML Assessment Method

As shown in Table 5, higher $\rho$ values were found between the PML predictions and the individual *E*-, *P*-, and *A*-scores compared to conventional metrics. For the *E*-score, the PML prediction ($\rho$ = 0.77) showed a 148% and 38% improvement over $L_{Aeq}$ ($\rho$ = 0.31) and *RC* ($\rho$ = 0.56), respectively. Similarly, for the *P*-score, the PML prediction ($\rho$ = 0.80) demonstrated a 166% and 38% improvement over $L_{Aeq}$ ($\rho$ = 0.30) and *RC* ($\rho$ = 0.58). For the *A*-score, the PHMLM ($\rho$ = 0.57) showed a 171% improvement over the psychoacoustic metric *S* ($\rho$ = 0.21).

**Table 5.** Spearman's rank correlation coefficients between the individual *E*-, *P*-, and *A*-scores and the psychoacoustic machine learning (PML) predictions and the individual acoustic and psychoacoustic metrics.

| Variable | PML | $L_{Aeq}$ | $L_{A10}$ | $L_{A50}$ | $L_{A90}$ | *NC* | *NR* | *RC* | $L_N$ | *N* | $N_5$ | $N_{95}$ | *S* | $S_5$ | $S_{95}$ | *R* | $R_5$ | $R_{95}$ | $f_{Mod}$ | $N_5$–$N_{95}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *E*-score | 0.77 | 0.31 | 0.56 | 0.56 | 0.56 | 0.56 | 0.55 | 0.56 | 0.56 | 0.38 | 0.38 | 0.38 | NA | NA | NA | 0.31 | 0.30 | 0.30 | NA | 0.33 |
| *P*-score | 0.80 | 0.30 | 0.58 | 0.58 | 0.58 | 0.58 | 0.58 | 0.58 | 0.58 | 0.37 | 0.37 | 0.37 | NA | NA | NA | 0.26 | 0.25 | 0.26 | NA | 0.32 |
| *A*-score | 0.57 | NA | 0.16 | 0.16 | 0.16 | 0.16 | 0.12 | 0.14 | 0.16 | 0.21 | 0.21 | 0.21 | 0.21 | 0.17 | 0.14 | 0.17 | 0.18 | 0.18 | 0.17 | 0.17 |

Notes. *n* = number of soundtracks, $L_Z$ = unweighted sound pressure level, $L_A$ = A-weighted sound pressure level, *NC* = noise criterion, *NR* = noise rating, *RC* = room criteria, $_{eq}$ = time-equivalent, $_{5/10/50/90/95}$ = percentiles of 5/10/50/90/95%, *N* = total loudness, *S* = sharpness, *R* = roughness, *FS* = fluctuation strength, $f_{Mod}$ = dominant modulation frequency determined in the calculation of *R*. NA = no significant correlation found.

#### 3.2.2. Prediction of *EPA*-Score

The $CRA_{EPA}$ model was established using stepwise linear regression to identify the most significant acoustic and psychoacoustic metrics as predictors for the *EPA*-score. The analysis identified *RC*, $N_5-N_{95}$, $f_{Mod}$, $L_{Zeq}$ and $L_{A90}$ as the significant independent variables. The final regression model achieved an adjusted $R^2$ = 0.315, $F(5, 426) = 40.6$, $p$-value < 0.001). Among the predictors, *RC* demonstrated the strongest positive influence (B = 3.96, $p$-value < 0.001), while $N_5-N_{95}$ showed a strong negative contribution (B = −4.16, $p$-value < 0.001). The background noise level ($L_{A90}$) also negatively impacted the *EPA*-score (B = −2.19, $p$-value = 0.006), with minor contributions from $L_{Zeq}$ (B = −0.40, $p$-value = 0.001) and $f_{Mod}$ (B = −0.03, $p$-value = 0.013). The resulting regression equation was found to be:

$$CRA_{EPA} = 3.96 \times RC - 4.16 \times (N_5 - N_{95}) - 0.03 \times f_{Mod} - 0.40 \times L_{Zeq} - 2.9 \times L_{A90} - 41.7 \quad (9)$$

The predictive performance of the overall sound quality (*EPA*-score) was evaluated by comparing the PML assessment methods against the $CRA_{EPA}$ (Equation (9)), $GPM_{EPA}$ (Equation (2)), and the metrics $L_{Aeq}$, *RC*, and *N*. As shown in Figure 3, the $PML_{EPA}$ demonstrated the highest goodness-of-fit among the models across all environment types. For all air-conditioned built environments, the $PML_{EPA}$ achieved a strong correlation ($\rho$ = 0.79) and explained the highest variance (adjusted $R^2$ = 0.61). In contrast, the $CRA_{EPA}$ showed moderate performance ($\rho$ = 0.58, adjusted $R^2$ = 0.32), while the $GPM_{EPA}$ showed the weakest predictive power ($\rho$ = 0.40, adjusted $R^2$ = 0.15). The performance of the $GPM_{EPA}$ was comparable to that of single-value acoustic metrics such as *N* ($\rho$ = 0.40), indicating that general models fail to capture the specific acoustic properties in different environments.

The superiority of the PML approach was consistent across individual environments. In the built environment of 3. Lecture Hall, the $PML_{EPA}$ achieved an exceptional fit (adjusted $R^2$ = 0.77), providing a 28% improvement over the $CRA_{EPA}$ (adjusted $R^2$ = 0.60) and a 43% improvement over the $GPM_{EPA}$ (adjusted $R^2$ = 0.54). Similarly, in the built environment of 1. Library, the $PML_{EPA}$ (adjusted $R^2$ = 0.69) outperformed the $CRA_{EPA}$ (adjusted $R^2$ = 0.23) by 200%, highlighting the strength of the heatmap-based features in quiet environments where traditional metrics often struggle. Furthermore, traditional acoustic metrics ($L_{Aeq}$, *RC*, and *N*) showed limited predictive capability, with adjusted $R^2$ values generally below

0.30 for all air-conditioned environment (see Figure 3d–f). The results confirmed that while conventional metrics and $CRA_{EPA}$ can provide a baseline assessment, they lack the sensitivity of the $PML_{EPA}$ to accurately predict multidimensional sound quality in complex air-conditioned environments.

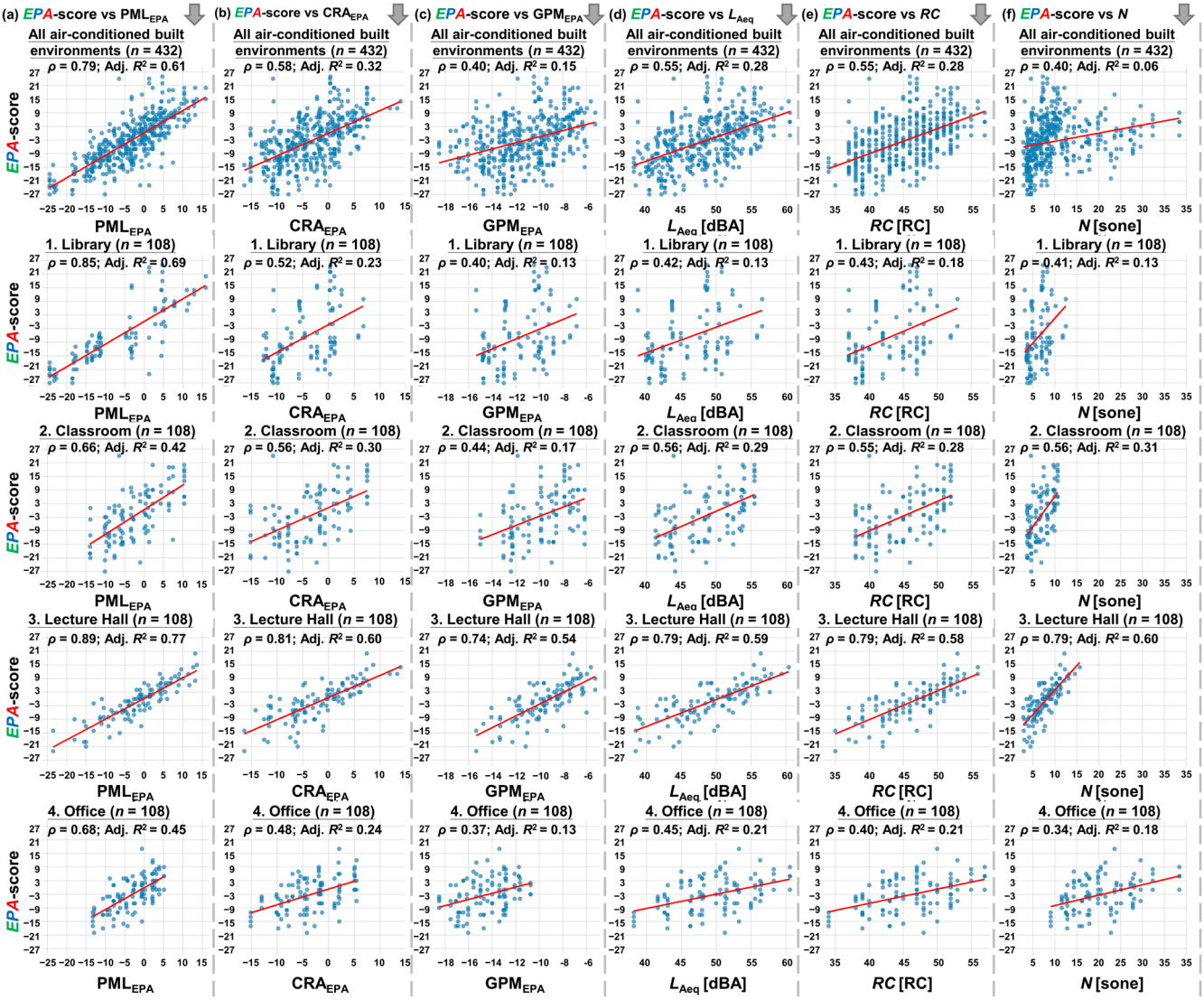



**Figure 3.** Comparative predictive performance for *EPA*-score prediction across four types of air-conditioned built environments ($n$ = 432): Scatter plots of observed vs. predicted *EPA*-scores from (**a**) psychoacoustic machine learning model ($PML_{EPA}$), (**b**) conventional regression approach ($CRA_{EPA}$), (**c**) general prediction model ($GPM_{EPA}$), and traditional acoustic metrics (**d**) A-weighted equivalent continues sound pressure level ($L_{Aeq}$), (**e**) room criteria (*RC*), and (**f**) total loudness (*N*). The regression line (red) and coefficient of determination (adjusted $R^2$) indicate the goodness-of-fit for each model.

### 3.2.3. Prediction of Occupant Well-Being-Related Noise Impacts

The capability of the models to predict occupant well-being, specifically the negative noise impacts (*O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*), was evaluated by integrating the predicted *EPA*-scores into the $GPM_{O1–4}$, Equations (3)–(6). Two hybrid models were established: $GPM_{O1–4\text{-}PML}$, which substitutes the $GPM_{EPA}$ in the equations with the $PML_{EPA}$, and $GPM_{O1–4\text{-}PCRA}$, which substitutes it with the $CRA_{EPA}$.

These hybrid models were then compared against the standalone $PML_{EPA}$ and $CRA_{EPA}$, as well as traditional acoustic metrics ($L_{Aeq}$, *RC*, and *N*), to assess their predictive performance for occupant well-being-related noise impacts.

Tables 6 and 7 demonstrated a consistent relationship where higher adjusted $R^2$ corresponds to lower root mean squared error (RMSE) and narrower 95% confidence intervals (CIs), confirming the reliability of adjusted $R^2$ as the primary statistics indicator across all methods for subsequent discussion in this paper. As shown in Table 6, the $PML_{EPA}$ consistently achieved the highest goodness-of-fit across all well-being-related noise impacts (*O1*–*O4*) for all air-conditioned built environments ($n$ = 432). For *O1*: *Discomfortable*, the $PML_{EPA}$ achieved a correlation of $\rho$ = 0.65 (adjusted $R^2$ = 0.42), significantly outperforming the $CRA_{EPA}$ ($\rho$ = 0.50, adjusted $R^2$ = 0.25). The hybrid model $GPM_{O1\text{-}PML}$ yielded nearly identical performance to the standalone $PML_{EPA}$ ($\rho$ = 0.64, adjusted $R^2$ = 0.41), whereas the $GPM_{O1\text{-}CRA}$ and traditional acoustic metrics ($L_{Aeq}$ and *RC*) showed weaker correlations ($\rho$ = 0.48). The superiority of the PML assessment method was even more pronounced for *O2*: *Annoying* and *O3*: *Stressful*. The $PML_{EPA}$ achieved correlations of 0.63 and 0.64 (adjusted $R^2$ = 0.40–0.41), respectively, providing a substantial improvement over the $CRA_{EPA}$ ($\rho$ = 0.46–0.50). In contrast, the $GPM_{O2–O3}$ and single-value metrics ($L_{Aeq}$, *RC*, and *N*) showed weak predictive power, with adjusted $R^2$ values dropping as low as 0.02 for *O3*: *Stressful* predictions using *N*. For *O4*: *Unacceptable*, the $PML_{EPA}$ remained the best predictor ($\rho$ = 0.60, adjusted $R^2$ = 0.36), effectively identifying the acoustic conditions that occupants deemed unacceptable. Notably, in the library environment, the $PML_{EPA}$ achieved exceptional accuracy across all dimensions, with correlations reaching $\rho$ = 0.74–0.78 (adjusted $R$ = 0.36 = 0.53–0.62), demonstrating its sensitivity to the subtle acoustic properties that drive dissatisfaction in quiet spaces. These results confirm that integrating the multidimensional *EPA*-scores predicted by the PML into well-being models $GPM_{O1–4\text{-}PML}$ offers a far more reliable tool for acoustic assessment than CRA or simple energy-based acoustic or psychoacoustic metrics.

**Table 6.** Comparison of Spearman's rank correlation coefficients $\rho$ and adjusted coefficients of determination (adjusted $R^2$ in parentheses) for predicting occupant well-being related noise impacts (*O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*) in the four types of air-conditioned built environment using different assessment methods.

| Dependent Variable | $PML_{EPA}$ | $CRA_{EPA}$ | $GPM_{O1–4}$ | $GPM_{O1–4\text{-}PML}$ | $GPM_{O1–4\text{-}CRA}$ | $L_{Aeq}$ | *RC* | *N* |
|---|---|---|---|---|---|---|---|---|
| *O1*: *Discomfortable* | | | | | | | | |
| 1. Library ($n$ = 108) | 0.74 (0.53) | 0.45 (0.18) | 0.32 (0.10) | 0.73 (0.53) | 0.46 (0.20) | 0.35 (0.11) | 0.44 (0.21) | 0.43 (0.16) |
| 2. Classroom ($n$ = 108) | 0.49 (0.23) | 0.38 (0.14) | 0.31 (0.09) | 0.48 (0.22) | 0.39 (0.13) | 0.42 (0.15) | 0.66 (0.45) | 0.65 (0.44) |
| 3. Lecture Hall ($n$ = 108) | 0.66 (0.49) | 0.71 (0.50) | 0.57 (0.26) | 0.65 (0.49) | 0.71 (0.49) | 0.64 (0.42) | 0.40 (0.14) | 0.41 (0.16) |
| 4. Office ($n$ = 108) | 0.66 (0.43) | 0.50 (0.26) | 0.36 (0.11) | 0.66 (0.43) | 0.46 (0.23) | 0.43 (0.20) | 0.37 (0.14) | 0.35 (0.10) |
| Total ($n$ = 432) | 0.65 (0.42) | 0.50 (0.25) | 0.32 (0.09) | 0.64 (0.41) | 0.49 (0.24) | 0.48 (0.22) | 0.48 (0.23) | 0.33 (0.06) |
| *O2*: *Annoying* | | | | | | | | |
| 1. Library ($n$ = 108) | 0.78 (0.59) | 0.45 (0.20) | 0.42 (0.13) | 0.70 (0.49) | 0.44 (0.21) | 0.38 (0.12) | 0.39 (0.15) | 0.36 (0.11) |
| 2. Classroom ($n$ = 108) | 0.46 (0.19) | 0.35 (0.12) | 0.30 (0.08) | 0.27 (0.08) | 0.23 (0.05) | 0.38 (0.12) | 0.36 (0.11) | 0.37 (0.12) |
| 3. Lecture Hall ($n$ = 108) | 0.67 (0.47) | 0.71 (0.48) | 0.56 (0.17) | 0.58 (0.31) | 0.54 (0.22) | 0.67 (0.43) | 0.68 (0.44) | 0.67 (0.45) |
| 4. Office ($n$ = 108) | 0.61 (0.36) | 0.55 (0.30) | 0.22 (0.05) | NA | NA | 0.48 (0.23) | 0.48 (0.24) | 0.48 (0.21) |
| Total ($n$ = 432) | 0.63 (0.40) | 0.50 (0.25) | 0.42 (0.16) | 0.35 (0.13) | 0.20 (0.04) | 0.46 (0.20) | 0.46 (0.21) | 0.18 (0.01) |
| *O3*: *Stressful* | | | | | | | | |
| 1. Library ($n$ = 108) | 0.76 (0.62) | 0.45 (0.20) | 0.35 (0.15) | 0.76 (0.62) | 0.45 (0.20) | 0.37 (0.12) | 0.37 (0.15) | 0.35 (0.11) |
| 2. Classroom ($n$ = 108) | 0.60 (0.33) | 0.46 (0.20) | 0.25 (0.03) | 0.60 (0.33) | 0.46 (0.20) | 0.46 (0.19) | 0.44 (0.18) | 0.46 (0.19) |
| 3. Lecture Hall ($n$ = 108) | 0.50 (0.27) | 0.46 (0.24) | 0.15 (0.02) | 0.50 (0.27) | 0.46 (0.24) | 0.44 (0.21) | 0.45 (0.22) | 0.45 (0.24) |
| 4. Office ($n$ = 108) | 0.61 (0.38) | 0.50 (0.26) | −0.27 (0.08) | 0.61 (0.38) | 0.50 (0.26) | 0.45 (0.21) | 0.46 (0.22) | 0.46 (0.20) |
| Total ($n$ = 432) | 0.64 (0.41) | 0.46 (0.22) | 0.17 (0.03) | 0.64 (0.41) | 0.46 (0.22) | 0.42 (0.17) | 0.43 (0.18) | 0.23 (0.02) |
| *O4*: *Unacceptable* | | | | | | | | |
| 1. Library ($n$ = 108) | 0.77 (0.61) | 0.40 (0.15) | 0.29 (0.08) | 0.75 (0.57) | 0.46 (0.22) | 0.32 (0.09) | 0.33 (0.11) | 0.31 (0.07) |
| 2. Classroom ($n$ = 108) | 0.50 (0.24) | 0.43 (0.19) | 0.34 (0.12) | 0.39 (0.15) | 0.27 (0.06) | 0.45 (0.19) | 0.44 (0.18) | 0.44 (0.19) |
| 3. Lecture Hall ($n$ = 108) | 0.69 (0.48) | 0.65 (0.45) | 0.65 (0.23) | 0.56 (0.33) | 0.49 (0.25) | 0.66 (0.45) | 0.68 (0.46) | 0.66 (0.47) |

**Table 6.** *Cont.*

| Dependent Variable | $PML_{EPA}$ | $CRA_{EPA}$ | $GPM_{O1–4}$ | $GPM_{O1–4-PML}$ | $GPM_{O1–4-CRA}$ | $L_{Aeq}$ | *RC* | *N* |
|---|---|---|---|---|---|---|---|---|
| 4. Office ($n$ = 108) | 0.39 (0.15) | 0.32 (0.10) | 0.22 (0.05) | 0.36 (0.13) | 0.32 (0.10) | 0.26 (0.07) | 0.27 (0.08) | 0.27 (0.07) |
| Total ($n$ = 432) | 0.60 (0.36) | 0.45 (0.20) | 0.42 (0.15) | 0.50 (0.25) | 0.36 (0.13) | 0.44 (0.19) | 0.44 (0.19) | 0.37 (0.12) |

Notes. $n$ = number of soundtracks, $CRA_{EPA}$ = conventional regression approach for *EPA*-score prediction, $PML_{EPA}$ = psychoacoustic machine learning for *EPA*-score prediction, $GPM_{O1–4}$ = general prediction model for *O1–O4* prediction, $GPM_{O1–4-PML}$ = hybrid models combining $GPM_{O1–4}$ and $PML_{EPA}$, $GPM_{O1–4-CRA}$ = hybrid models combining $GPM_{O1–4}$ and $CRA_{EPA}$, $L_{Aeq}$ = A-weighted equivalent continuous sound pressure level, *RC* = room criteria, *N* = total loudness, NA = no significant correlation found.

**Table 7.** Comparison of root mean squared error (RMSE) and confidence intervals (CIs) of predicted occupant well-being-related noise impacts (*O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*) in the four types of air-conditioned built environments using different assessment methods.

| Dependent Variable | $PML_{EPA}$ | $CRA_{EPA}$ | $GPM_{O1–4}$ | $GPM_{O1–4-PML}$ | $GPM_{O1–4-CRA}$ | $L_{Aeq}$ | *RC* | *N* |
|---|---|---|---|---|---|---|---|---|
| *O1*: *Discomfortable* ($n$ = 432, Mean = 3.87, SD = 1.52) | | | | | | | | |
| RMSE | 1.16 | 1.32 | 1.45 | 1.17 | 1.33 | 1.34 | 1.33 | 1.47 |
| 95% CIs | [3.77, 3.99] | [3.75, 4.00] | [3.74, 4.01] | [3.76, 3.99] | [3.75, 4.00] | [3.75, 4.00] | [3.75, 4.00] | [3.74, 4.01] |
| *O2*: *Annoying* ($n$ = 432, Mean = 3.91, SD = 1.49) | | | | | | | | |
| RMSE | 1.15 | 1.29 | 1.36 | 1.39 | 1.46 | 1.33 | 1.32 | 1.48 |
| 95% CIs | [3.8, 4.02] | [3.79, 4.03] | [3.78, 4.04] | [3.78, 4.04] | [3.77, 4.05] | [3.79, 4.04] | [3.79, 4.04] | [3.77, 4.05] |
| *O3*: *Stressful* ($n$ = 432, Mean = 3.70, SD = 1.43) | | | | | | | | |
| RMSE | 1.10 | 1.27 | 1.43 | 1.10 | 1.27 | 1.30 | 1.29 | 1.42 |
| 95% CIs | [3.59, 3.8] | [3.58, 3.82] | [3.56, 3.83] | [3.59, 3.08] | [3.58, 3.82] | [3.57, 3.82] | [3.57, 3.82] | [3.56, 3.83] |
| *O4*: *Unacceptable* ($n$ = 432, Mean = 3.56, SD = 1.53) | | | | | | | | |
| RMSE | 1.17 | 1.32 | 1.41 | 1.26 | 1.38 | 1.36 | 1.35 | 1.51 |
| 95% CIs | [3.44, 3.67] | [3.43, 3.69] | [3.42, 3.70] | [3.44, 3.68] | [3.42, 3.69] | [3.43, 3.69] | [3.43, 3.69] | [3.41, 3.71] |

Notes. $n$ = number of soundtracks, $CRA_{EPA}$ = conventional regression approach for *EPA*-score prediction, $PML_{EPA}$ = psychoacoustic machine learning for *EPA*-score prediction, $GPM_{O1–4}$ = general prediction model for *O1–O4* prediction, $GPM_{O1–4-PML}$ = hybrid models combining $GPM_{O1–4}$ and $PML_{EPA}$, $GPM_{O1–4-CRA}$ = hybrid models combining $GPM_{O1–4}$ and $CRA_{EPA}$, $L_{Aeq}$ = A-weighted equivalent continuous sound pressure level, *RC* = room criteria, *N* = total loudness.

## 4. Discussion

### *4.1. Multidimensional Sound Quality Assessment Framework Application*

The fundamental advantage of the PML assessment method lies in its integration of "psychoacoustics" with "machine learning" to create a holistic, multidimensional sound quality assessment framework (see Figure 1). The results of the multidimensional objective acoustic characterization demonstrated that psychoacoustic metrics (see Table 3) provide critical spectral and temporal information that complements traditional energy-based metrics (see Table 2). For instance, while traditional metrics indicated varying noise levels, the psychoacoustic analysis revealed that the built environment of 2. Classroom exhibited the lowest *S* values. This spectral characteristic, which traditional metrics failed to capture, aligns perfectly with the lower *A*-scores observed in the subjective responses (see Figure 2), confirming the agreement between objective psychoacoustic features and human perception. This supports recent findings that single-number ratings often fail to capture the acoustic characteristics of noise perception in complex indoor environments [54].

Furthermore, the study demonstrated that subjective perceptual responses cannot be explained solely by energy content-related metrics. A striking example is the built environment of 4. Office, which recorded the highest energy values in Table 2. Counter-intuitively, the subjective results in Figure 2 showed that the office environment had a *P*-score of −1, which is lower than that of the 2. Classroom and 3. Lecture Hall. This discrepancy can be attributed to the specific psychoacoustic characteristics of the Office noise, specifically the lowest $f_{Mod}$ observed in Table 3. The low loudness sensation associated with low-frequency noise leads to a lower perception of *Potency* despite high physical energy

levels, a phenomenon consistent with studies on low-frequency HVAC noise impacts [55]. Consequently, this resulted in a lower overall *EPA*-score for the office environment compared to university Classroom and Lecture Hall environments. These findings substantiate the necessity of using the PPS for multidimensional sound quality assessment, as simple energy metrics fail to account for the complex non-linear processing of human hearing [56]. The high internal consistency (Cronbach's $\alpha > 0.82$) observed across all environment types further validates the PPS as a reliable psychometric tool for evaluating air-conditioned built environments [12,49].

### *4.2. Predictive Performance of PML for Sound Quality Prediction*

The PML approach demonstrated superior performance in predicting the multidimensional sound quality scores (*E*, *P*, and *A*-scores) compared to traditional methods. Spearman's rank correlation tests confirmed that PML predictions achieved significantly higher correlations with the individual *E*-, *P*-, and *A*-scores than individual acoustic and psychoacoustic metrics (see Table 5). When evaluating the overall *EPA*-score prediction, the $PML_{EPA}$ was found to be better than both the $CRA_{EPA}$ and the $GPM_{EPA}$. While the $CRA_{EPA}$ showed reasonable performance because it is a case-specific model developed using the subjective assessment results from this study, it relies on obtaining subjective data, which is often impractical for design stages. The $GPM_{EPA}$, designed as a pure-objective method without the need for subjective inputs, showed lower performance in this study (adjusted $R^2 = 0.15$). This is likely because the GPM assessment method was originally designed based on outdoor soundscapes [37], and its current formulation may not fully capture the specific noise source characteristics of indoor air-conditioning systems. However, the GPM remains valuable for building acoustics design as it offers a generalized baseline without requiring jury testing.

The significantly better performance of the PML model can be attributed to its ability to learn non-linear noise features from 227 × 227-pixel psychoacoustic heatmaps representing 0.002 s resolution evolution of loudness, sharpness, roughness, and fluctuation strength. Unlike using time-averaged stationary metrics in linear regression models (CRA and GPM), the deep learning architecture of the PML captures the complex interactions between spectral and temporal variations, offering a solution to the feature engineering problem inherent in developing pure-objective prediction models [57]. This capability makes PML the most feasible method for predicting the *EPA*-score, as it effectively bridges the gap between objective signal properties and subjective human perception [58], accounting the human–environmental interaction.

### *4.3. Implications for Occupant Well-Being Prediction*

The predictive power of the PML approach extended to the assessment of occupant well-being. The standalone $PML_{EPA}$ model demonstrated the highest accuracy in predicting the negative noise impacts (*O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*), surpassing both the traditional metrics and the hybrid models $GPM_{O1–4}$. Although the hybrid models ($GPM_{O1–4\text{-}PML}$) utilized the *EPA*-scores predicted by the PML, they yielded slightly lower goodness-of-fit compared to the standalone $PHMLM_{EPA}$ model (see Tables 6 and 7). The same domain mismatch between the GPM developed for outdoor soundscapes [37] and the $PHMLM_{EPA}$ developed for indoor air-conditioned environments likely explains this result. The physics-based GPM parameters optimized for outdoor psychoacoustic relationships may conflict with the non-linear mappings learned by $PHMLM_{EPA}$ from indoor soundscapes, constraining hybrid performance. Future investigation should examine GPM and $PHMLM_{EPA}$ performance across both indoor and outdoor soundscapes before attempting integration. Only after domain-specific validation can data-

driven predictions be effectively incorporated into generalized frameworks applicable to both environments. This indicates that simply integrating high-quality EPA predictions into the existing GPM framework, which was originally designed for different acoustic contexts, may not fully capitalize on the precision of the PML features. Therefore, the combination mechanism between data-driven predictions (PML or CRA) and generalized frameworks (GPM) requires further investigation to optimize the predictive combined effect.

In addition, the high performance of the standalone $PML_{EPA}$ confirms that the method serves as a reliable assessment method for human comfort. The significance of these findings contributes directly to the field of built environment physics in studying the impact of acoustics on occupant well-being. By providing a validated, high-accuracy tool for predicting dissatisfaction and stress, the PML method empowers engineers and designers to prioritize occupant health in the acoustic design of sustainable buildings in the future [59].

### *4.4. Limitations and Future Work*

While the PML demonstrated superior performance as a standalone model, the hybrid approach combining PML predictions with the GPM framework $GPM_{O1–4}$ yielded slightly lower goodness-of-fit for the models. This suggests that the current method of integrating data-driven inputs into the generalized model structure requires optimization. Future work should investigate more advanced fusion techniques to better leverage the collaboration between the feature extraction of PML and the generalized structure of GPM. Although some soundtracks were randomly selected from the PML training datasets, independent listening tests with new subjects generated distinct perceptual labels, validating cross-panel prediction capability. Future work should validate PML performance using completely different soundtracks from diverse HVAC systems and building types to confirm generalizability beyond the original acoustic training domain.

Moreover, the subjective perceptual responses were collected from 36 participants with a mean age of 24.1 $\pm$ 2.2 years, of whom 69.4% were male. This represents a narrow demographic range, predominantly comprising young university students. While this participant profile facilitated standardized experimental control, it limits the generalizability of well-being predictions, particularly for diverse "Office" environments where actual occupant age and gender distributions are much wider. Future research should expand participant recruitment to include a broader demographic spectrum across age (e.g., 20–65 years), gender balance, occupational diversity, and varying hearing abilities to enhance ecological validity. Furthermore, field studies in actual office settings with long-term occupant monitoring would validate model performance across real-world demographic variability. These extensions would strengthen the practical applicability of the PML method for sustainable building design across diverse occupant populations.

In addition, the PML model was trained on a specific dataset of air-conditioned sounds. Further validation is required to test its generalizability across a broader range of building types (e.g., residential, hospitals) and noise sources [60] (e.g., traffic intrusion, machinery). The AlexNet transfer learning approach may have optimized feature extraction for this particular building inventory. Future research should employ cross-dataset validation with more independent field recordings. In addition, the current study highlights the "black box" nature of ML [61]. Future research could explore Explainable-AI techniques [62] to visualize which specific regions of the psychoacoustic heatmaps contribute most to the predictions, offering deeper insights into the physical drivers of sound perception. Therefore, the impact of noise source changes on PML performance needs systematic investigation to establish stronger thresholds for practical implementation in diverse real-world scenarios. While PML demonstrated better predictive performance (adjusted $R^2$ = 0.61 for *EPA*-score than

CRA and GPM, this is not a comparison flaw but a valid finding of method strengths. The results reflected PML's ability to capture non-linear temporal–spectral complexity (0.002 s resolution psychoacoustic heatmaps) inherent to HVAC noise that linear methods cannot address. However, simplicity and interpretability remain critical practical advantages of CRA and GPM. CRA offers straightforward implementation using time-averaged metrics with transparent regression coefficients that enable engineering design rules. GPM provides general applicability across environments without retraining through fixed equations that facilitate rapid assessment. PML's computational complexity represents a trade-off against CRA/GPM's practicality for real-time building applications. Future work should validate all three methods in more engineering-practical scenarios across additional perspectives, including real-time monitoring systems, diverse HVAC configurations, long-term field studies beyond controlled jury tests, and cross-cultural perceptual validation. Hybrid GPM-PML approaches combining non-linear feature extraction with interpretable linear outputs will balance predictive power with practical utility for sustainable building acoustic design.

## 5. Conclusions

This study comprehensively evaluated CRA, GPM, and PML for predicting sound quality (*EPA*-score) and negative noise impacts on occupant well-being (*O1*: *Discomfortable*, *O2*: *Annoying*, *O3*: *Stressful*, and *O4*: *Unacceptable*) across air-conditioned built environments. PML demonstrated better performance with Spearman $\rho = 0.79$ (adjusted $R^2 = 0.61$) for *EPA*-score prediction ($n = 432$) than CRA ($\rho = 0.58$, adjusted $R^2 = 0.32$) and GPM ($\rho = 0.40$, adjusted $R^2 = 0.15$). Furthermore, PML also demonstrated better performance in well-being predictions (O1: adjusted $R^2 = 0.42$, *O2*: adjusted $R^2 = 0.40$, and *O3*: adjusted $R^2 = 0.41$) and showed exceptional accuracy in quiet library environments (adjusted $R^2 = 0.53$–$0.62$). GPM's poor performance reflects an outdoor-to-indoor domain mismatch, while CRA's limitations stem from static time-averaged metrics. PML's psychoacoustic heatmaps capture critical temporal variations missed by both alternatives. This cross-method comparison identifies PML as the optimal objective assessment tool for building acoustic design, delivering reliable occupant well-being predictions without subjective testing to support health-focused sustainable building practices.

**Author Contributions:** Conceptualization, K.W.M. and C.M.M.; methodology, K.W.M.; software, C.M.M. and F.-L.C.; validation, C.M.M.; formal analysis, K.W.M.; investigation, K.W.M.; resources, H.M.W.; data curation, K.W.M.; writing—original draft preparation, K.W.M.; writing—review and editing, K.W.M., C.M.M., F.-L.C. and H.M.W.; visualization, K.W.M.; supervision, C.M.M.; project administration, C.M.M.; funding acquisition, C.M.M. and F.-L.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (PolyU15219421).

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

**Acknowledgments:** Students Hung Yeuk Lun, Au Wing Kin, Leung Wai Yuen, and Yip Lai Hei from The Hong Kong Polytechnic University are acknowledged for their assistance in sound measurements and data collection.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

PHMLM Psychoacoustic Heatmap Machine Learning Models

| | |
|---|---|
| ML | Machine Learning |
| PML | Psychoacoustic Machine Learning |
| GPM | General Prediction Model |
| CRA | Conventional Regression Approach |
| E | Evaluation |
| P | Potency |
| A | Activity |
| EPA | Evaluation, Potency, and Activity |
| PPS | Psychoacoustics Perception Scale |
| WHO | World Health Organization |
| HVAC | Heating, Ventilation, and Air-Conditioning |
| NC | Noise Criteria |
| NR | Noise Rating |
| RC | Room Criteria |
| ANN | Artificial Neural Network |
| SD | Standard Deviation |
| IQR | Interquartile Range |
| SGDM | Stochastic Gradient Descent with Momentum |
| ReLU | Rectified Linear Unit activation function |
| AI | Artificial Intelligence |
| RMSE | Root Mean Squared Error |
| CIs | Confidence Intervals |

## Appendix A

*Equations of the Psychoacoustic Metrics*

$$N'(t) = 0.08\left(\frac{E_{TQ}}{E_0}\right)^{0.23}\left[\left(0.5 + 0.5\frac{E(t)}{E_{TQ}}\right)^{0.23} - 1\right]\left(\frac{sone}{Bark}\right), \tag{A1}$$

where $N'(t)$ is the time-varying specific loudness in sone per Bark, $E(t)$ is the time-varying excitation of the sound, $E_{TQ}$ is the excitation of threshold in quiet, and $E_0$ is the reference excitation. The loudness of a 1 kHz pure tone sound at 40 dB is defined as 1 sone and 40 phon.

$$N(t) = \int_0^{24\ Bark} N'(t)dz(sone)\ , \tag{A2}$$

where $N(t)$ is the time-varying total loudness.

$$S(t) = 0.11\frac{\int_0^{24\ Bark} N'(t)g(z)zdz}{\int_0^{24\ Bark} N'dz}(acum), \tag{A3}$$

$$g(z) := \begin{cases} 1, z \le 14 \\ 0.00012Z^4 - 0.0056Z^3 + 0.1Z^2 - 0.81z + 3.51, z > 14 \end{cases} \tag{A4}$$

where $S(t)$ is the time-varying sharpness, and $g$(z) is critical-band-rate dependent. The sharpness of a 1 kHz pure tone sound at 60 dB is defined to be 1 acum.

$$R = 0.0003f_{\text{Mod}}\int_0^{24\ Bark} \Delta L(z)dz(asper), \tag{A5}$$

$$FS = 0.008\frac{\int_0^{24\ Bark} \Delta L(z)dz}{\left(\frac{f_{\text{Mod}}}{4} + \frac{4}{f_{\text{Mod}}}\right)}(vacil), \tag{A6}$$

where $f_{\text{Mod}}$ is the dominant modulation frequency for the effects of frequency resolution (Hz), and ΔL is the temporal depth for masking the effect of temporal resolution (dB). A

1 kHz pure tone at 60 dB with 100% amplitude modulation at 70 Hz and 4 Hz is equal to 1 asper and 1 vacil, respectively.

For the calculation of $N$, the time-averaged excitation of the sound is utilized.

$$L_{\mathrm{N}} = \begin{cases} 40\left(\frac{N}{sone} + 0.005\right)^{0.35}, N < 1 \text{ sone} \\ 40 + 10\log_2\left(\frac{N}{sone}\right), N \geq 1\ sone \end{cases} (phon) \quad \text{(A7)}$$

where $L_{\mathrm{N}}$ is the loudness level converted from the time-averaged $N$ in phon.

$N_{\mathrm{P\%}}$ and $S_{\mathrm{P\%}}$ are the values of instantaneous total loudness and sharpness that exceed P% of the elapsed time, respectively.

## References


1. Scrosati, C.; Scamoni, F. Managing measurement uncertainty in building acoustics. *Buildings* **2015**, *5*, 1389–1413. [CrossRef]
2. Park, J.; Loftness, V.; Aziz, A. Post-occupancy evaluation and IEQ measurements from 64 office buildings: Critical factors and thresholds for user satisfaction on thermal quality. *Buildings* **2018**, *8*, 156. [CrossRef]
3. Torresin, S.; Aletta, F.; Babich, F.; Bourdeau, E.; Harvie-Clark, J.; Kang, J.; Lavia, L.; Radicchi, A.; Albatici, R. Acoustics for supportive and healthy buildings: Emerging themes on indoor soundscape research. *Sustainability* **2020**, *12*, 6054. [CrossRef]
4. Aflaki, A.; Esfandiari, M.; Jarrahi, A. Multi-criteria evaluation of a library's indoor environmental quality in the tropics. *Buildings* **2023**, *13*, 1233. [CrossRef]
5. Fry, A. *Noise Control in Building Services*; Elsevier: Amsterdam, The Netherlands, 1988.
6. Munjal, M.J. *Acoustics of Ducts and Mufflers*; Wiley: New York, NY, USA, 2014.
7. Mak, C.M.; Ma, K.W.; Wong, H.M. *Prediction and Control of Noise and Vibration from Ventilation Systems*, 1st ed.; Taylor & Francis: Oxfordshire, UK, 2023.
8. *ISO 1996-1:2016*; Acoustics—Description, Measurement and Assessment of Environmental Noise—Part 1: Basic Quantities and Assessment Procedures. International Organization for Standardization: Geneva, Switzerland, 2016.
9. *ANSI/ASA S12.2-2019*; Criteria for Evaluating Room Noise. Acoustical Society of America: Melville, NY, USA, 2019.
10. Kim, A.; Wang, S.; McCunn, L.; Prozuments, A.; Swanson, T.; Lokan, K. Commissioning the acoustical performance of an open office space following the latest healthy building standard: A case study. *Acoustics* **2019**, *1*, 473–492. [CrossRef]
11. Thampanichwat, C.; Wongvorachan, T.; Bunyarittikit, S.; Chunhajinda, P.; Phaibulputhipong, P.; Wongmahasiri, R. The Architectural Design Strategies That Promote Attention to Foster Mindfulness: A Systematic Review, Content Analysis and Meta-Analysis. *Buildings* **2024**, *14*, 2508. [CrossRef]
12. Ma, K.W.; Mak, C.M.; Wong, H.M. Development of a subjective scale for sound quality assessments in building acoustics. *J. Build. Eng.* **2020**, *29*, 101177. [CrossRef]
13. Ma, K.W.; Wong, H.M.; Mak, C.M. Dental Environmental Noise Evaluation and Health Risk Model Construction to Dental Professionals. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1084. [CrossRef]
14. World Health Organization. *Environmental Noise Guidelines for the European Region*; WHO: Geneva, Switzerland, 2018.
15. World Health Organization. *Burden of Disease from Environmental Noise: Quantification of Healthy Life Years Lost in Europe*; WHO: Geneva, Switzerland, 2011.
16. Ma, K.W.; Mak, C.M.; Wong, H.M. The perceptual and behavioral influence on dental professionals from the noise in their workplace. *Appl. Acoust.* **2020**, *161*, 107164. [CrossRef]
17. Yonemura, M.; Lee, H.; Sakamoto, S. Subjective evaluation on the annoyance of environmental noise containing low-frequency tonal components. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7127.
18. Zhang, X.; Ba, M.; Kang, J.; Meng, Q. Effect of soundscape dimensions on acoustic comfort in urban open public spaces. *Appl. Acoust.* **2018**, *133*, 73–81. [CrossRef]
19. Erfanian, M.; Mitchell, A.J.; Kang, J.; Aletta, F. The psychophysiological implications of soundscape: A systematic review of empirical literature and a research agenda. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3533. [CrossRef]
20. *ISO/TS 15666:2003*; Acoustics—Assessment of Noise Annoyance by Means of Social and Socio-Acoustic Surveys. International Organization for Standardization: Geneva, Switzerland, 2003.
21. *ISO 12913-1:2014*; Acoustics—Soundscape—Part 1: Definition and Conceptual Framework. International Organization for Standardization: Geneva, Switzerland, 2014.
22. *ISO 12913-2:2018*; Acoustics—Soundscape—Part 2: Data Collection and Reporting Requirements. International Organization for Standardization: Geneva, Switzerland, 2018.
23. *ISO 12913-3:2019*; Acoustics—Soundscape—Part 3: Data Analysis. International Organization for Standardization: Geneva, Switzerland, 2019.

24. Soeta, Y.; Kagawa, H. Three dimensional psychological evaluation of aircraft noise and prediction by physical parameters. *Build. Environ.* **2020**, *167*, 106445. [CrossRef]
25. Lian, Y.; Ou, D.; Tan, R. The effects of sound source dominance and pressure level on cognitive performance and environmental perception in green space. *Appl. Acoust.* **2025**, *240*, 110897. [CrossRef]
26. Osgood, C.E. The nature and measurement of meaning. *Psychol. Bull.* **1952**, *49*, 197–273. [CrossRef]
27. Takada, M.; Tanaka, K.; Iwamiya, S.-I. Relationships between auditory impressions and onomatopoeic features for environmental sounds. *Acoust. Sci. Technol.* **2006**, *27*, 67–79. [CrossRef]
28. Galiana, M.; Llinares, C.; Page, Á. Subjective evaluation of music hall acoustics: Response of expert and non-expert users. *Build. Environ.* **2012**, *58*, 1–13. [CrossRef]
29. Ma, K.W.; Wong, H.M.; Mak, C.M. A systematic review of human perceptual dimensions of sound: Meta-analysis of semantic differential method applications to indoor and outdoor sounds. *Build. Environ.* **2018**, *133*, 123–150. [CrossRef]
30. Zwicker, E.; Fastl, H. *Psychoacoustics: Facts and Models*, 3rd ed.; Springer Science & Business Media: Berlin, Germany, 2007; Volume 22.
31. *ISO 532-1:2017*; Acoustics—Methods for Calculating Loudness—Part 1: Zwicker Method. International Organization for Standardization: Geneva, Switzerland, 2017.
32. Tang, S.K. Performance of noise indices in air-conditioned landscaped office buildings. *J. Acoust. Soc. Am.* **1997**, *102*, 1657–1663. [CrossRef]
33. Segura-Garcia, J.; Navarro-Ruiz, J.M.; Perez-Solano, J.J.; Montoya-Belmonte, J.; Felici-Castell, S.; Cobos, M.; Torres-Aranda, A.M. Spatio-temporal analysis of urban acoustic environments with binaural psycho-acoustical considerations for IoT-based applications. *Sensors* **2018**, *18*, 690. [CrossRef]
34. Ma, K.W.; Mak, C.M.; Wong, H.M. Acoustical measurements and prediction of psychoacoustic metrics with spatial variation. *Appl. Acoust.* **2020**, *168*, 107450. [CrossRef]
35. Ma, K.W.; Mak, C.M.; Wong, H.M. Effects of environmental sound quality on soundscape preference in a public urban space. *Appl. Acoust.* **2021**, *171*, 107570. [CrossRef]
36. Axelsson, Ö.; Nilsson, M.E.; Berglund, B. A principal components model of soundscape perception. *J. Acoust. Soc. Am.* **2010**, *128*, 2836–2846. [CrossRef]
37. Ma, K.W.; Mak, C.M.; Wong, H.M. Development of a sound quality model for noise impact prediction in building acoustics. *J. Build. Eng.* **2025**, *111*, 113183. [CrossRef]
38. Jiang, J.; Trundle, P.; Ren, J. Medical image analysis with artificial neural networks. *Comput. Med. Imaging Graph.* **2010**, *34*, 617–631. [CrossRef]
39. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105. [CrossRef]
40. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [CrossRef]
41. Vican, I.; Kreković, G.; Jambrošić, K. Improved fetal heartbeat detection using pitch shifting and psychoacoustics. *Biomed. Signal Process. Control* **2024**, *90*, 105850. [CrossRef]
42. Barros, A.; Geluykens, M.; Pereira, F.; Freitas, E.; Faria, S.; Goubert, L.; Vuye, C. Beyond noise levels: Vehicle classification using psychoacoustic indicators from pass-by road traffic noise and their correlations with speed and temperature. *Appl. Acoust.* **2023**, *214*, 109716. [CrossRef]
43. Souli, S.; Lachiri, Z. Audio sounds classification using scattering features and support vectors machines for medical surveillance. *Appl. Acoust.* **2018**, *130*, 270–282. [CrossRef]
44. Hvastja, A.; Ćirić, D.; Milivojčević, M.; Prezelj, J. Assessing air and noise pollution through acoustic classification of vehicles fuel types using deep learning. *Heliyon* **2025**, *11*, e43426. [CrossRef]
45. Potočnik, P.; Olmos, B.; Vodopivec, L.; Susič, E.; Govekar, E. Condition classification of heating systems valves based on acoustic features and machine learning. *Appl. Acoust.* **2021**, *174*, 107736. [CrossRef]
46. Wu, C.; Redonnet, S. A simple yet efficient data-driven model for the prediction of aircraft noise impact. *Aerosp. Sci. Technol.* **2025**, *163*, 110286. [CrossRef]
47. O'Reilly, D.; White, M.; Langenheim, N.; Alambeigi, P.; Huang, X.; Yang, T. Receiver-centric mapping of pedestrian noise annoyance: A cost-effective approach using random forest, psychoacoustic metrics and open-source data. *Sustain. Cities Soc.* **2025**, *130*, 106651. [CrossRef]
48. Dai, R.; Zhao, J.; Zhao, W.; Ding, W.; Huang, H. Exploratory study on sound quality evaluation and prediction for engineering machinery cabins. *Measurement* **2025**, *253*, 117684. [CrossRef]
49. Ma, K.W.; Mak, C.M.; Chung, F.-L.; Wong, H.M. Advanced Acoustic Monitoring Using Psychoacoustic Heatmap Machine Learning Models for Noise Impact Prediction in Air-Conditioned Building Environments. *Sensors* **2026**, *26*, 544. [CrossRef]

50. Zhang, T.; Feng, G.; Liang, J.; An, T. Acoustic scene classification based on Mel spectrogram decomposition and model merging. *Appl. Acoust.* **2021**, *182*, 108258. [CrossRef]
51. Dixon, W.J.; Massey, F.J., Jr. *Introduction to Statistical Aanalysis*, 4th ed.; McGraw-Hill Book Company: New York, NY, USA, 1983.
52. Razali, N.M.; Wah, Y.B. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *J. Stat. Model. Anal.* **2011**, *2*, 21–33.
53. Gliem, J.A.; Gliem, R.R. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. In *Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education*; The Ohio State University: Columbus, OH, USA, 2003.
54. Zannin, P.H.T.; Marcon, C.R. Objective and subjective evaluation of the acoustic comfort in classrooms. *Appl. Ergon.* **2007**, *38*, 675–680. [CrossRef]
55. Soeta, Y.; Shimokura, R. Sound quality evaluation of air-conditioner noise based on factors of the autocorrelation function. *Appl. Acoust.* **2017**, *124*, 11–19. [CrossRef]
56. Abbasi, A.M.; Motamedzade, M.; Aliabadi, M.; Golmohammadi, R.; Tapak, L. Study of the physiological and mental health effects caused by exposure to low-frequency noise in a simulated control room. *Build. Acoust.* **2018**, *25*, 233–248. [CrossRef]
57. Lionello, M.; Aletta, F.; Kang, J. A systematic review of prediction models for the experience of urban soundscapes. *Appl. Acoust.* **2020**, *170*, 107479. [CrossRef]
58. Jafari, M.J.; Khosrowabadi, R.; Khodakarim, S.; Mohammadian, F. The effect of noise exposure on cognitive performance and brain activity patterns. *Open Access Maced. J. Med. Sci.* **2019**, *7*, 2924. [CrossRef]
59. Starowicz, A.; Zieliński, M. Sustainable acoustics: The impact of AI on acoustics design and noise management. *Tech. Sci.* **2024**, *27*, 193–209. [CrossRef]
60. Piczak, K.J. Environmental sound classification with convolutional neural networks. In Proceedings of the 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), Boston, MA, USA, 17–20 September 2015; pp. 1–6.
61. Petch, J.; Di, S.; Nelson, W. Opening the black box: The promise and limitations of explainable machine learning in cardiology. *Can. J. Cardiol.* **2022**, *38*, 204–213. [CrossRef] [PubMed]
62. Holzinger, A.; Saranti, A.; Molnar, C.; Biecek, P.; Samek, W. Explainable AI methods-a brief overview. In Proceedings of the International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers, Vienna, Austria, 18 July 2020; pp. 13–38.


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.