

Vibration Condition Monitoring



Frequency-Domain Carpet Detection and Structural Looseness Diagnosis

Technical Case Report

Github Repository

<https://github.com/marioandrededeus/vibration-condition-monitoring.git>

Interactive Application

<https://vibration-monitoring.streamlit.app/>

Author: Mario André de Deus

Date: 2026

Abstract

This technical case study presents a two-stage vibration analysis framework designed to address broadband carpet detection and structural looseness diagnosis.

Rather than approaching the challenge purely as a classification task, the solution emphasizes:

- Physical interpretability
- Bias investigation prior to modeling
- Robustness under domain shift
- Deterministic deployment design

The final system integrates frequency-domain microstructure analysis, harmonic feature engineering, and physics-informed scoring to produce a production-ready diagnostic pipeline.

1. Introduction – Framing the Engineering Problem

Industrial vibration diagnostics is fundamentally an exercise in physical reasoning.

The challenge consists of two tasks:

1. Detect broadband “carpet” regions in frequency spectra.
2. Diagnose structural looseness using tri-axial vibration signals.

Although both tasks operate in the frequency domain, they reflect distinct physical mechanisms:

- Carpet behavior reflects spectral microstructure irregularity.
- Structural looseness reflects stiffness alteration and harmonic amplification.

The central engineering objective was:

To design metrics that reflect physical phenomena rather than statistical coincidence.

2. Part 1 – Broadband Carpet Detection

2.1 Signal Integrity Validation

Before implementing any spectral metric, signal integrity was verified:

- Uniform time sampling
- Stable acquisition duration
- Absence of corrupted values

This step ensures that spectral irregularities are not acquisition artifacts.

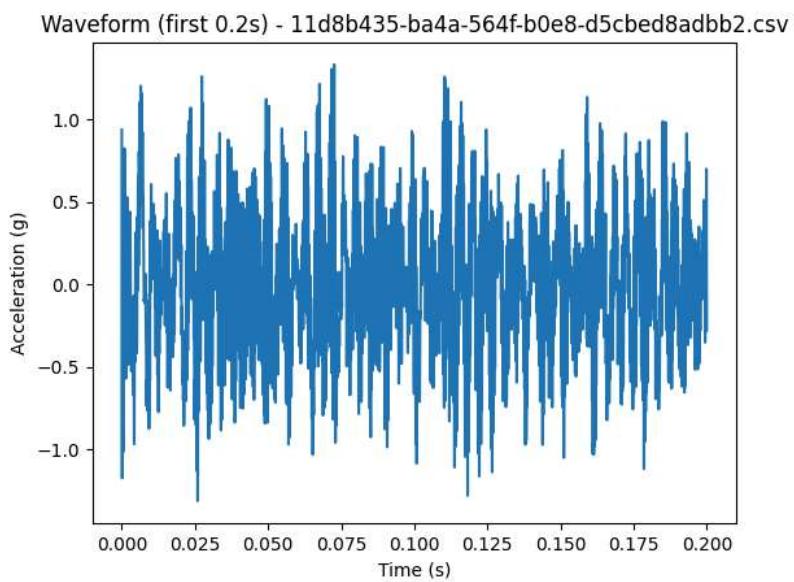


Figure 1 – Representative waveform segment

2.2 Spectral Representation Strategy

Welch PSD was initially used for exploratory diagnostics.

However, PSD smooths peak microstructure, which is central to the carpet definition.

The final detection strategy relies on FFT magnitude, which preserves exact peak positions.



Figure 2 – PSD (≥ 1000 Hz focus)FFT - Spectral representation

2.3 Carpet Definition and Operationalization

Carpet regions consist of spectral peaks randomly close to each other.

In contrast, harmonic systems exhibit regular peak spacing.

Operational constraints:

- Frequency ≥ 1000 Hz
 - Multiple disjoint regions allowed
 - No overlapping intervals
-

2.4 Peak Spacing Irregularity

For consecutive FFT peaks:

$$\Delta f_i = f_{i+1} - f_i$$

Regular harmonic structures produce near-constant spacing.

Carpet structures produce high variability.

Irregularity is quantified via:

$$CV = \frac{\sigma(\Delta f)}{\mu(\Delta f)}$$



Figure 3 – Regular vs Irregular Peak Spacing Comparison

2.5 Dynamic Clustering and Region Extraction

Rather than predefining static frequency bands, a dynamic clustering rule was implemented.

Peaks are grouped when:

$$\Delta f < \text{gap_max_hz}$$

This ensures:

- Disjoint regions
- No overlap
- Physical interpretability

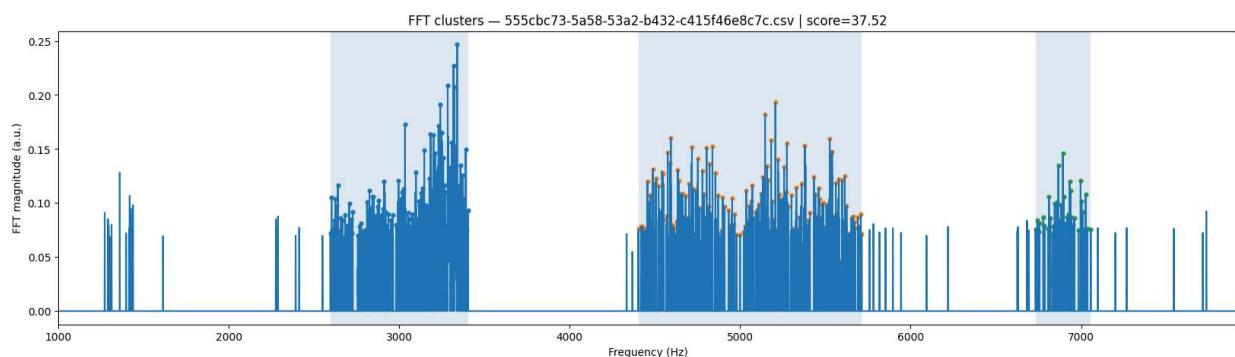


Figure 4 – Clustered Regions Overlaid on FFT

2.6 Carpet Severity Metric

Severity increases with:

- Peak density
- Spacing irregularity

- Absence of dominant single resonance

Per-sample severity is defined as the maximum cluster score.

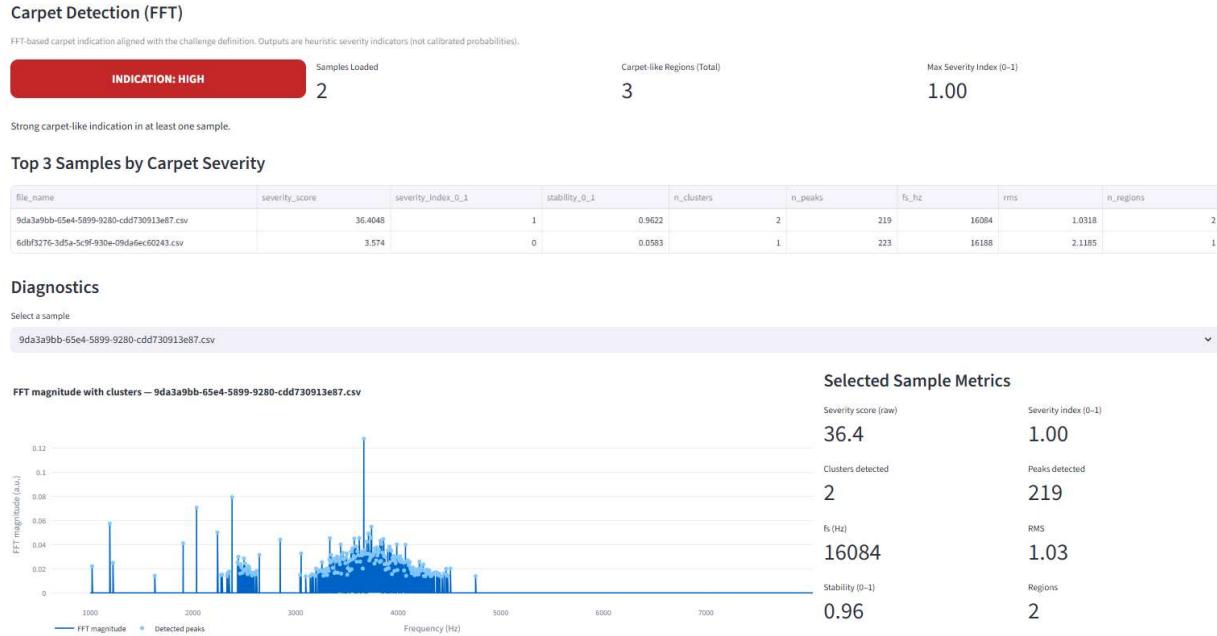


Figure 5 – Carpet Severity Ranking Across Samples

2.7 Robustness Analysis

Clustering tolerance (gap_max_hz) was evaluated at:

- 25 Hz
- 35 Hz
- 50 Hz

The worst-case sample remained invariant across parameter variation.

This indicates ranking stability.

2.8 Final Result

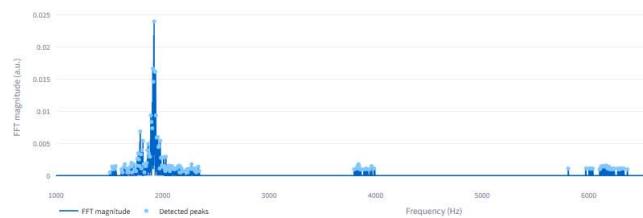
The sample:

Diagnostics

Select a sample

3186c48d-fc24-5300-910a-6d0bafdd87ea.csv

FFT magnitude with clusters — 3186c48d-fc24-5300-910a-6d0bafdd87ea.csv

**Selected Sample Metrics**

Severity score (raw)	50.8	Severity index (0-1)	0.00
Clusters detected	3	Peaks detected	149
fs (Hz)	16188	RMS	0.16
Stability (0-1)	0.95	Regions	3

 Show metric definitions**Regions (by cluster score)**

start_hz	end_hz	n_peaks	density_per_khz	spacing_cv	dominance	score
6102.6112	6361.9709	26	100.2469	0.7169	1.4142	50.8165
3798.5087	3987.7178	16	84.5625	0.5939	1.7023	29.5039
1615.4407	2345.1061	95	130.1967	0.4481	17.3277	3.3667

Figure 6 – Sensitivity Analysis Across Clustering Tolerances

consistently presents the strongest carpet symptom.

The result is stable and consistent with spectral inspection.

3. Part 2 – Structural Looseness Diagnosis

3.1 Structural Looseness – Physical Framing and Modeling Strategy

Structural looseness alters system stiffness.

Expected vibration signatures include:

- Amplified 1× rotational frequency
- Enhanced 2× and 3× harmonics
- Increased crest factor due to nonlinear contact

Unlike carpet behavior, looseness modifies harmonic structure rather than broadband high-frequency energy.

3.2 Sampling-Rate Domain Shift

A structural difference was observed between train and test splits:

Split	Sampling Rate	Duration
Train	~4 kHz	~0.5 s
Test	~8 kHz	~2.0 s

Implications:

- Different spectral resolution
- Different Nyquist frequency
- Different time window length

Mitigation:

- Harmonics anchored to rpm
 - Resolution-aware tolerances
 - Ratio-based features
-

3.3 Feature Engineering Anchored in Physics

Extracted features:

- 1×, 2×, 3× harmonic amplitudes
- 2×/1× and 3×/1× ratios
- Crest factor
- Low-frequency band metrics

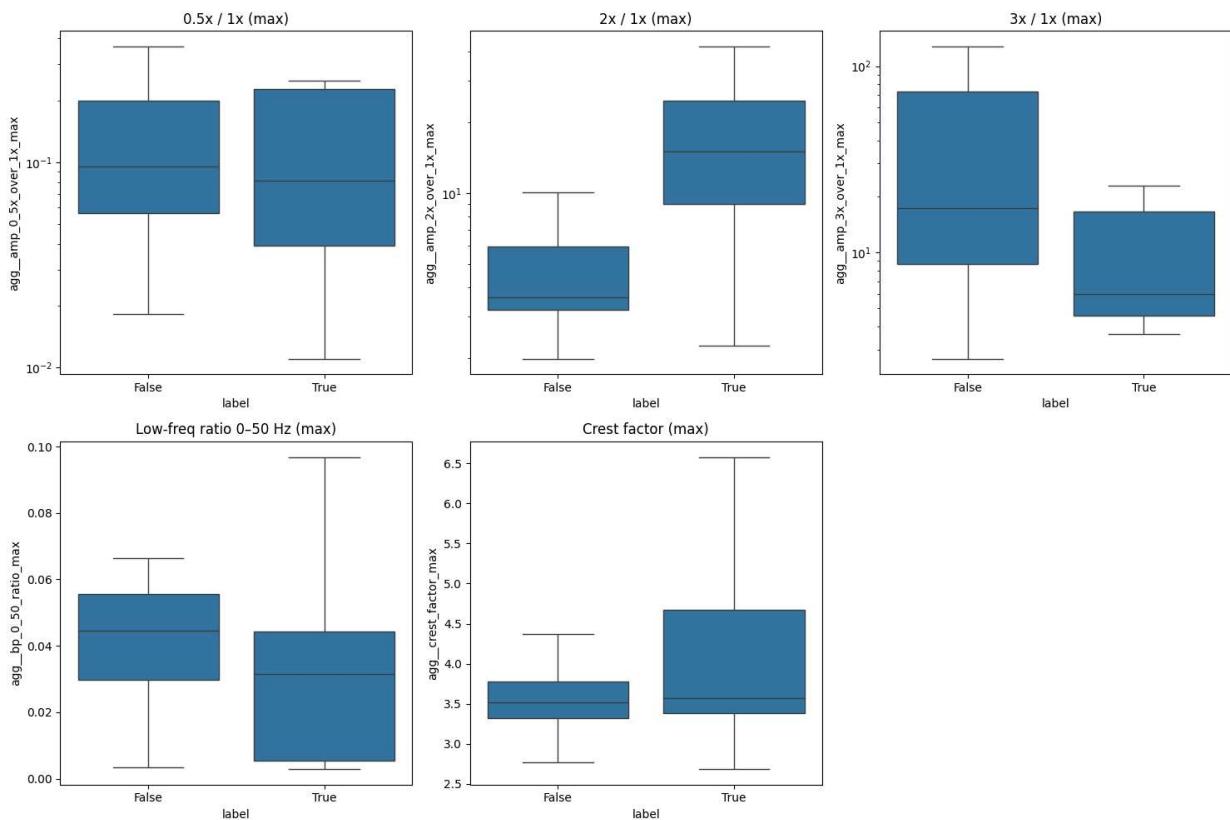


Figure 7 – Harmonic amplification

3.4 Structural Bias – Sensor and RPM

Sensor distribution was balanced.

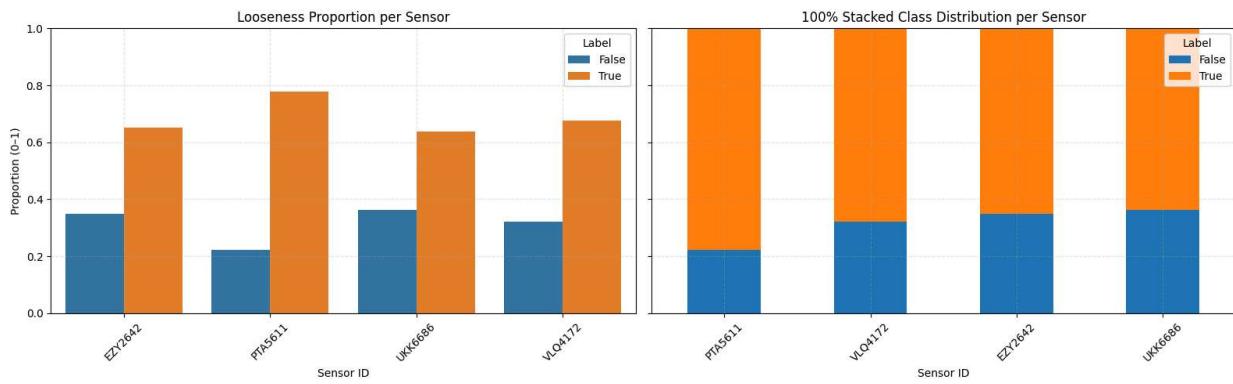


Figure 8 – Sensor bias

However, rpm revealed structural correlation:

- All 1595 rpm samples labeled looseness.
- 1598 rpm contained both classes.

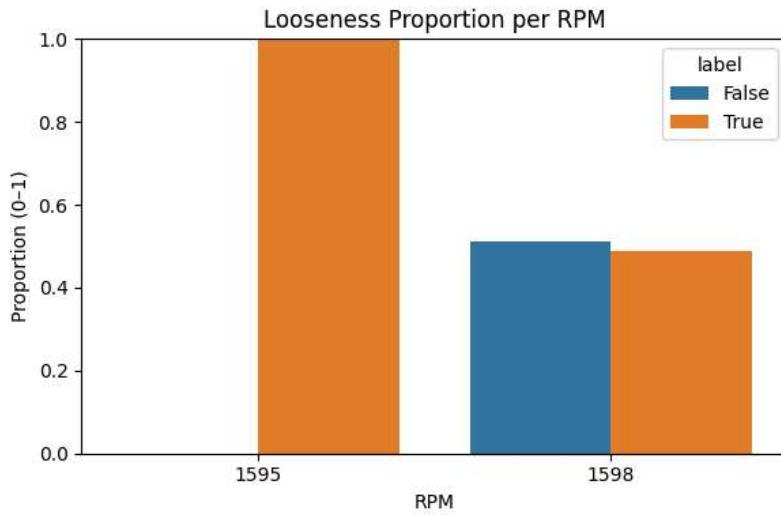


Figure 9 – RPM bias

To prevent regime-based shortcut learning, evaluation was performed in two scenarios:

- Full dataset
 - rpm = 1598 only
-

3.5 Baseline – Single Harmonic Threshold

A simple 2x/1x threshold achieved strong recall.

This confirms harmonic amplification as the dominant looseness signature.

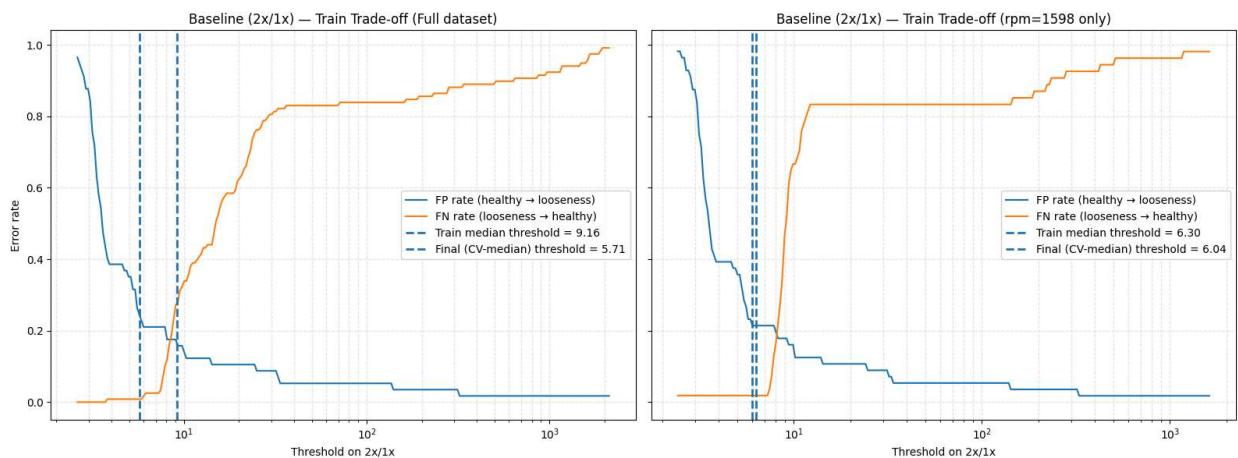


Figure 10 – Baseline treshold

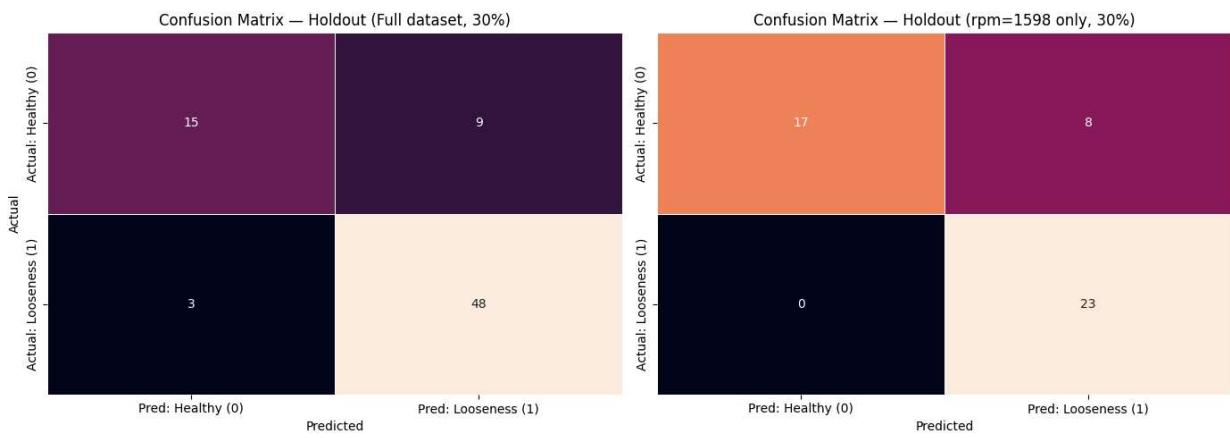


Figure 11 – Baseline confusion matrices

3.6 Logistic Regression and Random Forest Benchmarking

Supervised models were evaluated using:

- Stratified splits
- Cross-validation
- Holdout validation

Performance gains over baseline were moderate.

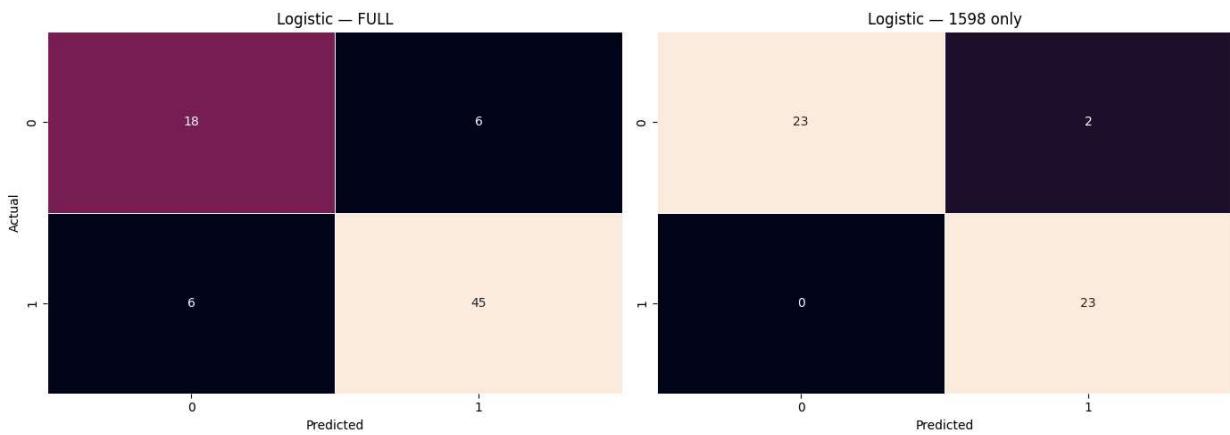


Figure 12 – Logistic regression confusion matrices

3.7 Physics-Informed Heuristic v2

A soft-scoring model combining:

- 2×/1× harmonic amplification
- 3×/1× nonlinear amplification
- Crest factor

was implemented.

Severity score normalized to [0,1].

This approach achieved performance comparable to Random Forest while offering:

- Interpretability
- Lower complexity
- Deterministic inference

	Scenario	Model	Precision (Looseness)	Recall (Looseness)	F1	ROC-AUC	PR-AUC	Threshold
0	Full dataset (1595+1598)	Physics Baseline (2x/1x, tuned threshold)	0.842105	0.941176	0.888889	0.759804	0.805764	5.707472
1	Restricted (rpm=1598 only)	Physics Baseline (2x/1x, tuned threshold)	0.741935	1.000000	0.851852	0.780870	0.738692	6.039541
2	Full dataset (1595+1598)	Logistic Regression (tuned)	0.882353	0.882353	0.882353	0.843137	0.884883	0.658027
3	Restricted (rpm=1598 only)	Logistic Regression (tuned)	0.920000	1.000000	0.958333	0.998261	0.998188	0.453344
4	Full dataset (1595+1598)	Random Forest (RandomizedSearchCV)	0.916667	0.862745	0.888889	0.933007	0.971626	NaN
5	Restricted (rpm=1598 only)	Random Forest (RandomizedSearchCV)	0.920000	1.000000	0.958333	0.996522	0.996522	NaN
6	Full dataset (1595+1598)	Physics Heuristic v2 (sigmoid: 2x+3x+crest)	0.859649	0.960784	0.907407	0.758987	0.820254	0.400000
7	Restricted (rpm=1598 only)	Physics Heuristic v2 (sigmoid: 2x+3x+crest)	0.766667	1.000000	0.867925	0.841739	0.809928	0.520000

Figure 13 – Models comparison

3.8 Final Inference Strategy

The deployed heuristic was tuned using the rpm=1598 subset to reduce regime artifact risk.

This ensures:

- Physical feature reliance
- Reduced shortcut learning
- Better operational robustness

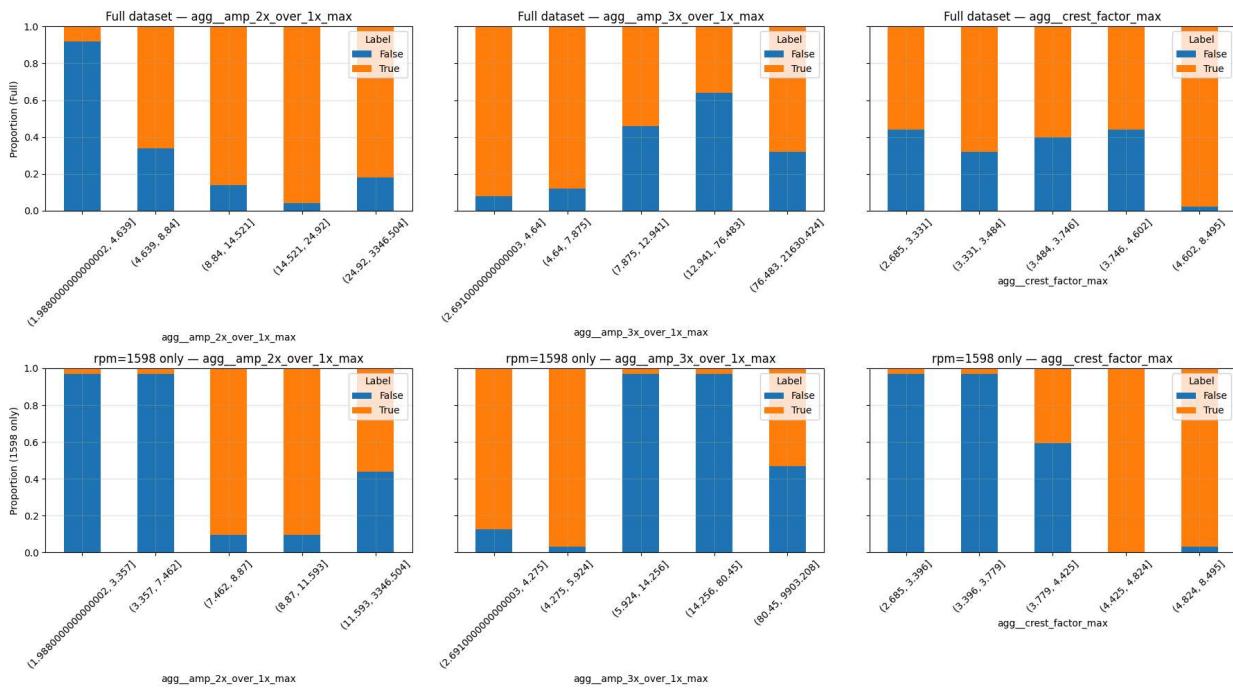


Figure 14 – Physical feature reliance

4. Engineering Insights

Key insights from the investigation:

1. Spectral structure is more informative than broadband energy for carpet detection.
2. Dataset bias must be investigated before model selection.
3. Harmonic amplification is the dominant looseness indicator.
4. Marginal ROC gains do not justify black-box deployment.
5. Physics-informed heuristics can match ML performance with greater transparency.

	sample_id	rpm	prediction_looseness	score
0	e057600e-3b4e-58ba-b8b8-357169ae6bf6	1800.0	True	0.742132
1	9f3b933a-1bc3-5093-9dee-800cc03c6b1d	1590.0	True	0.654978
2	33542920-30ea-5844-861d-2c82d79087b8	1170.0	True	0.601234
3	1dab1534-b8a8-5962-b01c-bff0782d54a9	3545.0	False	0.144208
4	2211750b-6672-5a94-bd40-cda811f69d01	2025.0	False	0.074849
5	680bbcbf-b1c8-544d-8f80-bf763cdcd128	3573.0	False	0.066461
6	01e98ad9-23c9-5986-ace0-4519bad71198	1785.0	False	0.016721

Figure 15 – Test data prediction - Score

5. Conclusion

The final solution integrates:

- FFT-based microstructure analysis
- Dynamic spectral clustering
- Harmonic feature engineering
- Bias-aware modeling
- Physics-informed scoring

Part 1 delivers a deterministic and robust carpet detection metric.

Part 2 delivers an interpretable looseness classifier grounded in physical vibration behavior.

The outcome is not merely predictive — it is structurally diagnostic and deployment-ready.

References

Xu, X., Liao, X., Zhou, T., He, Z., & Hu, H. (2024).

Vibration-based identification of lubrication starved bearing conditions. *Measurement*.

Randall, R. B. (2011).

Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications. Wiley.