

Article

Normalized Frequency-Domain Acoustic Features for Same-Vehicle Load-State Difference Analysis

Yuzhe Pan , Zifang Yuan, Yuhao Liu and Hui Zhang *

School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen 518107, China; panyzh25@mail2.sysu.edu.cn (Y.P.); yuanzf6@mail2.sysu.edu.cn (Z.Y.); liuyh386@mail2.sysu.edu.cn (Y.L.)

* Correspondence: zhanghui@mail.sysu.edu.cn

Featured Application

This study provides a low-cost and interpretable acoustic feature extraction approach for same-vehicle load-state difference analysis in enclosed logistics yards, industrial parks, mining roads, and port-transport scenarios. It can support roadside non-contact monitoring systems where the same vehicle repeatedly passes through a fixed road-excitation location under comparable acquisition conditions.

Abstract

Roadside acoustic sensing may provide a low-cost approach to vehicle load-state monitoring. This study evaluates a normalized frequency-domain descriptor for controlled same-vehicle comparison. Vehicle-pass signals underwent resampling, pre-emphasis, framing, and windowing; event-level power spectral density from 500 to 5000 Hz was normalized and partitioned into 500–1000, 1000–3000, and 3000–5000 Hz bands. Field measurements comprised 60 passes of one 4×2 box truck at 20, 30, and 40 km/h with payloads of approximately 1000 and 4000 kg. Heavy load consistently increased the low-frequency proportion and decreased the high-frequency proportion. Two-way analysis of the additive log-ratio $\ln(P_L/P_H)$ identified significant speed, load, and speed $\times$ load effects, with heavy–light separation decreasing as speed increased. A CarSim template was reparameterized at total masses of 9540 and 12,540 kg and a wheelbase of 7.15 m. Wheel-force impulses were convolved with a mass-dependent modal response to generate uncalibrated relative acoustic signals. None of the 48 sensitivity evaluations reproduced the full field direction. Thus, the descriptor supports same-vehicle comparison under controlled conditions, whereas the current dynamics-to-acoustic model is a reproducible discrepancy analysis rather than a calibrated representation of field sound radiation.

Keywords: vehicle load state; acoustic signature; normalized power spectrum; frequency-domain energy ratio; same-vehicle comparison; non-contact monitoring



Academic Editor: John S. Allen

Received: 26 June 2026

Revised: 21 July 2026

Accepted: 22 July 2026

Published: 28 July 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Monitoring the load state of freight vehicles is important for transport safety, overload supervision, and fleet dispatching in enclosed industrial parks, logistics yards, mining roads, and port-transport scenarios. Conventional contact-based weighing devices, such as static scales and weigh-in-motion systems, can provide high measurement accuracy, but their deployment may involve high construction and maintenance costs, limited layout flexibility, and potential effects on traffic efficiency [1,2]. Vehicle load information is also closely related to pavement structural response, vehicle operating safety, and traffic management,

and recent studies have explored roadside sensors, pavement vibration responses, and camera-assisted information to infer load-related characteristics [3–6]. Therefore, low-cost and non-contact approaches for load-state sensing are of practical interest.

Compared with contact-based weighing and vision-based methods, roadside acoustic sensing offers low hardware cost, flexible deployment, and robustness to illumination variations. Acoustic methods have been applied to vehicle detection, speed estimation, and traffic monitoring [7–10]. A moving vehicle generates tire–road interaction noise, engine and driveline noise, body-structure vibration, and local impact sounds. When the vehicle passes over a speed bump or local road irregularity, a change in payload modifies the coupled vibration–acoustic response. The resulting spectrum may contain contributions from sprung-mass motion, suspension and tire forces, structural radiation, and local impacts; however, assigning a particular acoustic band to one mechanism requires synchronized structural and acoustic measurements. Off-road obstacle impacts and whole-body vibration studies further show that road excitation, structural design, and suspension or seat characteristics strongly influence vehicle vibration exposure and response [11,12]. Accordingly, frequency-domain energy distribution is treated here as an empirical and interpretable descriptor for controlled same-vehicle comparison rather than as a direct measurement of an individual vibration mechanism.

Existing vehicle-acoustics studies mainly focus on vehicle detection, vehicle-type recognition, traffic-noise assessment, acoustic event detection, or data-driven sound classification [13–16]. In load-state analysis, however, direct use of sound pressure level or raw spectral magnitude can be sensitive to vehicle speed, microphone distance, device gain, and background noise. A compact and interpretable feature representation is therefore needed to describe the energy redistribution caused by changes in vehicle load state.

The scope of this study is not to build a universal load-classification model applicable to all vehicles, roads, and speeds. Instead, it targets relatively enclosed scenarios in which vehicles repeatedly travel along fixed routes and pass the same excitation location. In such settings, vehicle type, trajectory, road-excitation position, and sensor placement can be relatively stable, making same-vehicle comparison a more defensible application framework. The main contributions of this paper are as follows: (1) an event-level normalized frequency-domain feature extraction method is developed for vehicle-pass acoustic signals; (2) the 500–5000 Hz analysis band is partitioned into low-, mid-, and high-frequency sub-bands to form a three-dimensional compositional feature vector; (3) a balanced field experiment with 60 passes evaluates light- and heavy-load states at 20, 30, and 40 km/h using confidence intervals, log-ratio analysis, and a speed $\times$ load model; (4) a reproducible CarSim-to-acoustic convolution procedure is evaluated against the field direction; and (5) sensitivity analysis and comparison with conventional spectral descriptors and mel-frequency cepstral coefficients are used to characterize the robustness and limitations of the synthesized response.

2. Acoustic Signal Preprocessing and Feature Extraction

2.1. Preprocessing of Vehicle Acoustic Signals

The preprocessing workflow consists of downsampling, pre-emphasis, framing, and windowing. First, the original acoustic recording, typically sampled at 44.1 kHz, is downsampled to 16 kHz to reduce data redundancy and computational cost while retaining the frequency range relevant to the subsequent analysis. Second, a first-order high-pass pre-emphasis filter is applied to compensate for high-frequency attenuation and improve spectral balance:

$$H(z) = 1 - \alpha z^{-1}, \quad (1)$$

where the pre-emphasis coefficient is set to $\alpha = 0.96$ in this study. Third, the non-stationary vehicle-pass signal is divided into short-time frames. The frame length is set to 30 ms and the frame shift to 10 ms; at a sampling frequency of 16 kHz, these settings correspond to 480 and 160 samples, respectively. Finally, each frame is multiplied by a Hann window to reduce spectral leakage caused by truncation and to improve the continuity at both frame boundaries.

2.2. Power Spectral Density Estimation

The acoustic response generated by a truck passing over a speed bump is transient and non-stationary. To obtain a compact representation of its frequency-domain energy distribution, the preprocessed frame signal is transformed into the frequency domain using the discrete Fourier transform (DFT). Let $x_m(n)$ denote the m th windowed frame, where $N = 480$ is the frame length. The DFT is defined as

$$X_m(k) = \sum_{n=0}^{N-1} x_m(n) e^{-j2\pi nk/N}, \quad k = 0, 1, \dots, N-1. \quad (2)$$

The one-sided short-time power spectral density (PSD) of the m th frame is calculated as

$$P_m(k) = \frac{c_k}{f_s \sum_{n=0}^{N-1} w^2(n)} |X_m(k)|^2, \quad (3)$$

where $w(n)$ is the Hann window and $c_k = 2$ for positive-frequency bins except the Nyquist bin, while $c_k = 1$ for the direct-current and Nyquist bins. This formulation explicitly accounts for window energy and one-sided spectral scaling. The frequency-bin width is $\Delta f = f_s/N$. For one complete vehicle-pass event containing M frames, the event-level average PSD is obtained by temporal averaging:

$$\bar{P}(k) = \frac{1}{M} \sum_{m=1}^M P_m(k). \quad (4)$$

The discrete frequency index k is mapped to physical frequency as

$$f_k = \frac{k f_s}{N}, \quad (5)$$

where $f_s = 16,000$ Hz is the sampling frequency.

2.3. In-Band Power-Spectrum Normalization

In roadside acoustic acquisition, the absolute signal amplitude can be affected by vehicle speed, microphone distance, device gain, and background noise. Directly using raw spectral magnitude may therefore introduce non-load-related interference. This study normalizes the average PSD within the core analysis band of 500–5000 Hz so that the feature represents a relative energy distribution. Let K_{total} denote the set of discrete frequency indices within 500–5000 Hz. The normalized spectrum is defined as

$$p(k) = \frac{\bar{P}(k) \Delta f}{\sum_{j \in K_{\text{total}}} \bar{P}(j) \Delta f}, \quad k \in K_{\text{total}}. \quad (6)$$

After normalization, the in-band energy ratios satisfy

$$\sum_{k \in K_{\text{total}}} p(k) = 1. \quad (7)$$

Thus, the analysis is shifted from absolute acoustic amplitude to the relative distribution of energy across frequencies, which is more appropriate for same-vehicle comparison under similar acquisition conditions.

3. Same-Vehicle Load-State Difference Analysis Method

3.1. Normalized Frequency-Domain Feature Vector

After obtaining the normalized spectrum $p(k)$, the 500–5000 Hz analysis band is divided into three sub-bands: low frequency (500–1000 Hz), mid frequency (1000–3000 Hz), and high frequency (3000–5000 Hz). The corresponding index sets are denoted by K_L , K_M , and K_H , respectively. Their acoustic interpretations are summarized in Table 1.

Table 1. Frequency-band partitioning and possible contributing acoustic sources.

Feature	Frequency Range	Index Set	Possible Contributing Sources
Low-frequency ratio P_L	500–1000 Hz	K_L	Candidate contribution from body, suspension, cargo-box, and other low-order structural responses; not treated as a uniquely identified structural mode.
Mid-frequency ratio P_M	1000–3000 Hz	K_M	Tire–road interaction, tread compression, air-pumping effects, and partial mechanical radiation; used as an auxiliary band.
High-frequency ratio P_H	3000–5000 Hz	K_H	Candidate contribution from local impacts, mechanical clearances, tire–road transients, and other high-frequency sources.

The normalized energy ratios of the three bands are calculated as

$$P_L = \sum_{k \in K_L} p(k), \quad P_M = \sum_{k \in K_M} p(k), \quad P_H = \sum_{k \in K_H} p(k), \quad (8)$$

forming the feature vector

$$\mathbf{V}_{\text{freq}} = [P_L, P_M, P_H]. \quad (9)$$

Compared with raw sound pressure level or absolute spectral magnitude, this vector emphasizes the redistribution of acoustic energy among frequency regions and is therefore suitable for analyzing load-associated changes in the same vehicle. Because $P_L + P_M + P_H = 1$, the three components are compositional rather than statistically independent; an increase in one component necessarily decreases at least one other component. The present analysis therefore interprets the vector jointly and does not treat the individual ratios as independent causal indicators.

3.2. Relative Energy Redistribution Between Light and Heavy Loads

A change in vehicle load modifies the coupled dynamic response when the vehicle passes over a localized road excitation. Changes in body motion, suspension force, tire force, structural damping, and local contact can all affect the radiated sound. A shift toward low- or high-frequency energy may therefore be associated with a change in load, but the acoustic ratios alone do not identify which structural mechanism produced that shift.

For a light-load and heavy-load pair of the same vehicle under similar speed and acquisition conditions, the relative change coefficient of each band is defined as

$$\Delta P_i = \frac{P_{i,\text{heavy}} - P_{i,\text{light}}}{P_{i,\text{light}}}, \quad i \in \{L, M, H\}, \quad (10)$$

where $P_{i,\text{heavy}}$ and $P_{i,\text{light}}$ are the normalized energy ratios of the i th band under heavy and light loads, respectively. Positive or negative values describe relative redistribution within the selected band only; they do not establish an independent increase or decrease in absolute radiated energy. Table 2 summarizes the descriptive interpretation used in this study.

Table 2. Qualitative interpretation of frequency-domain energy redistribution.

Feature Pattern	Possible Contributing Sources	Load-State Tendency
P_L increases and P_H decreases	Relative spectral energy shifts toward the low-frequency band; synchronized vibration data are required for a structural attribution.	Empirical heavy-load tendency in the present field data
Lower P_L and higher P_H	Relative spectral energy shifts toward the high-frequency band; several impact and radiation mechanisms may contribute.	Empirical light-load tendency in the present field data
Higher P_M	Mid-band sources contribute a larger fraction of the selected 500–5000 Hz energy.	Auxiliary descriptive evidence

This method does not set a single fixed threshold for all vehicles, speeds, and environments. The frequency distribution of vehicle sound can be affected by speed, vehicle type, tire condition, pavement roughness, microphone distance, background noise, and acoustic propagation path. Accordingly, the proposed analysis emphasizes relative differences between light- and heavy-load samples under comparable conditions rather than absolute classification based on a universal threshold.

4. Experimental Results and Analysis

4.1. Experimental Design and Data Sources

The experiments comprise simulated and field-acquired acoustic data. The simulation component uses a parameterized 4×2 box-truck model passing over a localized road excitation at target speeds of 20, 30, and 40 km/h. The CarSim template was reparameterized using the field vehicle's 7.15 m wheelbase and 8540 kg curb mass. Payloads of approximately 1000 and 4000 kg therefore gave total simulated masses of 9540 and 12,540 kg, respectively. The corresponding center-of-gravity heights were engineering assumptions of 1.05 and 1.125 m because condition-specific field measurements were unavailable. The road input was a 0.075 m high and 0.50 m long half-sine speed-bump profile. The selected CarSim solver template was reparameterized to represent these vehicle-scale properties; its suspension, tire, inertia, and center-of-gravity parameters were not manufacturer-calibrated. Each of the six conditions was simulated once, and the native CarSim outputs were stored at 1000 Hz. The field component used the photographed 4×2 box truck traversing the same excitation location at nominal speeds of 20, 30, and 40 km/h. The field payloads were approximately 1000 kg (light load) and 4000 kg (heavy load), rather than total vehicle masses. The microphone was positioned adjacent to the edge of the speed bump, with its sensing height 0.10 m above the road surface. Ten independent vehicle-pass feature records were retained for each speed–load combination, giving 60 pass-level observations in a balanced 3×2 design. All field records were processed using the same 16 kHz resampling, $\alpha = 0.96$ pre-emphasis, 30 ms Hann window, 10 ms frame shift, event-level PSD averaging, and 500–5000 Hz normalization procedure. Figure 1 shows the experimental workflow.

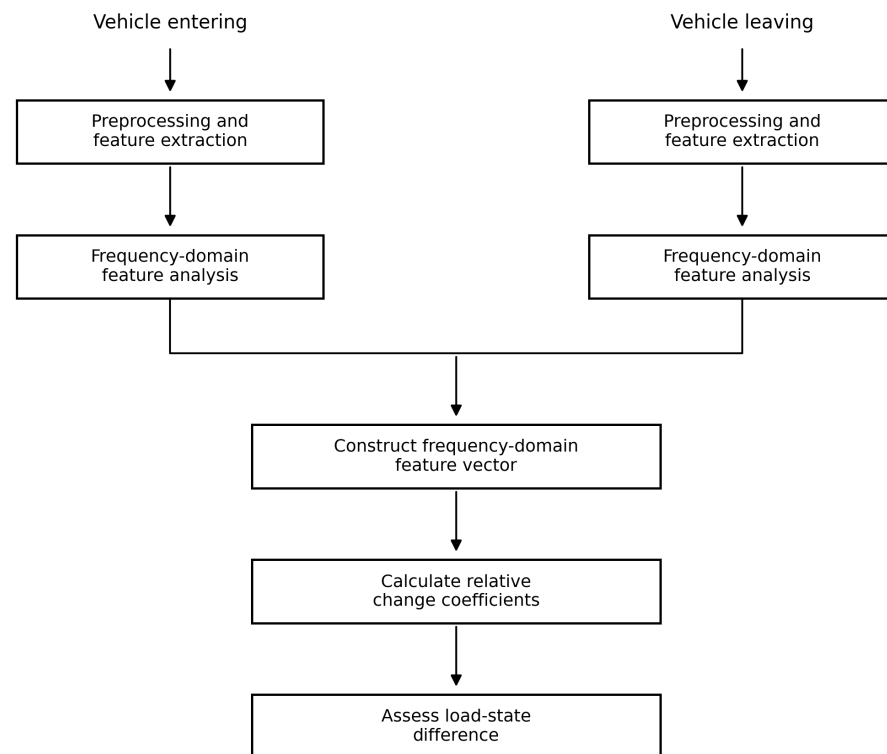


Figure 1. Experimental workflow for same-vehicle load-state difference analysis based on acoustic features.

4.1.1. CarSim-to-Acoustic Coupling Procedure

The simulated acoustic signals are not direct CarSim audio outputs. Body vertical acceleration, pitch rate, and the four wheel vertical forces were exported from CarSim. For each wheel-force channel $F_i(t)$, a static baseline $\bar{F}_i$ was estimated from the initial 2 s, and the positive dynamic component was defined as

$$F_i^+(t) = \max[F_i(t) - \bar{F}_i, 0]. \quad (11)$$

Candidate impacts were detected using a median-plus-five-median-absolute-deviation threshold, with a minimum separation of 0.08 s and a maximum of eight events per wheel channel. The impulse of event e was integrated over ± 0.035 s around the detected peak:

$$I_e = \int_{t_e - 0.035}^{t_e + 0.035} F_i^+(t) dt. \quad (12)$$

The discrete excitation sequence was $q(t) = \sum_e I_e \delta(t - t_e)$. A mass-dependent modal impulse response was then defined as

$$h(t; m) = \sum_{r=1}^5 G_r(m) e^{-\zeta_r \omega_r(m) t} \sin \left[\omega_r(m) \sqrt{1 - \zeta_r^2} t \right] u(t), \quad (13)$$

where $u(t)$ is the unit-step function. With the 9540 kg case as the reference mass m_0 , modal frequency and gain were scaled as $f_r(m) = f_{r0}(m_0/m)^{a_r}$ and $G_r(m) = G_{r0}(m_0/m)^{b_r}$. The synthesized pressure was obtained by convolution:

$$p_{\text{sim}}(t) = q(t) * h(t; m). \quad (14)$$

Table 3 lists the fixed reference parameters used for all six simulations. These parameters provide a transparent physics-informed synthesis model, but the gains have not been

calibrated using a hammer test or synchronized accelerometer–microphone measurements. No empirical background-noise model was added to $p_{\text{sim}}(t)$; the synthesized waveform therefore represents only the deterministic response of the assumed force-to-pressure model. Consequently, $p_{\text{sim}}(t)$ is used for within-model relative comparison and not as an absolute sound-pressure prediction.

Table 3. Reference modal parameters of the physics-informed acoustic impulse response ($m_0 = 9540$ kg).

Mode	f_{r0} (Hz)	ζ_r	G_{r0} [Pa/(N · s)]	(a_r, b_r)
1	85	0.120	1.50×10^{-6}	(0.50, 0.50)
2	220	0.100	1.00×10^{-6}	(0.50, 0.50)
3	720	0.080	5.50×10^{-7}	(0.30, 0.35)
4	1650	0.060	3.00×10^{-7}	(0.10, 0.20)
5	3400	0.045	1.50×10^{-7}	(0.00, 0.10)

4.1.2. Field Statistical Analysis

The statistical analysis was based on pass-level frequency-band features obtained from the real roadside truck experiment. For each field speed–load condition, the three band proportions are reported as the mean $\pm$ sample standard deviation (SD) and the 95% confidence interval (CI) of the mean, calculated using the Student- t distribution ($n = 10$). Because P_L , P_M , and P_H sum to unity and are therefore compositional, the additive log-ratio coordinate

$$\text{ALR}_{L/H} = \ln\left(\frac{P_L}{P_H}\right) \quad (15)$$

was specified as the primary inferential outcome. A categorical two-way analysis of variance (ANOVA) was fitted with speed, load, and their interaction as fixed effects; pass-level observations were treated as experimental repetitions. Partial η^2 was reported as the effect-size measure. The homogeneity-of-variance assumption was acceptable for the ALR outcome (median-centered Levene test, $p = 0.465$). A reliable one-to-one pairing identifier between the light- and heavy-load passes was not retained. Speed-specific load contrasts were therefore treated as independent and evaluated using Welch tests, with Holm adjustment across the three speeds, rather than imposing an unverified paired structure. The individual band proportions were retained as descriptive and secondary outcomes rather than interpreted as statistically independent responses.

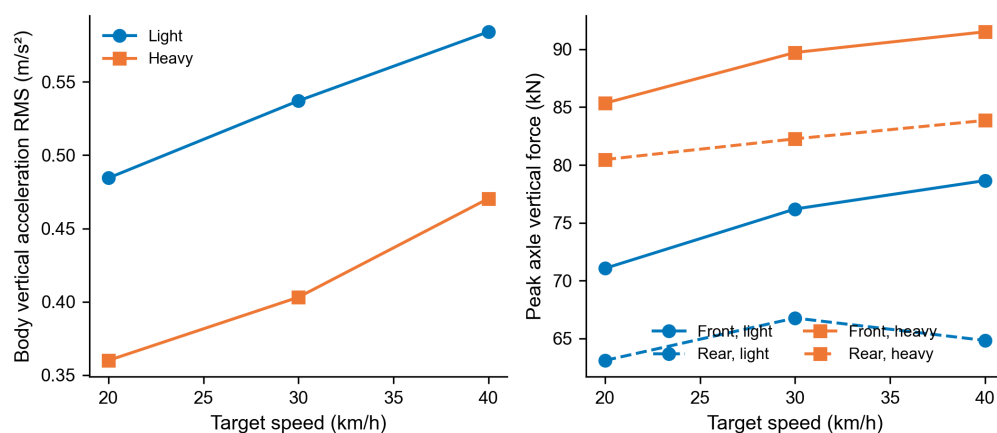
4.2. Simulated Acoustic Response Analysis

4.2.1. Vehicle-Dynamic Response

Table 4 summarizes the CarSim response while the vehicle traversed the speed bump. The achieved mean speeds were within 0.05 km/h of the target values. Heavy load increased the absolute front- and rear-axle force peaks, whereas its body vertical-acceleration RMS and peak values were lower than those of the corresponding light-load condition. No front- or rear-axle contact loss occurred in the six nominal runs. This result shows that a larger static and impact force does not necessarily produce a larger body-acceleration response, and it cautions against assigning an acoustic band directly to sprung-mass acceleration. Figure 2 visualizes the corresponding body-acceleration and axle-force measures for the six conditions.

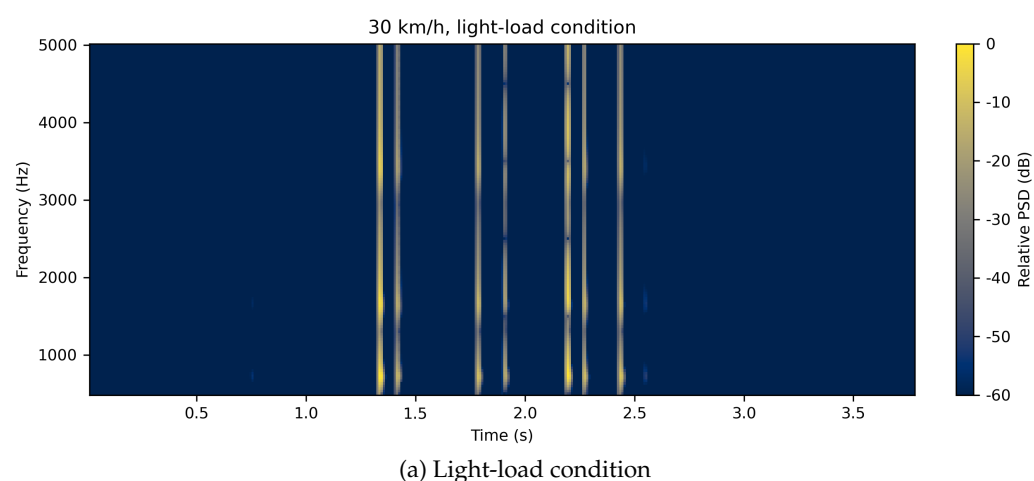
Table 4. CarSim dynamic-response summary for the six speed–load conditions.

Condition	Mean Speed (km/h)	Body Acc. RMS (m/s ²)	Body Acc. Peak (m/s ²)	Front-Axle Peak (kN)	Rear-Axle Peak (kN)
20 km/h, light load	19.985	0.484	2.746	71.07	63.10
20 km/h, heavy load	19.988	0.360	2.083	85.35	80.47
30 km/h, light load	30.002	0.537	3.008	76.20	66.77
30 km/h, heavy load	30.028	0.403	2.219	89.73	82.26
40 km/h, light load	39.978	0.584	3.006	78.65	64.83
40 km/h, heavy load	39.960	0.470	2.593	91.52	83.85

**Figure 2.** CarSim body vertical-acceleration RMS and peak axle vertical forces under the six nominal conditions.

4.2.2. Time–Frequency Spectrograms

To examine the time–frequency structure of the synthesized vehicle-pass sound, STFT spectrograms were generated at 30 km/h under light and heavy loads using the same pre-emphasis, frame length, and frame shift as the field-signal pipeline. Figure 3 shows discrete broadband responses at the detected wheel-force impulse times. Differences in event timing and decay are visible, but the spectrograms alone do not establish a load-specific radiation mechanism because the modal transfer function is uncalibrated.

**Figure 3.** Cont.

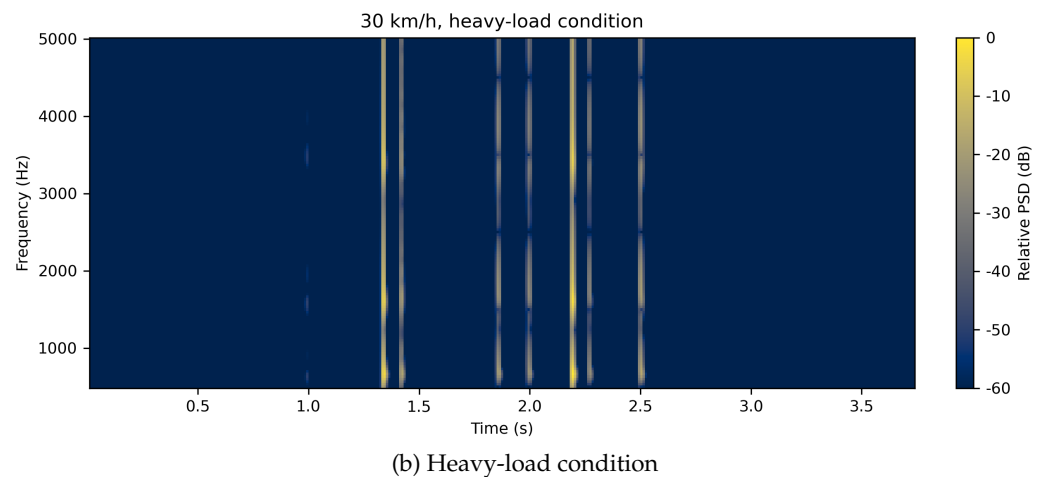


Figure 3. Time-frequency spectrograms of the same vehicle passing over the speed bump at 30 km/h: (a) light-load condition; (b) heavy-load condition.

The spectrograms provide qualitative evidence only and are not used as a standalone quantitative criterion. Therefore, the following subsection further extracts normalized energy ratios from low-, mid-, and high-frequency bands.

4.2.3. Frequency-Domain Feature Variation Under Different Speeds

Figure 4 directly overlays the normalized power spectral distributions for the light- and heavy-load simulations at 30 km/h. The common axes and distinct line styles make the load-associated differences visible without relying on comparisons between separate panels. Table 5 lists the three-band feature vectors under all simulated speed and load conditions.

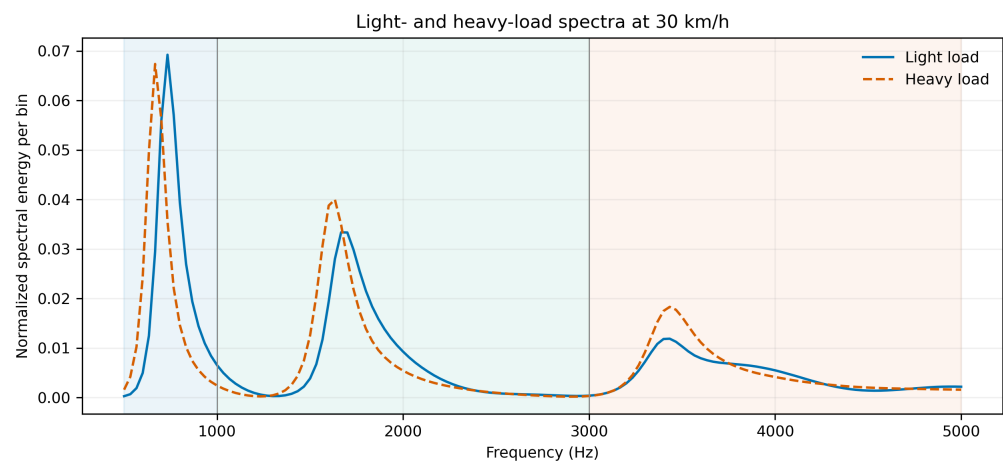


Figure 4. Overlay of the normalized power spectral distributions for the synthesized light- and heavy-load cases at 30 km/h. The light-load spectrum is shown as a blue solid line and the heavy-load spectrum as an orange dashed line; shaded regions denote the 500–1000, 1000–3000, and 3000–5000 Hz analysis bands.

The newly generated signals were processed with the same $\alpha = 0.96$ pre-emphasis, 30 ms Hann-windowed frames, 10 ms frame shift, and event-level average PSD used for the field recordings. The overlaid curves have similar principal peak regions, but the heavy-load spectrum is relatively lower in the 500–1000 Hz band and higher in the 3000–5000 Hz band. At 20 km/h, P_L decreased from 34.88% under light load to 28.58% under heavy load, while P_H increased from 26.81% to 32.46%. The same direction occurred at 30 and 40 km/h. Thus, the present uncalibrated physics-informed synthesis does not reproduce the field-data tendency of higher P_L and lower P_H under heavy load.

Table 5. Normalized frequency-domain features under different simulated speeds.

Condition	P_L (500–1000 Hz)	P_M (1000–3000 Hz)	P_H (3000–5000 Hz)	Feature Vector
20 km/h, light load	34.88%	38.31%	26.81%	[0.3488, 0.3831, 0.2681]
20 km/h, heavy load	28.58%	38.96%	32.46%	[0.2858, 0.3896, 0.3246]
30 km/h, light load	35.15%	38.07%	26.77%	[0.3515, 0.3807, 0.2677]
30 km/h, heavy load	31.64%	37.41%	30.95%	[0.3164, 0.3741, 0.3095]
40 km/h, light load	33.55%	37.64%	28.80%	[0.3355, 0.3764, 0.2880]
40 km/h, heavy load	31.61%	37.40%	30.98%	[0.3161, 0.3740, 0.3098]

The relative shift coefficients in Table 6 quantify this model–field discrepancy. Because the ratios are compositional, the positive high-frequency shift cannot be interpreted independently of the negative low-frequency shift. The similarity of the 30 and 40 km/h feature vectors also indicates that the normalized spectral shape is governed mainly by the assumed modal transfer function, whereas speed primarily changes event timing and amplitude in the present synthesis.

Table 6. Relative shifts from light-load to heavy-load condition.

Speed	ΔP_L	ΔP_M	ΔP_H	Main Observation
20 km/h	−0.1805	0.0167	0.2109	Synthesized energy shifts from the low to the high band.
30 km/h	−0.0998	−0.0174	0.1558	The same model-dependent shift remains.
40 km/h	−0.0578	−0.0064	0.0758	The high-band increase remains at the highest speed.

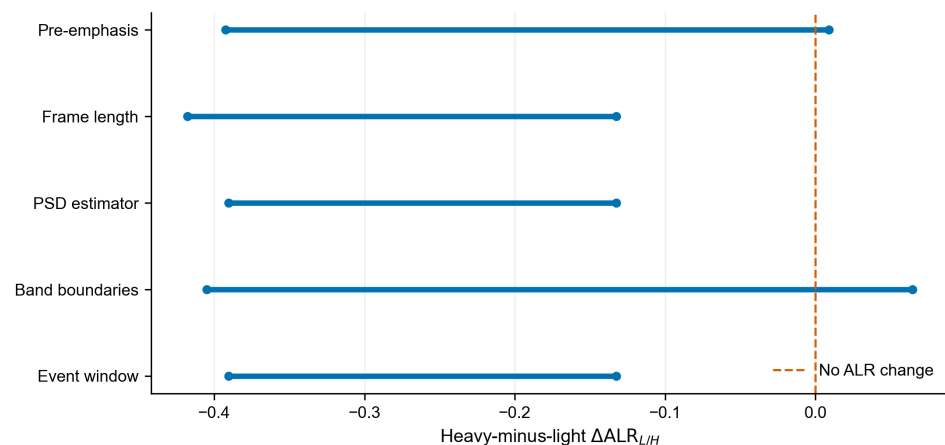
4.2.4. Robustness Checks and Descriptor Comparison

To determine whether the opposite simulated and field trends resulted from a particular signal-processing choice, a one-factor-at-a-time sensitivity analysis was conducted around the reference setting ($\alpha = 0.96$, 30 ms Hann window, 10 ms frame shift, 500–1000/1000–3000/3000–5000 Hz bands, and a 0.75/1.25 s pre-/post-event window). The tested factors were the pre-emphasis coefficient, frame length, PSD estimator, frequency-band boundaries, and event-window duration. For each setting, the heavy-minus-light differences in P_L , P_H , and $\text{ALR}_{L/H}$ were calculated at 20, 30, and 40 km/h. A setting was considered consistent with the field direction only when $\Delta P_L > 0$, $\Delta P_H < 0$, and $\Delta \text{ALR}_{L/H} > 0$. This sensitivity grid was applied to the synthesized signals only. The retained field dataset supports pass-level normalized band analysis but not reprocessing under alternative waveform-level settings; field invariance to these preprocessing choices is therefore not claimed.

None of the 48 speed-by-setting evaluations reproduced the full field direction. All evaluations using the reference pre-emphasis coefficient ($\alpha = 0.96$) remained opposite to the field direction. A small number of alternative pre-emphasis or band-boundary settings produced a positive $\Delta \text{ALR}_{L/H}$, but they did not meet the full criterion because P_L and P_H did not change simultaneously in the required directions. Welch PSD and the frame-averaged periodogram gave nearly identical band ratios. The simulated–field disagreement is therefore not attributable to the PSD estimator and remained under every complete direction criterion tested. Table 7 summarizes the tested factors and outcome ranges, and Figure 5 visualizes the corresponding ranges of $\Delta \text{ALR}_{L/H}$.

Table 7. Sensitivity of the synthesized heavy–light spectral direction to signal-processing choices.

Factor	Tested Levels	Field-Direction Matches	$\Delta\text{ALR}_{L/H}$ Range
Pre-emphasis coefficient	0, 0.90, 0.96, 0.98	0/12	−0.393 to 0.009
Frame length	20, 30, 40, 50 ms	0/12	−0.418 to −0.133
PSD estimator	Frame-averaged periodogram, Welch	0/6	−0.391 to −0.133
Band boundaries	500–800–2500–5000, 500–1000–3000–5000, and 500–1200–3500–5000 Hz	0/9	−0.405 to 0.065
Event window	0.25/0.75, 0.75/1.25, and 1.25/1.50 s pre-/post-event	0/9	−0.391 to −0.133

**Figure 5.** Range of the synthesized heavy-minus-light additive log-ratio, $\Delta\text{ALR}_{L/H}$, under each sensitivity factor. Positive values occurred for a small subset of settings; full field-direction agreement required simultaneous $\Delta P_L > 0$, $\Delta P_H < 0$, and $\Delta\text{ALR}_{L/H} > 0$.

Conventional spectral descriptors and MFCCs were extracted from the same six synthesized event signals as a descriptive comparison. Spectral centroid, 85% roll-off, normalized entropy, and flatness were computed from the 500–5000 Hz Welch spectrum. MFCC extraction used $\alpha = 0.96$ pre-emphasis, 25 ms Hamming frames, a 10 ms frame shift, a 512-point transform, 26 triangular mel filters spanning 300–5000 Hz, and 13 coefficients; Table 8 reports the means of the first five coefficients. Relative to the light-load synthesis, the heavy-load synthesis had a higher mean spectral centroid, roll-off frequency, and spectral entropy, together with a lower $\text{ALR}_{L/H}$ and slightly lower spectral flatness. The centroid, roll-off, and ALR results indicate that the assumed transfer function shifted the synthesized heavy-load spectrum toward higher frequencies. Because each simulated speed–load condition contains only one deterministic realization, no classifier performance or inferential descriptor comparison is claimed. This six-signal comparison is intended only as a model-diagnostic description and does not demonstrate that the proposed descriptor outperforms conventional spectral or cepstral features.

Overall, the CarSim results demonstrate load-dependent differences in axle force and body acceleration, while the acoustic synthesis provides a reproducible coupling procedure from wheel-force impulses to a relative pressure waveform. The reference analysis showed the opposite simulated and field direction, and none of the 48 sensitivity evaluations met the complete field-direction criterion. The simulation is therefore retained as a transparent feasibility and model-discrepancy analysis rather than as independent acoustic validation.

Table 8. Descriptive comparison of conventional and MFCC features for the synthesized signals.

Descriptor	Light-Load Mean	Heavy-Load Mean	Heavy-Light Direction
Spectral centroid (Hz)	1974.5	2054.8	Higher
85% spectral roll-off (Hz)	3622.2	3655.6	Higher
Normalized spectral entropy	0.8520	0.8597	Higher
Spectral flatness	0.4429	0.4426	Lower
ALR _{L/H}	0.2294	−0.0283	Lower
MFCC 1 mean	−295.089	−296.517	Lower
MFCC 2 mean	1.124	1.036	Lower
MFCC 3 mean	0.245	0.234	Lower
MFCC 4 mean	0.035	0.039	Higher
MFCC 5 mean	0.054	0.068	Higher

4.3. Comparison with Real Roadside Truck Measurements

To examine the empirical load-associated pattern separately from the uncalibrated synthesis, the same 4×2 box truck was measured under light and heavy loads at nominal speeds of 20, 30, and 40 km/h. The field vehicle was a 2025-model-year Dongfeng Tianjin KR Plus 4×2 box truck (DFH5170XXYEX16), manufactured by Dongfeng Commercial Vehicle Co., Ltd. (Shiyan, China) in 2026. It had a 9.8 m cargo body and was powered by a Dongfeng DDi75E300-60 7.5 L inline six-cylinder China VI diesel engine rated at 220 kW, coupled to a Fast 8JS125TA manual transmission. Previous road testing of a comparable Dongfeng Tianjin commercial truck characterized cab-suspension and road-induced vibration responses, providing relevant vehicle-dynamic context for this vehicle family [17]. Ten pass-level feature records were retained for each of the six speed-load conditions. Figure 6 shows the field truck, and Table 9 reports the condition-level frequency-band statistics. Because the field and simulation domains differ and the acoustic transfer model was not calibrated to the measurement setup, their comparison is used to assess directional agreement rather than to claim model validation.

Table 9. Pass-level field frequency-band results for the six speed-load conditions. Values are mean $\pm$ SD [95% CI], with $n = 10$ passes per condition.

Speed	Load	P_L (500–1000 Hz)	P_M (1000–3000 Hz)	P_H (3000–5000 Hz)
20 km/h	Light	2.91 ± 0.08 [2.86, 2.96]%	76.80 ± 0.50 [76.44, 77.16]%	20.29 ± 0.51 [19.93, 20.65]%
20 km/h	Heavy	39.38 ± 1.22 [38.51, 40.26]%	56.37 ± 1.18 [55.52, 57.22]%	4.25 ± 0.14 [4.15, 4.35]%
30 km/h	Light	2.53 ± 0.08 [2.48, 2.58]%	73.72 ± 0.85 [73.11, 74.33]%	23.75 ± 0.83 [23.16, 24.34]%
30 km/h	Heavy	36.88 ± 1.05 [36.13, 37.63]%	56.94 ± 1.14 [56.12, 57.76]%	6.18 ± 0.17 [6.06, 6.30]%
40 km/h	Light	2.08 ± 0.05 [2.04, 2.12]%	70.35 ± 1.05 [69.60, 71.10]%	27.57 ± 1.04 [26.83, 28.31]%
40 km/h	Heavy	32.89 ± 0.82 [32.30, 33.48]%	58.11 ± 0.87 [57.49, 58.73]%	9.00 ± 0.31 [8.78, 9.22]%

The heavy-load condition consistently showed a markedly larger 500–1000 Hz proportion and a smaller 3000–5000 Hz proportion than the light-load condition at all three speeds. Under light load, P_L decreased from 2.91% at 20 km/h to 2.08% at 40 km/h, whereas P_H increased from 20.29% to 27.57%. Under heavy load, P_L decreased from 39.38% to 32.89%, while P_H increased from 4.25% to 9.00%. Thus, increasing speed shifted the normalized field spectrum toward the high-frequency band under both load states, although the load-associated separation remained visible at 40 km/h.

Figure 7 shows the condition means with 95% CIs. The nonparallel ALR profiles indicate that the magnitude of the load contrast changed with speed. The heavy-minus-light ALR difference declined from 4.168 at 20 km/h to 4.026 at 30 km/h and 3.880 at 40 km/h, but every contrast remained positive and its CI excluded zero (Table 10; Holm-adjusted $p < 0.001$ at every speed).



Figure 6. The 2025-model Dongfeng Tianjin KR Plus 4×2 box truck (DFH5170XXYEX16), manufactured in 2026, powered by a Dongfeng DDi75E300-60 diesel engine, and used in the roadside acoustic experiment.

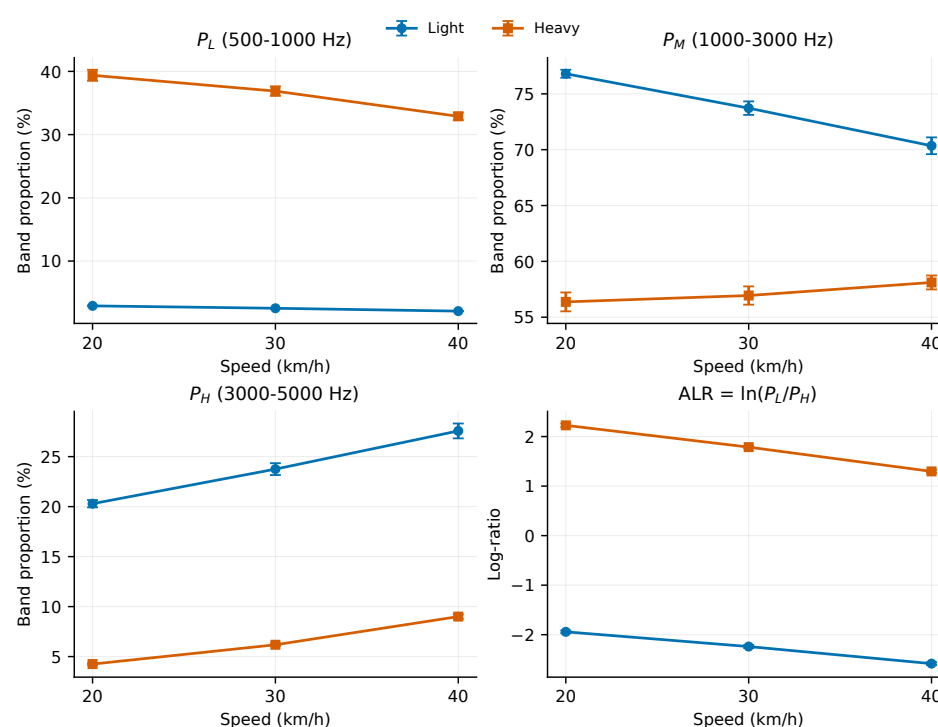


Figure 7. Field frequency-band proportions and additive log-ratio across speed and load conditions. Points denote pass-level condition means and error bars denote 95% confidence intervals ($n = 10$ per condition).

Table 10. Speed-specific heavy-minus-light contrasts for $ALR_{L/H} = \ln(P_L/P_H)$. Confidence intervals and tests are based on independent pass-level observations; p values were adjusted by the Holm method.

Speed	Mean Difference	95% CI	Holm-Adjusted p
20 km/h	4.168	[4.125, 4.212]	<0.001
30 km/h	4.026	[3.991, 4.060]	<0.001
40 km/h	3.880	[3.841, 3.919]	<0.001

The two-way ANOVA on the ALR outcome identified significant effects of speed ($F_{2,54} = 1788.48$, $p < 0.001$, partial $\eta^2 = 0.985$), load ($F_{1,54} = 140468.89$, $p < 0.001$, partial $\eta^2 > 0.999$), and the speed $\times$ load interaction ($F_{2,54} = 60.03$, $p < 0.001$, partial $\eta^2 = 0.690$). The interaction confirms that speed altered the magnitude of the load-associated low-to-high frequency redistribution: the heavy–light separation weakened progressively with increasing speed rather than remaining constant across the three operating conditions. The field direction is opposite to the present uncalibrated modal-convolution result, indicating that the assumed force-to-pressure transfer function does not yet represent the combined structural-radiation, propagation, background-noise, and microphone effects of the field experiment.

5. Discussion

The proposed method provides an interpretable representation for same-vehicle load-state comparison. The expanded field experiment replaces the former small set of measurements at one nominal speed with 60 pass-level observations covering 20, 30, and 40 km/h. The load-associated direction was reproduced at every tested speed: heavy load was associated with a larger 500–1000 Hz proportion, a smaller 3000–5000 Hz proportion, and a positive $ALR_{L/H}$ contrast. Because the three ratios sum to unity, they must be interpreted jointly. The narrow within-condition confidence intervals indicate that the pattern was repeatable under the controlled acquisition conditions used in this experiment. However, the acoustic data alone cannot assign the pattern uniquely to sprung-mass motion, suspension response, mechanical clearance, or another physical source. References to body vibration, suspension response, tire force, and structural radiation therefore denote possible contributing sources only and should not be interpreted as identified band-specific mechanisms.

The significant speed $\times$ load interaction also shows that the descriptor is not speed invariant. The low-to-high frequency separation between load states decreased as speed increased. Accordingly, a practical decision rule should either control for speed or include speed as an explicit model variable. For an operational implementation, a vehicle- and site-specific light-load reference distribution could first be established within defined speed intervals. A pass could then be flagged when its ALR value exceeds a speed-specific threshold or the prediction limit of a speed-adjusted reference model. The present dataset supports this conceptual approach but is insufficient to select and independently validate a deployment threshold. The results therefore strengthen controlled same-vehicle comparison but do not establish a vehicle-independent load classifier or performance under unobserved roads, vehicles, sensor arrangements, or environments.

The vehicle-dynamic simulation adds a separate physical perspective. Heavy load produced higher absolute axle-force peaks but lower body vertical acceleration. After the detected wheel-force impulses were passed through the assumed modal transfer function, the reference synthesized heavy-load signals had lower P_L , higher P_H , a higher spectral centroid, and a higher spectral roll-off frequency, opposite to the field trend. None of the 48 sensitivity evaluations met the complete field-direction criterion. A reproducible force-to-pressure convolution is therefore not equivalent to an acoustically calibrated

vehicle model, and the current simulation is better interpreted as a transparent model-discrepancy analysis.

The analysis does not support a classification-performance claim. Although conventional descriptors and MFCC means provide a useful model-diagnostic comparison, they were calculated from only six deterministic synthesized signals. They therefore cannot support claims of classifier accuracy or comparative feature superiority. The 60 field observations support uncertainty estimates and formal speed–load inference for the proposed compositional descriptor, but all measurements still concern one truck at one excitation location.

Several limitations remain. Exact field axle loads, suspension and tire states, microphone sensitivity and calibration, microphone orientation, ambient-noise level, weather, pavement state, and actual pass-speed distributions were not retained for every observation. Calibrated absolute-amplitude field data were not available for a defensible comparison of non-normalized band energy. The archived field dataset contains pass-level normalized frequency-band features rather than complete raw audio. Independent waveform-level reprocessing under alternative preprocessing settings therefore cannot be performed. This does not affect the reported statistical analysis, which is based on the retained features from the real field passes, but it limits reproducibility of the complete waveform-processing pipeline. Although the vehicle model, payload conditions, microphone position, wheelbase, and test masses are now documented, each CarSim condition is a single nominal run. The CarSim model remains an engineering parameterization rather than a manufacturer-calibrated representation of the DFH5170XXYEX16, and the tire-model logs contain extrapolation warnings at the imposed wheel loads. Finally, the modal frequencies, damping ratios, and gains require identification using a hammer test or synchronized force, structural-acceleration, and microphone measurements before causal or absolute sound-pressure claims can be made.

6. Conclusions

This study developed a normalized frequency-domain acoustic descriptor for comparing load states of the same vehicle under controlled operating conditions. In-band PSD normalization converts the 500–5000 Hz spectrum into the compositional vector $\mathbf{V}_{\text{freq}} = [P_L, P_M, P_H]$. Across 60 field passes at 20, 30, and 40 km/h, heavy load consistently produced a larger 500–1000 Hz fraction and a smaller 3000–5000 Hz fraction than light load. Log-ratio analysis identified significant speed, load, and speed $\times$ load effects, with the heavy–light low-to-high frequency separation decreasing as speed increased.

The main contribution of this study is an evidence-bounded framework that distinguishes an empirical load-sensitive acoustic signature from an unvalidated physical interpretation. The framework combines a compositional three-band spectral descriptor, replicated speed-by-load field measurements, and an explicit simulation–field discrepancy analysis. The field results establish a repeatable within-vehicle redistribution of normalized spectral energy and quantify its speed dependence through the additive log-ratio. In contrast, the opposite direction produced by the CarSim-to-acoustic synthesis, with no complete field-direction match in 48 sensitivity evaluations, shows that wheel-force excitation alone is insufficient to reproduce the roadside spectrum without calibrated structural radiation, propagation, background-noise, and microphone-response components. This distinction clarifies both the information carried by the normalized acoustic descriptor and the physical mechanisms that remain unresolved.

Accordingly, the proposed descriptor should be interpreted as an exploratory same-vehicle monitoring feature under controlled acquisition conditions, with speed controlled or explicitly modeled, rather than as a universal load classifier or calibrated acoustic-radiation model. The findings have not yet been generalized across vehicle types, suspension

configurations, pavement conditions, or sensor layouts. Future work should evaluate additional vehicles and load levels, retain complete raw audio and calibration metadata, and use synchronized force, structural-acceleration, and calibrated acoustic measurements to identify the force-to-pressure transfer relationship [18–21].

Author Contributions: Conceptualization, Y.P. and H.Z.; methodology, Y.P.; validation, Y.P.; formal analysis, Y.P., Z.Y. and Y.L.; investigation, Y.P.; data curation, Z.Y. and Y.L.; writing—original draft preparation, Y.P., Z.Y. and Y.L.; writing—review and editing, Y.P., Z.Y., Y.L. and H.Z.; supervision, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The CarSim input and output files, extracted dynamics, detected impact events, physics-informed pressure signals, synthesized audio files, processing scripts, and pass-level normalized frequency-band features derived from the real roadside experiments are available from the corresponding author upon reasonable request. The archived field dataset for this revision consists of the pass-level normalized frequency-band features used in the statistical analysis. Complete raw field audio was not retained and is therefore unavailable for independent waveform-level reprocessing under alternative signal-processing settings.

Acknowledgments: The authors thank the field experiment participants for their assistance with data collection.

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

Abbreviations

The following abbreviations are used in this manuscript:

DFT	Discrete Fourier Transform
PSD	Power Spectral Density
STFT	Short-Time Fourier Transform
ALR	Additive Log-Ratio
MFCC	Mel-Frequency Cepstral Coefficient

References

1. Burnos, P.; Gajda, J.; Sroka, R.; Wasilewska, M.; Dolega, C. High Accuracy Weigh-In-Motion Systems for Direct Enforcement. *Sensors* **2021**, *21*, 8046. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Yang, H.; Yang, Y.; Hou, Y.; Liu, Y.; Liu, P.; Wang, L.; Ma, Y. Investigation of the Temperature Compensation of Piezoelectric Weigh-In-Motion Sensors Using a Machine Learning Approach. *Sensors* **2022**, *22*, 2396. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Gajda, J.; Sroka, R.; Burnos, P.; Daniol, M. Long-Term Assessment of the Properties of Load Sensors Applied in Weigh-In-Motion Systems. *Sensors* **2025**, *25*, 2421. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Hagmanns, M.P.M.; Lamberty, S.; Fazekas, A.; Oeser, M. Enhancing Weigh-in-Motion Systems Accuracy by Considering Camera-Captured Wheel Oscillations. *Sensors* **2024**, *24*, 8151. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Zhao, Q.; Ye, Z.; Tan, Z.; Xu, J.; Wang, L. Vehicle Load Information Acquisition Using Roadside Micro-Electromechanical Systems Accelerometers. *Sensors* **2025**, *25*, 4901. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Li, P.; Wang, L.; Yang, S.; Ye, Z. Three-Axis Vibration Characteristics and Vehicle Axle Shape Identification of Cement Pavement under Heavy Vehicle Loads Based on EMD–Energy Decoupling Method. *Sensors* **2025**, *25*, 4066. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Szwoch, G.; Kotus, J. Acoustic Detector of Road Vehicles Based on Sound Intensity. *Sensors* **2021**, *21*, 7781. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Kotus, J.; Szwoch, G. Estimation of Average Speed of Road Vehicles by Sound Intensity Analysis. *Sensors* **2021**, *21*, 5337. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Corera, I.; Piñeiro, E.; Navallas, J.; Sagues, M.; Loayssa, A. Long-Range Traffic Monitoring Based on Pulse-Compression Distributed Acoustic Sensing and Advanced Vehicle Tracking and Classification Algorithm. *Sensors* **2023**, *23*, 3127. [\[CrossRef\]](#) [\[PubMed\]](#)

10. Deng, J.; Jin, L.; Wang, H.; Zhang, Z.; Liu, Y.; Meng, F.; Wang, J.; Li, Z.; Wu, J. Distributed Acoustic Sensing for Road Traffic Monitoring: Principles, Signal Processing, and Emerging Applications. *Infrastructures* **2025**, *10*, 228. [\[CrossRef\]](#)
11. Nowakowski, M.; Kosiuczenko, K. Evaluation of TAERO UGV Structural Collision Resistance Using FEM Analysis. *Bull. Pol. Acad. Sci. Tech. Sci.* **2025**, *73*, 153431. [\[CrossRef\]](#)
12. Davies, H.W.; Wang, F.; Du, B.B.; Viventi, R.; Johnson, P.W. Exposure to Whole-Body Vibration in Commercial Heavy-Truck Driving in On- and Off-Road Conditions: Effect of Seat Choice. *Ann. Work Expo. Health* **2022**, *66*, 69–78. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Ghaffarpasand, O.; Almojarkesh, A.; Morris, S.; Stephens, E.; Chalabi, A.; Almojarkesh, U.; Almojarkesh, Z.; Pope, F.D. Traffic Noise Assessment Using Intelligent Acoustic Sensors (Traffic Ear) and Vehicle Telematics Data. *Sensors* **2023**, *23*, 6964. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Ma, Q.; Ma, L.; Liu, F.; Sun, D. A Novel Driving Noise Analysis Method for On-Road Traffic Detection. *Sensors* **2022**, *22*, 4230. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Wang, C.; Song, Y.; Liu, H.; Liu, H.; Liu, J.; Li, B.; Yuan, X. Real-Time Vehicle Sound Detection System Based on Depthwise Separable Convolution Neural Network and Spectrogram Augmentation. *Remote Sens.* **2022**, *14*, 4848. [\[CrossRef\]](#)
16. Marciniuk, K.; Kostek, B. Machine Learning Applied to Acoustic-Based Road Traffic Monitoring. *Procedia Comput. Sci.* **2022**, *207*, 1087–1095. [\[CrossRef\]](#)
17. Sun, Q.; Yin, C.; Wang, B. Experimental Validation of Truck Cab Suspension Model and Ride Comfort Improvement under Various Semi-Active Control Strategies. *Processes* **2024**, *12*, 1880. [\[CrossRef\]](#)
18. Antony, A.; Theimer, W.; Grossetti, G.; Friedrich, C.M. Acoustic Event Detection in Vehicles: A Multi-Label Classification Approach. *Sensors* **2025**, *25*, 2591. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Badawy, M.; Rashed, A.; Bamaqa, A.; Sayed, H.A.; Elagamy, R.; Almaliki, M.; Farrag, T.A.; Elhosseini, M.A. From Sensors to Insights: Interpretable Audio-Based Machine Learning for Real-Time Vehicle Fault and Emergency Sound Classification. *Machines* **2025**, *13*, 888. [\[CrossRef\]](#)
20. Rossi, D. Coupling Different Road Traffic Noise Models with a Multilinear Regressive Model: A Measurements-Independent Technique for Urban Road Traffic Noise Prediction. *Sensors* **2024**, *24*, 2275. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Guarnaccia, C.; Catherin, U.; Mascolo, A.; Rossi, D. Improvement and Validation of a Smart Road Traffic Noise Model Based on Vehicles Tracking Using Image Recognition: EAgLE 3.0. *Sensors* **2025**, *25*, 1750. [\[CrossRef\]](#) [\[PubMed\]](#)

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.