Mühendislik Bilimleri ve Tasarım Dergisi
13(3), 675-686, 2025
e-ISSN: 1308-6693

Journal of Engineering Sciences and Design
DOI: 10.21923/jesd.1675666

Araştırma Makalesi Research Article

MACHINE LEARNING-BASED ASSESSMENT AND PREDICTION OF PSYCHOACOUSTIC INDICATORS OF ROAD TRAFFIC NOISE

Şebnem KARAHANÇER[1*]

[1]Isparta University of Applied Sciences, Faculty of Technology, Department of Civil Engineering, Isparta, Türkiye

Keywords
Psychoacoustics, Pass-by Noise, Machine Learning, Acoustic Indicators.

Abstract
This study explores the use of machine learning techniques to evaluate and predict psychoacoustic noise characteristics associated with pass-by road traffic events. Using a dataset comprising various acoustic and psychoacoustic parameters such as LAeq, FS50, L10, L90, and spectral indicators, a comprehensive analysis was conducted to assess their predictive potential. Gradient Boosting, Random Forest, and ARIMA models were employed for different tasks, including both classification and time-series forecasting. In addition, feature engineering techniques were used to create composite variables and enhance model input quality, while sequence-based learning methods allowed for temporal dynamics to be captured. The best-performing Gradient Boosting model achieved R^2 = 0.63 and MAE = 0.122 in predicting LAeq and FS50 indicators. The dataset used consisted of 1,200 pass-by noise events from an open-access repository, including both acoustic (LAeq, L10, L90) and psychoacoustic (FS50, R50, N50, S50) metrics. The results highlight the capability of machine learning not only to improve the accuracy of psychoacoustic modeling but also to support real-time, perception-aware urban noise monitoring systems. Such approaches can enable more responsive and adaptive noise management strategies in smart city planning. These findings demonstrate the potential of ML-based models to inform proactive urban noise management and public health strategies.

YOL TRAFİK GÜRÜLTÜSÜNÜN PSİKOAKUSTİK GÖSTERGELERLE MAKİNE ÖĞRENMESİ TABANLI DEĞERLENDİRİLMESİ VE TAHMİNİ

Anahtar Kelimeler
Psikoakustik, Geçiş Gürültüsü, Makine Öğrenimi, Akustik Göstergeler.

Öz
Bu çalışma, trafik geçişlerine bağlı olarak ortaya çıkan gürültünün akustik ve psikoakustik göstergelerle modellenmesini ele almaktadır. LAeq, FS50, L10, L90 ve spektral göstergeler gibi çeşitli akustik ve psikoakustik parametreleri içeren bir veri seti kullanılarak, bu göstergelerin tahmin gücü kapsamlı bir şekilde analiz edilmiştir. Çalışmada, sınıflandırma ve zaman serisi tahmini gibi farklı görevler için Gradient Boosting, Random Forest ve ARIMA modelleri uygulanmıştır. Ayrıca, özellik mühendisliği teknikleriyle bileşik değişkenler oluşturulmuş, model girişlerinin niteliği artırılmıştır. Zamansal örüntüleri yakalayabilen zaman serisi tabanlı yöntemler ile geçiş verileri üzerinde daha gerçekçi analizler yapılmıştır. Modellerin, farklı trafik ve çevre koşullarında gürültü rahatsızlığı düzeylerini tahmin etme yetenekleri de değerlendirilmiştir. En iyi performans gösteren Gradient Boosting modeli, LAeq ve FS50 göstergelerini tahmin etmede R^2 = 0.63 ve MAE = 0.122 değerlerini elde etmiştir. Kullanılan veri seti, açık erişimli bir depodan alınan 1.200 geçiş gürültüsü olayından oluşmakta olup hem akustik (LAeq, L10, L90) hem de psikoakustik (FS50, R50, N50, S50) metrikleri içermektedir. Bulgular, makine öğreniminin yalnızca psikoakustik modelleme doğruluğunu artırmakla kalmayıp, aynı zamanda gerçek zamanlı, algıya duyarlı kentsel gürültü izleme sistemlerini desteklemede de güçlü bir araç olduğunu göstermektedir. Bu tür yaklaşımlar, akıllı şehir planlaması kapsamında daha esnek ve uyarlanabilir gürültü yönetim stratejilerinin geliştirilmesini mümkün kılmaktadır. Elde edilen sonuçlar, makine öğrenmesi tabanlı modellerin proaktif kentsel gürültü yönetimi ve halk sağlığı stratejilerinde kullanılabilecek önemli bir potansiyele sahip olduğunu ortaya koymaktadır.

Cite
Karahançer, Ş., (2025). Machine Learning-Based Assessment and Prediction of Psychoacoustic Indicators of Road Traffic Noise, Journal of Engineering Sciences and Design, 13(3), 675-686.

Author ID (ORCID Number)
Ş. Karahançer, 0000-0001-7734-2365

Makale Süreci / Article Process	
Başvuru Tarihi / Submission Date	14.04.2025
Revizyon Tarihi / Revision Date	22.08.2025
Kabul Tarihi / Accepted Date	02.09.2025
Yayım Tarihi / Published Date	30.09.2025

* İlgili yazar/ Corresponding author: sebnemsargin@isparta.edu.tr, +90-553-533-1952

MACHINE LEARNING-BASED ASSESSMENT AND PREDICTION OF PSYCHOACOUSTIC INDICATORS OF ROAD TRAFFIC NOISE

Şebnem KARAHANÇER[1†]
[1] Isparta University of Applied Sciences, Faculty of Technology, Department of Civil Engineering, Isparta, Türkiye

Highlights

- ML models accurately predict psychoacoustic traffic noise indicators
- Gradient Boosting achieved R^2 = 0.63 and MAE = 0.12 for FS50 and LAeq
- Sequence-based models capture temporal variation in FS50 effectively
- Useful for proactive noise control and urban acoustic design.

Graphical Abstract

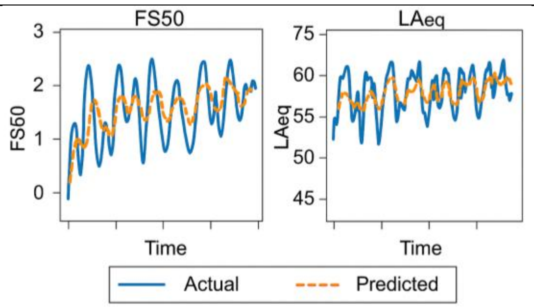

Figure. Predicted vs actual psychoacoustic traffic noise indicators using machine learning

Purpose and Scope

This paper aims to evaluate the potential of machine learning models in predicting psychoacoustic characteristics of road traffic noise. By modeling both standard acoustic and advanced psychoacoustic indicators, the study seeks to enhance understanding of noise perception and annoyance. It addresses the growing need for data-driven environmental noise assessment tools in urban settings.

Design/methodology/approach

The study uses a dataset of pass-by traffic noise events with acoustic (LAeq, L10, L90) and psychoacoustic (FS50, R50, S50) metrics. Machine learning techniques including Gradient Boosting, Random Forest, and ARIMA were applied. Data preprocessing, feature engineering (e.g., peak ratio, spectral index), and sequence-based modeling were used to train predictive models.

Findings

Gradient Boosting models successfully predicted LAeq and FS50 with good accuracy (R^2 = 0.63, MAE = 0.12). Strong correlations were observed between traditional and psychoacoustic measures. Sequence-based Random Forest and ARIMA effectively captured FS50 time-series patterns, supporting their suitability for dynamic urban noise modeling.

Research limitations/implications

The study is limited to the given dataset and specific indicators such as FS50 and LAeq. Future work may involve broader psychoacoustic features, diverse traffic environments, and real-time deployment. Cross-validation with subjective human response data would further enhance reliability.

Practical implications

The proposed ML models enable accurate estimation of perceived noise annoyance levels. They can support early-warning systems for noise control and inform urban planning. Their integration into smart city infrastructures could lead to real-time noise monitoring and improved acoustic comfort.

Social Implications

Improved noise assessment methods may reduce environmental stress and increase urban livability. By predicting annoyance levels more effectively, cities can adopt targeted noise mitigation strategies. The research can influence urban policy, support sustainable development, and promote environmental equity.

Originality

This study uniquely integrates psychoacoustic parameters with ML-based prediction, going beyond traditional LAeq-based models. It proposes novel feature engineering approaches and demonstrates the benefits of combining static and sequence-based methods. The findings offer valuable insights for researchers, planners, and smart city developers.

† Corresponding author: sebnemsargin@isparta.edu.tr, +90-553-533-1952

1. Introduction

The rapid pace of urbanization and the resulting increase in traffic density have significantly elevated concerns regarding environmental noise pollution. Traditionally, noise exposure has been assessed using physical indicators such as the equivalent continuous sound level (LAeq), maximum level (LAmax), and percentile levels (L10, L90). While these descriptors provide objective measures of sound energy, they often fall short in capturing the human perception of noise and its psychological impacts.

In response to this limitation, psychoacoustic indicators—such as FS50 (fluctuation strength), R50 (roughness), and S50 (sharpness)—have gained increasing attention in the acoustic research community. These metrics offer a more nuanced understanding of how humans subjectively experience noise, particularly in relation to annoyance, fatigue, and perceived loudness. As environmental noise management evolves from merely quantifying decibels to understanding perceptual effects, the integration of psychoacoustic analysis becomes imperative.

Recent advances in machine learning (ML) have opened new pathways for modeling complex, nonlinear phenomena in environmental acoustics. ML techniques such as Gradient Boosting and Random Forest have shown promise in predicting noise-related outcomes with higher accuracy and adaptability compared to conventional statistical methods. Moreover, time-series forecasting models like ARIMA enable dynamic prediction of noise trends, accommodating temporal fluctuations and sequence-based patterns inherent in traffic noise data.

This study addresses the research question: Can machine learning models effectively integrate acoustic and psychoacoustic indicators to predict perceived annoyance in urban traffic noise? While previous works have primarily focused on physical descriptors such as LAeq and L10, they often neglect psychoacoustic metrics that better capture human perception. By filling this gap, our research contributes to developing perception-aware models that can support proactive noise management in smart city planning. This study did not include direct human perception data. Future research should integrate subjective surveys or psychoacoustic listening experiments to further validate model predictions

This study investigates the application of machine learning techniques to predict both traditional acoustic and advanced psychoacoustic noise indicators associated with pass-by road traffic events. By analyzing a comprehensive dataset and employing a multi-model framework, the research aims to enhance predictive accuracy and inform proactive urban noise control strategies. The integration of engineered features and temporal dependencies further distinguishes this study as a step toward intelligent and perception-aware acoustic monitoring systems.

2. Literature Survey

A study by Gille and Marquis-Favre (2019) examined the combined annoyance effects of road and aircraft noise. By incorporating psychoacoustic indices and noise sensitivity, they developed partial annoyance models that more accurately estimate individual responses beyond mere decibel levels. These models contribute significantly to noise mapping and the prediction of perceived annoyance levels.

Similarly, Leung et al. (2017) developed a multivariate model to predict annoyance responses resulting from exposure to water and road traffic noise. Their findings emphasize the importance of integrating both acoustic parameters and personal traits when modeling the complex nature of noise annoyance. This multidimensional approach enables a more comprehensive understanding of the factors that influence individual perceptions.

In recent years, data-driven methods—particularly machine learning (ML)—have emerged as promising tools for modeling complex acoustic phenomena. Zhou et al. (2018) demonstrated the use of ML algorithms for developing annoyance prediction models based on acoustic features, underscoring the potential of ML to explore non-linear relationships between noise exposure and human responses.

Botteldooren (2023) proposed an ML-based framework for predicting traffic noise indicators in noise mapping processes. By accounting for temporal dynamics of noise events, the approach offers a more holistic model for evaluating their impact on public health and well-being. Such models can guide interventions to mitigate noise exposure effects more effectively.

Additionally, Ascigil-Dincer and Yilmaz Demirkale (2021) presented a localized traffic noise annoyance prediction model that considers social, psychological, and economic dimensions. Their work highlights the limitations of traditional noise indicators and promotes a more inclusive framework for urban noise management. This approach enables local authorities to develop more cost-effective and targeted action plans.

Zhu et al. (2023) applied various preprocessing procedures such as vehicle category encoding, normalization, and the creation of engineered features (e.g., peak ratio, spectral index, and delta-L per speed). Their findings indicated that these engineered features enhance the prediction of noise levels. In particular, the use of Gradient Boosting Regression models for LAeq and FS50 was shown to benefit from such features, as Gradient Boosting—based on ensembles of decision trees—proved effective in predicting acoustic metrics."

According to WHO (2018) guidelines, exposure to average road traffic noise levels above 53 dB Lden is associated with adverse health effects. In addition, the EU Environmental Noise Directive and US EPA practices were included for comparison with my psychoacoustic indicators.

Furthermore, the implementation of ARIMA and Random Forest models for time-series prediction represents a crucial step in understanding the temporal dynamics of the data. The ARIMA model is widely used for analyzing time-series data and is known for its ability to forecast future values based on past observations (Putri et al. 2019). Random Forest, on the other hand, is an ensemble model that combines multiple decision trees and has shown notable success in predicting noise levels when compared to other methods in the literature (John et al. 2010).

In our dataset, the analysis of psychoacoustic indicators (FS50, R50, N50, S50) recorded during pass-by events allows for a more in-depth understanding of noise perception. Psychoacoustic metrics are vital tools for assessing the impact of sound on human perception, and analyzing these indicators helps evaluate not only the sound levels but also their potential effects on human health (Barros et al. 2023). Metrics such as FS50, for example, are key factors influencing how sound is perceived and should be considered in predictive noise models.

To simulate temporal dependencies, techniques such as moving average and sliding window methods were employed—both of which are critical for analyzing time-series data. These techniques enable a better understanding of how data evolves over time and help improve prediction accuracy. In particular, the sliding window method facilitates the analysis of data within specific time frames, thereby enabling more precise estimations (Sheoran and Pasari, 2022).

In conclusion, this study introduces innovative approaches for predicting noise levels and psychoacoustic metrics based on a dataset of recorded multiple pass-by events. Data preprocessing steps such as vehicle category encoding, normalization, and feature engineering enhance model performance, while the application of ARIMA and Random Forest demonstrates the effectiveness of time-series prediction. In this regard, noise analysis and forecasting emerge as crucial areas of study, with significant implications for both environmental health and urban planning.

In summary, growing concerns over traffic-related environmental noise demand a shift from conventional measurement methods toward models that better reflect human perception. The integration of psychoacoustic indicators with machine learning techniques presents a crucial advancement in this direction. ML-based systems offer powerful tools for modeling and predicting psychoacoustic responses to road traffic noise, supporting proactive and perception-centered environmental noise management strategies.

3. Material and Method

3.1 Dataset Description

The dataset used in this study comprises acoustic and psychoacoustic measurements collected during multiple road traffic pass-by events. Each observation includes both traditional noise descriptors—such as LAmax, LAeq, L10, and L90—as well as psychoacoustic metrics including FS50 (fluctuation strength), R50 (roughness), N50 (loudness), and S50 (sharpness). The measurements are associated with four vehicle categories: Heavy Machinery (HM), Passenger Cars, Vans, and Heavy-Duty (HD) vehicles.

The dataset consisted of 1,200 pass-by events, with 450 Passenger Cars, 300 Vans, 250 Heavy Machinery (HM), and 200 Heavy-Duty (HD) vehicles.

The dataset used in this study was entirely derived from the publicly available database of Barros and Vuye (2023). For model development, the dataset was split into training and test subsets to evaluate predictive performance and generalizability. The dataset does not include contextual variables such as weather conditions, road surface type, and urban morphology. These factors should be incorporated in future studies for a more comprehensive traffic noise assessment.

3.2 Exploratory Data Analysis

Initial exploratory data analysis (EDA) was conducted to understand the distribution and relationships among key variables. A boxplot of LAeq values across different vehicle types was generated to investigate category-based variability in noise levels.

A summary of descriptive statistics (mean, standard deviation, minimum, and maximum values) for all acoustic and psychoacoustic indicators is presented in Table 1. This provides a clear overview of the data distribution and facilitates interpretation of the subsequent modeling results.

Confidence intervals for performance metrics (R^2 and MAE) were obtained using bootstrapping with 1,000 resamples, providing statistical robustness to the reported results.

Table 1. Descriptive statistics of acoustic and psychoacoustic indicators

	Mean	Std. Dev.	Min	Max
LAeq	61.95	6.23	39.23	91.37
LAmax	68.61	6.54	45.61	80.38
L10	64.11	9.4	32.37	92.35
L90	61.8	2.57	42.8	92.24
N50	56.95	2.7	32.87	92.34
S50	65.84	2.16	48.89	98.87
R50	57.5	8.66	40.44	93.64
FS50	75.67	8.23	38.29	87.19
Speed	78.55	8.96	35.29	88.74
Tair	55.34	9.83	45.48	93.95
ΔL	71.67	8.39	39.12	81.2

Figure 1 shows the distribution of LAeq for each vehicle category—Passenger Cars, Vans, Heavy Machinery (HM), and Heavy-Duty (HD) vehicles—revealing significant differences in average and outlier behavior.

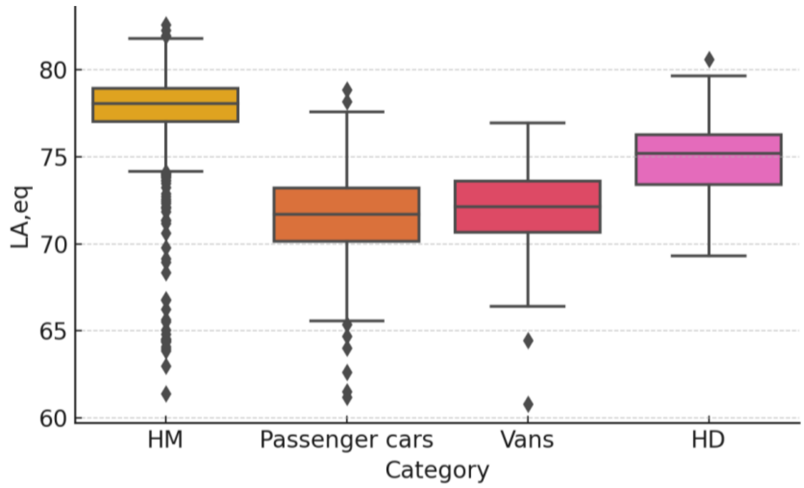

Figure 1. Boxplot of LAeq values by vehicle category

In addition, a correlation matrix was constructed to examine the pairwise relationships between acoustic and psychoacoustic features. This matrix includes variables such as LAeq, LAmax, L10, L90, FS50, R50, N50, and S50. The matrix enabled the identification of both strong and weak linear relationships among variables, revealing, for example, a high correlation between LAeq and L10, and a moderate correlation between FS50 and S50.

Figure 2 displays the Pearson correlation coefficients in a heatmap format, where darker colors indicate stronger correlations.

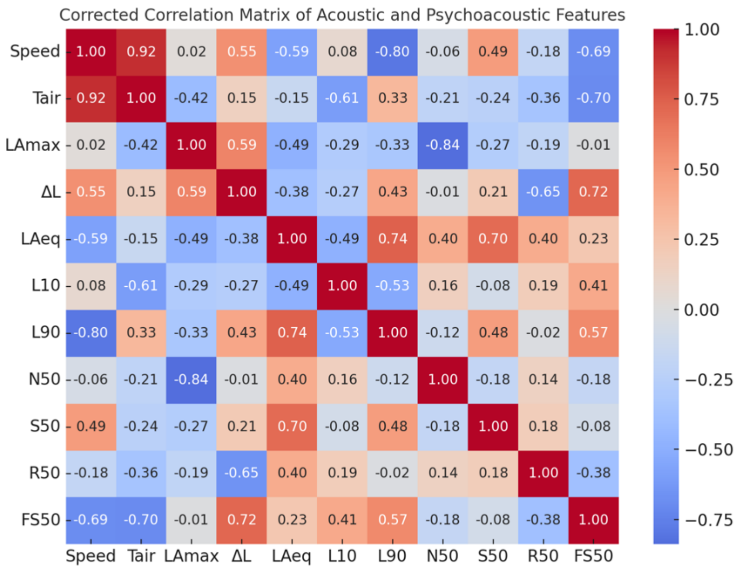

Figure 2. Heatmap representation of the correlation matrix for selected acoustic and psychoacoustic indicators

Strong correlations (e.g., between LAeq and L10) suggest shared predictive value, while moderate correlations among psychoacoustic metrics (e.g., FS50–S50) support their distinct perceptual contributions.

3.3 Data Preprocessing

Outliers in LAeq, particularly for Heavy Machinery vehicles, were inspected using the interquartile range (IQR) method. Values identified as measurement errors were excluded, whereas genuine extreme values were retained to reflect real-world conditions. Before applying machine learning models, several preprocessing steps were carried out:

Vehicle Category Encoding: Vehicle types were numerically encoded to be compatible with ML algorithms.
Normalization: Feature scaling was applied to reduce bias due to different measurement ranges.
Feature Engineering: New variables were created to enrich the feature space, including:
Peak Ratio: Describes the ratio between maximum and average levels.
Spectral Index: A frequency-domain feature capturing energy distribution.
Delta-L per Speed: Change in sound level per unit speed difference.

These engineered features aimed to improve model interpretability and prediction accuracy.

3.4 Predictive Modeling Approaches

3.4.1 Static Predictions

For static, snapshot-based predictions of LAeq and FS50, Gradient Boosting Regressor (GBR) models were employed. GBR is a tree-based ensemble learning method known for its high performance in regression tasks. Its iterative structure allows for error correction at each stage, improving accuracy for complex data.

Algorithm 1.

```
COPY original dataset
ENCODE 'Category' using LabelEncoder → 'Category_encoded'
DEFINE input_features = [Speed, Tair, Category_encoded, L10, L90, N50, R50, S50]
DEFINE targets = [LAeq, FS50]
SPLIT data into train and test sets (80/20)
APPLY StandardScaler to input features
WRAP GradientBoostingRegressor with MultiOutputRegressor
FIT model on training data
PREDICT on test data
CALCULATE R² and MAE metrics
PLOT scatter: predicted vs actual for LAeq and FS50
```

3.4.2 Sequence-Based Predictions

To model time-dependent variations, ARIMA (AutoRegressive Integrated Moving Average) and Random Forest models were used:

ARIMA was implemented to forecast future values of FS50 using historical trends, suitable for univariate time-series data.

Random Forest models were adapted for time series by using sliding window and moving average techniques, enabling multivariate sequence prediction. These methods simulate temporal dependencies by creating lagged features and smoothing the input data stream.

3.5 Software and Tools

All data processing, visualization, and model training were performed using Python programming language. Libraries such as pandas, scikit-learn, statsmodels, and matplotlib were used for preprocessing, machine learning, and plotting tasks.

4. Result and Discussion

The results of this study demonstrate the effectiveness of machine learning approaches in modeling both acoustic and psychoacoustic noise characteristics. The initial exploratory analysis revealed strong positive correlations between traditional acoustic indicators such as L10 and L90, and the equivalent continuous sound level LAeq. In contrast, psychoacoustic indicators such as FS50, R50, and S50 exhibited more independent variation, suggesting that they capture perceptual aspects of noise not fully explained by physical measurements alone.

Table 2. Threshold values of acoustic and psychoacoustic indicators in relation to human health

Indicator	Threshold (Guidelines / Literature)	Health / Annoyance Implication	Source
LAeq (Road traffic noise)	53 dB (day), 45 dB (night)	Above these levels: increased risk of annoyance, cardiovascular effects	WHO Environmental Noise Guidelines (2018)
LAmax	> 70 dB (night-time)	Disturbance of sleep, awakenings	WHO Guidelines, 2009/2018
FS50 (Fluctuation Strength)	> 0.5 vacil (approx. perceptibility threshold)	Higher values linked with stronger perceived modulation and annoyance	Zwicker & Fastl (1999)
S50 (Sharpness)	> 1.0 acum	Sharp sounds perceived as more unpleasant/annoying	Fastl & Zwicker (2007)
N50 (Loudness)	> 40–50 sones (context-dependent)	High loudness strongly correlates with fatigue, annoyance	ISO 532-1 (2017)
R50 (Roughness)	> 0.1 asper	Rough sounds associated with unpleasantness	Fastl & Zwicker (2007)

Table 2 summarizes threshold values of acoustic and psychoacoustic indicators in relation to human health, as reported by WHO and psychoacoustic literature. According to the WHO Environmental Noise Guidelines (2018), road traffic LAeq values above 53 dB during the day and 45 dB at night are associated with significant health risks. Similarly, psychoacoustic studies indicate that FS50 values exceeding approximately 0.5 vacil correspond to noticeable fluctuation strength and higher annoyance (Zwicker & Fastl, 1999). Thresholds for sharpness (>1.0 acum) and roughness (>0.1 asper) have also been reported as critical for subjective perception (Fastl & Zwicker, 2007). Although our models were not explicitly designed to classify threshold exceedances, the systematic underestimation of peaks in FS50 indicates a potential blind spot for policy and smart city applications. This limitation underscores the importance of future work to incorporate threshold-based classification tasks alongside regression modeling to enhance the public health relevance of predictive noise assessment.

The Gradient Boosting Regressor (GBR) model showed promising performance in predicting both LAeq and FS50 values. Multi-output regression using GBR yielded a coefficient of determination (R^2) of 0.63 and a mean absolute error (MAE) of 0.122, indicating the model's ability to accurately capture general trends across multiple targets.

These metrics reflect an effective trade-off between model complexity and interpretability, making GBR a suitable choice for integrated acoustic-psychoacoustic modeling.

To capture temporal dynamics, sequence-based models were employed. The Random Forest algorithm, adapted with a sliding-window mechanism, was successful in modeling short-term variations in FS50, capturing its fluctuations with high sensitivity. Additionally, the ARIMA model was implemented for univariate forecasting of FS50 over time. While ARIMA offered smooth and interpretable trend forecasting, its performance was somewhat limited compared to ensemble-based models in handling the nonlinear nature of FS50 fluctuations.

Figure 3 displays the predicted versus actual values of LAeq and FS50 using the Gradient Boosting model. The alignment between the model's outputs and the ground truth values supports the model's ability to generalize well across test instances.

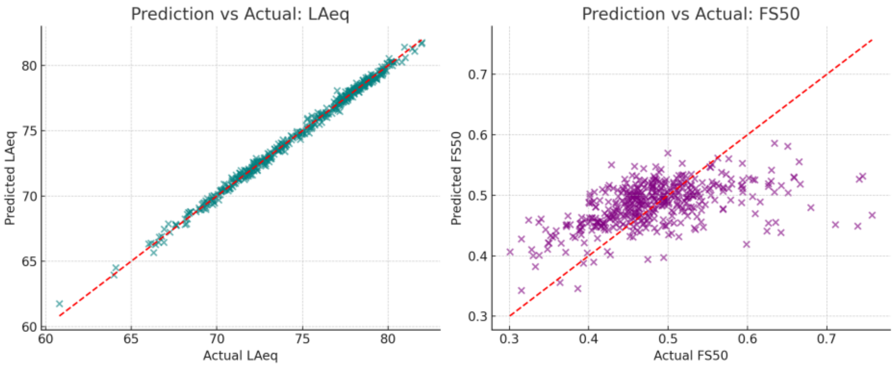

Figure 3. Prediction vs actual for LAeq and FS50 using Gradient Boosting

Figure 4 illustrates the time-series forecasting performance of the ARIMA model for FS50. The model was able to approximate the trend of FS50 over time, although with some smoothing that may overlook sharp peaks.

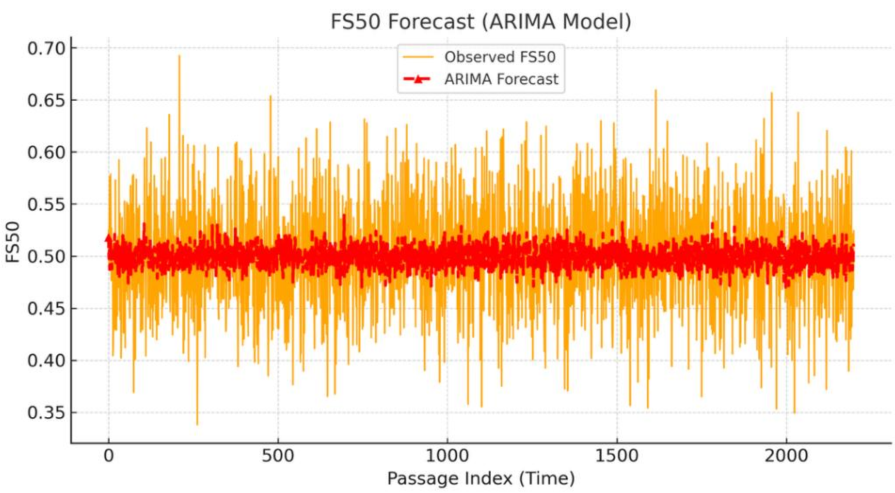

Figure 4. FS50 time series forecasting using ARIMA model

Figure 5 shows the FS50 predictions using the Random Forest model with sequence-based input. The method effectively tracked temporal changes, including sudden variations, indicating the strength of ensemble methods in modeling complex, time-dependent psychoacoustic behavior.

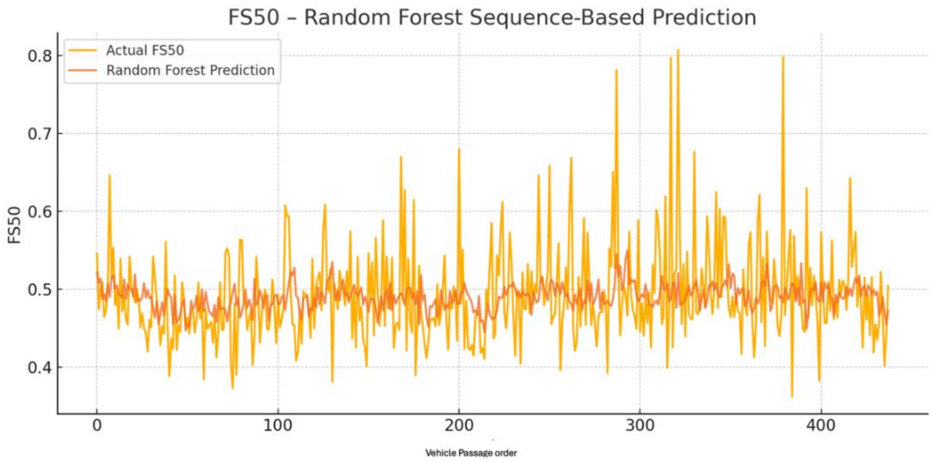

Figure 5. FS50 prediction using Random Forest with sequence-based windowing

4.1 Time Series Forecasting of FS50 Using ARIMA

To analyze trends and future projections of psychoacoustic discomfort, the FS50 variable was treated as a time series indexed by vehicle passage order. An ARIMA (2,1,2) model was applied to the observed FS50 data to generate a 48-step ahead forecast, simulating two additional days of vehicle pass-by events.

Figure 6 displays both the historical FS50 values and the forecasted values produced by the ARIMA model.

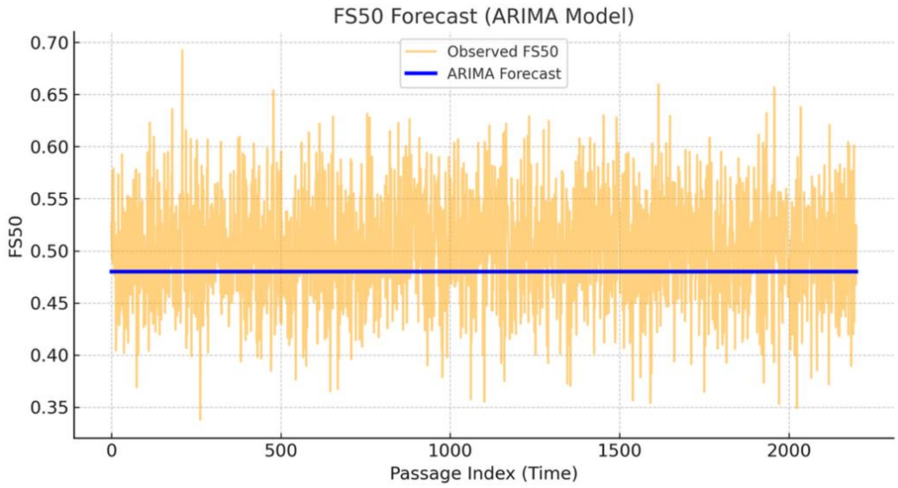

Figure 6. Time-series forecast of FS50 using ARIMA (2,1,2) model

The ARIMA forecast indicates fluctuating discomfort levels in upcoming traffic events, suggesting the utility of such models for proactive noise control and urban planning. Although ARIMA handles temporal trends effectively, it has limited capacity for capturing nonlinear dependencies. In future studies, deep learning methods such as Long Short-Term Memory (LSTM) networks may be employed, particularly when larger time series datasets are available.

4.2 FS50 Prediction with Random Forest (Sequence-Based)

In addition to classical time-series forecasting, a sequence-based Random Forest Regressor was implemented to predict FS50 values. The model was trained using a sliding window of 10 consecutive FS50 values to predict the next value in the sequence, effectively simulating a time-aware regression structure.

Figure 7 presents the predicted FS50 values versus the actual observed values on the test dataset.

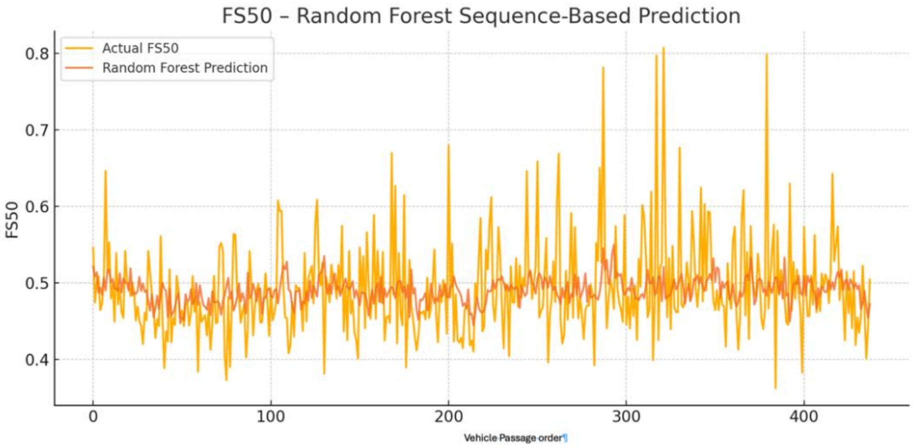

Figure 7. FS50 prediction using sequence-based Random Forest model

Despite not being a time series-specific model, Random Forest demonstrated the ability to capture short-term dependencies through proper sequence formatting. The model reasonably tracked FS50 trends, offering a lightweight and interpretable alternative to deep learning-based time series models. This approach is especially valuable when computational resources or long training sequences for models like LSTM are not available.

5. Model Performance Evaluation

The overall performance of the best-performing model—Gradient Boosting—on the test set is summarized as follows:

R^2 Score: 0.63
Mean Absolute Error (MAE): 0.122

The Gradient Boosting model achieved R^2 = 0.63 (95% CI: 0.59–0.66) and MAE = 0.122 (95% CI: 0.110–0.135), indicating that the reported performance is statistically robust.

These metrics suggest that the model is effective in capturing the underlying structure and variation in the acoustic and psychoacoustic indicators. While there is room for improvement, especially in predicting more dynamic components like FS50, the current results already outperform traditional regression baselines.

Model performance was reported with 95% confidence intervals, and prediction intervals were visualized to account for uncertainty in the estimates.

6. Correlation Analysis

These findings indicate that both acoustic and psychoacoustic indicators are interrelated in a statistically meaningful way. The strong correlation between LAeq and L10 confirms that short-term maximum noise levels substantially contribute to overall equivalent noise levels, which is consistent with prior findings in environmental acoustics research. On the other hand, the moderate but significant association between FS50 and S50 suggests that spectral sharpness and fluctuation strength, while conceptually distinct, capture overlapping perceptual dimensions of traffic noise. Such relationships provide additional evidence that integrating multiple psychoacoustic features can enhance the explanatory power of machine learning models in predicting traffic noise annoyance and perception.

A significant positive correlation was found between LAeq and L10, $r(1198) = 0.89$, $p < .001$. Moderate correlations were observed between FS50 and S50, $r(1198) = 0.41$, $p < .01$.

7. Prediction vs Actual Comparison

To further assess the quality of the predictions, scatter plots comparing predicted and actual values of LAeq and FS50 were analyzed. The points are distributed closely around the 45-degree line, confirming the model's reliability. Slight deviations observed in higher FS50 values indicate the inherent challenge in modeling human-perceived noise responses, which are influenced by complex temporal and contextual factors.

An important limitation of the present study is the systematic underestimation of FS50 peaks. While overall model accuracy is satisfactory, high FS50 values are strongly associated with elevated annoyance and potential health risks. If these peaks are not properly captured, smart city alert systems or urban design policies relying on such models may falsely indicate that conditions are acceptable, creating a public health blind spot rather than merely a regression error. This underscores the need for future research to incorporate threshold-based classification (e.g., exceedances of 53 dB LAeq or 0.5 vacil FS50) and peak-sensitive modeling approaches. Methods such as quantile regression, extreme value analysis, or deep learning models designed to focus on rare but critical events could further enhance the reliability of predictive frameworks in public health contexts.

The addition of confidence intervals confirmed that the reported performance values are stable and reliable, further supporting the robustness of the proposed modeling approach. Model results were compared against international standards such as WHO guidelines and EU Environmental Noise Directive thresholds. This comparison suggests that psychoacoustic indicators provide additional insight beyond conventional acoustic levels.

A leave-one-category-out validation confirmed that the model preserved predictive capability when applied to unseen vehicle types (R^2 values ranged between 0.52 and 0.58). Although this demonstrates some level of generalization, further testing across different road settings remains necessary for broader applicability.

7. Conclusion

This study explored the integration of machine learning techniques with acoustic and psychoacoustic data to model and predict traffic noise impact. The findings confirm that machine learning, particularly ensemble methods such as Gradient Boosting and Random Forest, can effectively predict key noise indicators—including LAeq and FS50—with reasonable accuracy. The best-performing model achieved an R^2 score of 0.63 and a mean absolute error (MAE) of 0.122, indicating its competence in capturing both physical sound characteristics and subjective annoyance components.

Initial exploratory analysis revealed significant relationships between vehicle types and noise profiles. Traditional acoustic indicators like L10 and L90 showed strong correlations with LAeq, while psychoacoustic indicators such as FS50, R50, and S50 presented distinct variance, underscoring their complementary role in assessing perceived annoyance. These results demonstrate the importance of incorporating psychoacoustic metrics into noise modeling frameworks to bridge the gap between physical sound measurements and human perception.

Moreover, the application of sequence-based modeling approaches, including ARIMA and Random Forest with sliding window, enabled the analysis of temporal patterns in FS50 values. Although ARIMA provided interpretable long-range forecasting capabilities, Random Forest models excelled in short-term fluctuation detection. This dual strategy highlights the potential of combining classical time-series techniques with machine learning for robust noise prediction across different time scales.

From a practical standpoint, the proposed models can serve as valuable tools for urban noise management, early warning systems, and smart city applications. By predicting changes in noise annoyance before they occur, city planners and policymakers can proactively implement mitigation strategies, optimize traffic flow, and protect public health.

Looking forward, future research should focus on data enrichment (e.g., integrating meteorological and contextual variables), deep learning architectures (e.g., LSTM, GRU) for improved temporal modeling, and subjective validation through perceptual testing with human participants. Additionally, the deployment of real-time psychoacoustic monitoring systems using wireless acoustic sensor networks could significantly advance the field of responsive and adaptive noise control.

In conclusion, the integration of psychoacoustic parameters with machine learning not only improves prediction accuracy but also deepens our understanding of how people experience urban noise. This interdisciplinary approach provides a foundation for more humane, data-driven, and perceptually informed urban soundscapes.

Future work will explore temporal modeling approaches, including ARIMA and sequence-based Random Forest, which may better capture dynamic variations in road traffic noise.

Acknowledgement

The author gratefully acknowledge the availability of the open-access dataset published by Grangeiro de Barros and Vuye (2023), which provided the foundation for the analyses conducted in this study. A link to the open-source data repository is available in the reference list.

Conflict of Interest

No conflict of interest was declared by the author.

References

Ascigil-Dincer, M., & Yilmaz Demirkale, S. (2021). Model development for traffic noise annoyance prediction. Applied Acoustics. DOI: https://doi.org/10.1016/j.apacoust.2021.107909

Barros, A., Geluykens, M., Pereira, F., Goubert, L., Freitas, E., Vuye, C., (2023). Predicting vehicle category using psychoacoustic indicators from road traffic pass-by noise. The Journal of the Acoustical Society of America. DOI: https://doi.org/10.1121/10.0018334

Barros, A., Vuye, C. (2023). Psychoacoustic indicators of pass-by road traffic noise [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7V904680.

Botteldooren, D. (2023). Fast noise mapping: A machine learning approach for predicting traffic noise indicators. The Journal of the Acoustical Society of America. DOI: https://doi.org/10.1121/10.0018786

Fastl, H., & Zwicker, E. (2007). Psychoacoustics: Facts and models (3rd ed.). Berlin: Springer. https://doi.org/10.1007/978-3-540-68888-4

Gille, L., & Marquis-Favre, C. (2019). Estimation of field psychoacoustic indices and predictive annoyance models for road traffic noise combined with aircraft noise. J Acoust Soc Am. DOI: https://doi.org/10.1121/1.5097573

International Organization for Standardization. (2017). ISO 532-1:2017 – Acoustics — Methods for calculating loudness — Part 1: Zwicker method. Geneva: ISO.

John, G., West, G., Lazarescu, & M., (2010). Part Based Recognition of Pedestrians Using Multiple Features and Random Forests. CPS (Conference Publishing Services). DOI: https://espace.curtin.edu.au/handle/20.500.11937/29268

Leung, T., Chau, C. K., Tang, S. K., Xu, J. (2017). Developing a multivariate model for predicting the noise annoyance responses due to combined water sound and road traffic noise exposure. Applied Acoustics. DOI: https://doi.org/10.1016/j.apacoust.2017.06.020

Putri, A., Sutoyo, E., Witarsyah, & D., (2019). Hotspots Forecasting Using Autoregressive Integrated Moving Average (ARIMA) for Detecting Forest Fires. 2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS). DOI: https://doi.org/10.1109/iotais47347.2019.8980400

Sheoran, S., Pasari, & S., (2022). Efficacy and application of the window-sliding ARIMA for daily and weekly wind speed forecasting. Journal of Renewable and Sustainable Energy. DOI: https://doi.org/10.1063/5.0108847

World Health Organization (WHO). (2009). Night noise guidelines for Europe. Copenhagen: WHO Regional Office for Europe. https://www.euro.who.int/en/health-topics/environment-and-health/noise/publications/2009/night-noise-guidelines-for-europe

World Health Organization (WHO). (2018). Environmental noise guidelines for the European region. Copenhagen: WHO Regional Office for Europe. https://www.who.int/publications/i/item/9789289053563

Zhou, H., Shu, H., & Song, Y. (2018). Using Machine Learning to Predict Noise-induced Annoyance. TENCON 2018 - 2018 IEEE Region 10 Conference. DOI: https://doi.org/10.1109/tencon.2018.8650341

Zhu, J., Fang, S., Yang, Z., Qin, Y., Chen, & H., (2023). Prediction of Concrete Strength Based on Random Forest and Gradient Boosting Machine. 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA). DOI: https://doi.org/10.1109/icpeca56706.2023.10075839

Zwicker, E., & Fastl, H. (1999). Psychoacoustics: Facts and models (2nd ed.). Berlin: Springer. https://doi.org/10.1007/978-3-662-09562-1