Influence of Concurrent Boiling Events on Acoustic Nucleation Rhythms

Andrew Martinez[1] and Michael Khasin[2]

[1]Department of Statistics, California Polytechnic State University, San Luis Obispo, CA
[2]NASA Ames Research Center, Moffett Field, CA

January 15, 2026

Abstract

We study whether the presence of multiple, simultaneous nucleation sources is *associated* with changes in the temporal regularity of acoustic nucleation rhythms. Using hand-labeled accelerometer data from bench-top experiments designed to inform cryogenic propellant management for space applications, we compare the dispersion of inter-nucleation intervals across single-source and concurrent-source conditions. Nonparametric comparisons (Mann–Whitney U) indicate larger dispersion under concurrency for several pre-specified contrasts at the $\alpha = 0.05$ level. We frame these results as *associations* that provide statistical support for a thermal-recovery mechanism, in which neighboring detachment events intermittently cool the wall and broaden timing variability. The analysis is observational and pre-specified; exact statistics are reported without post-hoc reprocessing, and limitations for inference and generalization are detailed.

Keywords: event-time analysis, nonparametric inference, dispersion of inter-event intervals, effect size estimation, bootstrap confidence intervals, thermal cross-talk

1 Introduction

Effective Cryogenic Fuel Management (CFM) is a critical enabler for long-duration space exploration, where the storage and transfer of volatile propellants in low-gravity environments present significant engineering challenges [1, 2]. A key operational risk arises from localized heat leaks into propellant tanks, which can induce unwanted incipient boiling. Such phenomena can lead to inefficient propellant utilization, unpredictable tank pressure fluctuations, and potentially hazardous thermal instabilities. Consequently, developing passive, non-intrusive diagnostic techniques to detect and characterize the earliest stages of boiling is essential for ensuring the safety and reliability of future space systems.

Acoustic sensing, using externally mounted accelerometers, offers a promising modality for this purpose. Foundational work by Khasin et al. [1] established an experimental framework to link acoustic signatures to underlying boiling physics. Using superheated water in a smooth-walled vessel to emulate cryogenic conditions, they identified two distinct boiling regimes: a stochastic "random boiling" pattern, attributed to the activation of residual gas nuclei, and a highly periodic "rhythmic boiling" pattern, which emerged under conditions of high superheat. They proposed a comprehensive physical model for rhythmic boiling, positing that it originates from heterogeneous nucleation on microscopic hydrophobic contamination spots on the vessel wall.

The periodicity of this rhythmic boiling is explained by a thermal-recovery limit cycle. In this model, (i) a vapor bubble nucleates and grows at a specific hot site; (ii) upon detachment, cooler bulk liquid rushes in, locally quenching the site's temperature; and (iii) a "waiting time" governed by thermal diffusion ensues as the site reheats to the nucleation threshold, at which point the cycle repeats. The duration of this thermal recovery, estimated to be on the order of seconds, sets the fundamental period of the acoustic rhythm [3].

This study builds directly upon that physical model by investigating a key, testable consequence: if a site's rhythm is governed by a sensitive local thermal process, it should be susceptible to thermal perturbations from its surroundings. The presence of a second, concurrently active nucleation site provides a natural experiment for testing this. A detachment event at a neighboring site can be viewed as an external thermal disturbance, which we hypothesize will interfere with the primary site's regular recovery cycle - a thermal "cross-talk" hypothesis. We therefore ask the specific question: *is the presence of an additional active boiling source associated with greater temporal irregularity in a given site's nucleation rhythm?* Guided by the thermal-recovery model, we had formulated the directional hypothesis that concurrency increases the dispersion of inter-event intervals. Preliminary analyses of the dataset from [1], reported in [4], provided an initial evidence supporting the hypothesis of thermal cross-talk between nucleation sites. The present study subjects that pre-analysis hypothesis to a formal statistical evaluation on the same dataset. By comparing the dispersion of inter-event timings between single-source and concurrent-source conditions, this work provides a targeted statistical test of the thermal crosstalk predicted by the recovery-cycle model.

2 Dataset

2.1 Experimental Setup

We analyze 441 short-duration experiments (1 s to 30 s) from the testbed developed by Khasin et al. [1]. Each run consists of a heated spot on a vessel containing degassed water, with two wall-mounted accelerometers sampling at 10 kHz (Fig. 1). A controlled heat input elicits nucleation activity, which was correlated with high-speed video in the original study to validate the acoustic signatures. One accelerometer data channel was used to maintain continuity with the pre-specified analysis pipeline.

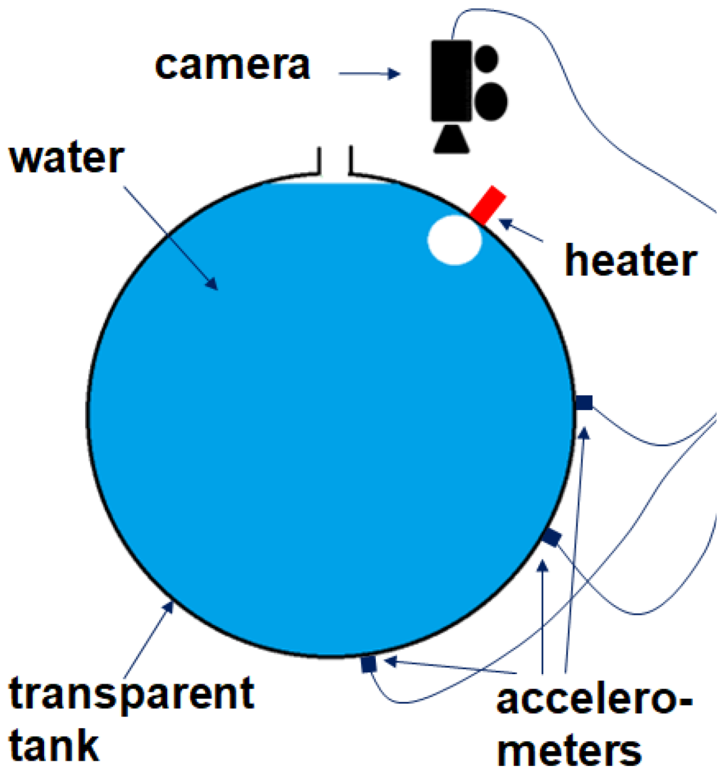

Figure 1: Schematic of the experimental arrangement with external accelerometers, adapted from [1].

2.2 Signal and Labels

Raw acceleration traces show sparse, high-amplitude transients corresponding to nucleation/collapse events (Fig. 2). Events were *hand-labeled* using a Python Dash interface to ensure temporal precision.

2.3 Within-Run Site Identification by Amplitude

In dual-source runs, we distinguish two rhythmic sources via *within-run* amplitude ranking: the *Dominant* rhythm is the site producing consistently larger peaks; the *Lesser* rhythm produces smaller peaks. We use amplitude *as a proxy for site identity only within a run*, and do not compare absolute amplitudes across runs due to geometry-dependent propagation.

3 Methods

3.1 Estimand and Groups

Our primary estimand is the difference in *dispersion* of inter-event intervals between conditions. We pre-specify contrasts of interest: Single vs Dominant, Single vs Lesser, and Single

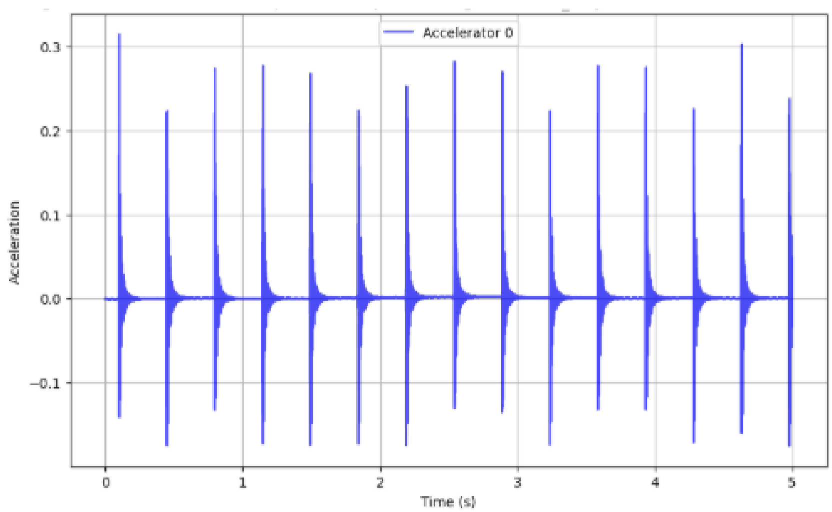

Figure 2: Representative time-domain signal from a single run. Peaks correspond to nucleation events.

vs Random. Group definitions follow the original pipeline and are retained to respect the pre-specified analysis.

Furthermore, the Mann–Whitney U tests are performed on the *per-run standard deviations (SDs) of inter-event intervals*, comparing the distribution of these SDs across groups rather than testing variances directly.

3.2 Distributional Diagnostics and Tests

We assessed distributional shape via the Shapiro–Wilk test [6] and heterogeneity of variance via the Brown–Forsythe test (median-centered Levene) [7]. Because several groups deviated from normality and variances were heterogeneous, we abandoned Welch's t-test in favor of the nonparametric Mann–Whitney U test for location comparisons.

All Mann–Whitney U tests [5] were computed using the `asymptotic` method, which compares the standardized test statistic against the normal distribution with tie correction. Additionally, one-sided (upper-tailed) alternatives were specified under the directional hypothesis that dispersion is greater under concurrent-source conditions.

Cliff's δ [8] was used to estimate the magnitude and direction of pairwise differences in dispersion. To quantify uncertainty 95% confidence intervals were obtained via nonparametric bootstrapping [9]. For each comparison, Cliff's δ was recalculated across 1,000 bootstrap iterations, and the empirical 2.5th and 97.5th percentiles of the resulting distribution were taken as the confidence bounds.

We report exact statistics and p-values as computed under the pre-specified analysis along with Cliff's δ. Multiple-comparison adjustments were not applied; we therefore designate Single vs (Dominant, Lesser, Random) as *primary* and all other pairwise contrasts as *exploratory*.

3.3 Run Inclusion and Stationarity Considerations

Intervals were computed over entire runs (1 s to 30 s). Short windows can under-sample slow rhythms; we therefore interpret dispersion estimates conservatively for fast or dominant rhythms. Because each analyzed segment is an arbitrary excerpt of a longer experiment (the segment boundaries have no physical meaning), start/end "warm-up/cool-down" detrending is not applicable. We therefore applied no boundary-based detrending; any within-segment nonstationarity (slow drift in level or variance) may inflate dispersion and is noted as a limitation.

All rhythms within an experimental run were required to exhibit at least 5 peaks (yielding 4 inter-event intervals) to ensure stable dispersion estimates. This criterion was satisfied by every analyzed run, so no subsetting of the original corpus was necessary. Final group sample sizes (n) are shown below and are consistent across all figures and statistical tables.

Table 1: Sample sizes (n) contributing to each group after applying inclusion criteria.

Group	**n (runs)**
Single	99
Dominant	42
Lesser	42
Random	29

These sample sizes reconcile with the 441-run experimental corpus reported in Section 2 and described in [4]. The original experiment corpus includes six classifications of runs. For this study we analyzed three classifications, including Single Rhythmic (Single), Double Rhythmic (Dominant, Lesser), and lastly Rhythmic and Random (Random) runs. These three classes comprise of 212 of the 441 total runs. The remaining 229 experimental runs and three categories, which were excluded from analysis, consist of Purely Random, Noise, and Rhythmic Climax runs. Purely Random runs were excluded because they do not contain a dominant rhythmic nucleation signal to analyze. Similarly, noise runs were excluded because they lack any identifiable nucleation signal. Lastly, Rhythmic Climax runs were excluded because only three examples exist within the corpus, which is insufficient for any meaningful statistical comparison.

3.4 Blinding and Labeling Protocol

Events were hand-labeled using a custom dashboard under a written guideline specifying peak identification and merge/split decisions. During labeling, group assignment (e.g., Single, Dominant, Lesser, Random) and any downstream statistics were not displayed to the annotator. All labels were produced by a single trained domain expert blinded to group assignment; no inter-rater reliability assessment was performed, which we acknowledge as a limitation. To reduce subjectivity, the same decision rules were applied uniformly across runs, and ambiguous cases were resolved by rule-based criteria rather than outcome appearance.

4 Results

4.1 Visual Summaries

Kernel density and violin plots of inter-event interval dispersion show distinct distributional shapes across groups (Figs. 3 and 4). Qualitatively, Single exhibits a narrower distribution than concurrent conditions.

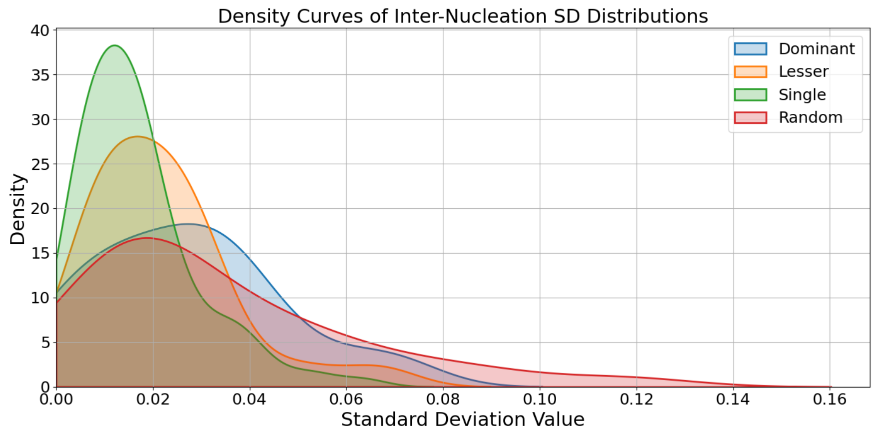

Figure 3: Kernel density of standard deviations (SD) of inter-event intervals across groups.

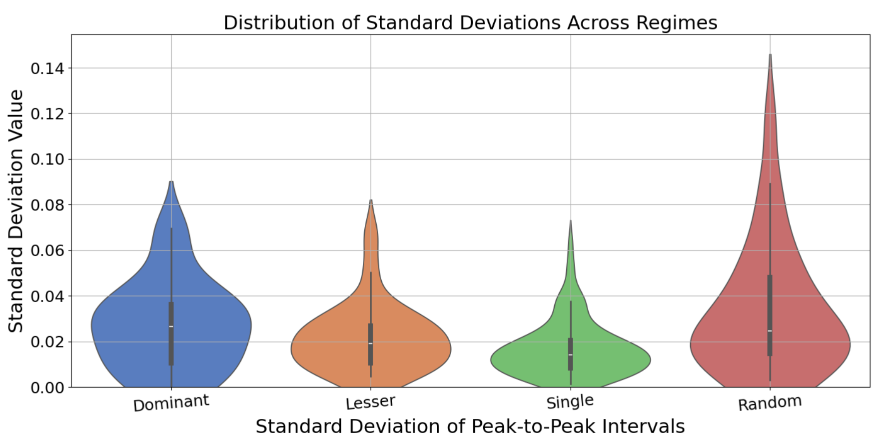

Figure 4: Violin plots of SD distributions across groups.

4.2 Normality and Variance Checks

The Shapiro–Wilk test rejected normality for most groups except Dominant (Table 2). Levene's test indicated unequal variances for some contrasts, notably Dominant vs Single (Table 3).

Table 2: Shapiro–Wilk Normality Tests.

Group	n	Statistic	p-value
Dominant	42	0.945	0.215
Lesser	42	0.871	0.005*
Random	29	0.851	<0.001*
Single	99	0.887	<0.001*

Table 3: Levene's Tests for Equality of Variances.

Group Comparison	n (G1, G2)	Statistic	p-value
Single vs *Lesser*	99, 42	0.713	0.399
Dominant vs *Random*	42, 29	0.870	0.357
Dominant vs *Lesser*	42, 42	2.490	0.122
Dominant vs *Single*	42, 99	9.840	0.002*

4.3 Primary Nonparametric Comparisons

Directional Mann–Whitney U tests (upper-tail) support greater dispersion under concurrency for Single vs Dominant and Single vs Random; Single vs Lesser is also statistically significant (Table 4). Moreover, Cliff's δ effect sizes and their respective CI's support greater dispersion under concurrency for Single vs Dominant and Single vs Random (Table 5).

Table 4: Upper-tailed: (Concurrency > Single) Mann–Whitney U Tests.

Group Comparison	n (G1, G2)	U	p-value
Dominant vs *Lesser*	42, 42	234	0.135
Lesser vs *Random*	42, 29	263	0.092
Single vs *Lesser*	99, 42	918	0.043*
Single vs *Dominant*	99, 42	809	0.008*
Single vs *Random*	99, 29	835	0.001*

Table 5: Cliff's δ Effect Sizes for Group Comparisons of Dispersion.

Group Comparison	n (G1, G2)	Cliff's δ	95% CI
Dominant vs *Lesser*	42, 42	-0.186	[-0.524, 0.128]
Lesser vs *Random*	42, 29	-0.221	[-0.521, 0.083]
Single vs *Lesser*	99, 42	-0.231	[-0.473, 0.014]
Single vs *Dominant*	99, 42	-0.324	[-0.596, -0.032]
Single vs *Random*	99, 29	-0.401	[-0.631, -0.170]

5 Mechanism Background (Qualitative)

The physical mechanism for rhythmic boiling proposed by Khasin et al. [1] provides the theoretical basis for our hypothesis. Their model describes a thermal-recovery limit cycle: (i) a bubble nucleates and detaches from a hot spot on the wall; (ii) this detachment allows cooler bulk liquid to contact the site, locally quenching the temperature; (iii) a "waiting time" ensues, governed by thermal diffusion, as the site reheats to the nucleation threshold; (iv) a new bubble forms, repeating the cycle. The duration of this reheating phase, τ, sets the fundamental period of the rhythm.

Our study tests a direct consequence of this model. If the rhythm is governed by a sensitive local thermal process, it should be susceptible to external thermal perturbations. We treat a neighboring detachment event as such a perturbation. The influx of cool liquid from a nearby event could transiently cool the primary site via diffusion or short-lived convection, thereby lengthening its recovery time and broadening the overall distribution of inter-event intervals.

6 Discussion

6.1 Interpretation

We find that concurrent-source conditions are *associated* with statistically significant increases in the dispersion of inter-event intervals compared to a single rhythmic source. This result provides strong statistical support for the thermal-recovery limit cycle model proposed by Khasin et al. [1]. That model predicts the rhythm's period is set by a sensitive thermal recovery time. Our findings confirm a key consequence of this model: an external thermal disturbance—in this case, a neighboring boiling event—perturbs the cycle and measurably increases its timing irregularity.

With a pre-specified, single-channel analysis, we limit inference to associations rather than causation. However, the observed distributional differences serve as compelling observational evidence consistent with the hypothesis that intermittent cooling from neighboring events acts as phase "kicks" to a primary rhythm.

6.2 Limitations

(i) Single-channel analysis with amplitude-based within-run identity may misassign events in rare configurations; we mitigated this by requiring stability of amplitude rank (see Methods). (ii) Lack of multiplicity correction inflates Type I error across exploratory contrasts; primary contrasts were pre-specified. (iii) Stationarity was not enforced; warm-up/cool-down may inflate dispersion.

6.3 Implications for Space Applications

For propellant tank health monitoring, the dispersion of acoustic inter-event intervals may serve as a conservative indicator of concurrency and evolving thermal instability. Extrapolation to microgravity should account for altered buoyancy and thermal transport [1]; the present results are best viewed as context-specific associations that inform sensor fusion strategies within CFM frameworks.

7 Conclusion

In a hand-labeled acoustic corpus generated by the experiments of Khasin et al. [1], we find that concurrent boiling sources are *associated* with greater dispersion of inter-event intervals relative to a single rhythmic source. These distributional differences provide statistical support for the underlying thermal-recovery mechanism and suggest a potential operational diagnostic for monitoring complex boiling states. Future work should report nonparametric effect sizes, adjust for multiplicity, and, where possible, separate sources across channels to strengthen inference.

8 Data availability

The acoustic, visual, and metadata files used in this study are publicly available through NASA's Intelligent Systems Division datasets page under the entry "Acoustic and Visual Data for Incipient Boiling at Local Heat Leaks in a Water Tank" [10].

References

[1] M. Khasin, J. Rogers, and V. Osipov, "Acoustic data-based characterization of incipient boiling for space applications," NASA/TM–20250009668, NASA Ames Research Center, Moffett Field, CA, Sep. 2025.

[2] National Aeronautics and Space Administration, "NASA 2024 Technology Taxonomy," Report, NASA, July 2024.

[3] J. H. Lienhard, IV and J. H. Lienhard, V, *A Heat Transfer Textbook*, 5th ed. Mineola, NY: Dover Publications, 2019.

[4] K. Gupta, J. Lamkin, A. Martinez, Z. Weinfeld, K. Bodwin,A. Dekhtyar, M. Khasin, "Interpretable Machine Learning for Acoustic Classification of Incipient Boiling Regimes," NASA/TM-20250011359, NASA Ames Research Center, Moffett Field, CA, Dec. 2025.

[5] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, vol. 18, 1947.

[6] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, 1965.

[7] M. B. Brown and A. B. Forsythe, "Robust tests for the equality of variances," *Journal of the American Statistical Association*, vol. 69, 1974.

[8] N. Cliff, "Dominance statistics: Ordinal analyses to answer ordinal questions," *Psychological Bulletin*, vol. 114, 1993.

[9] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall/CRC, 1994.

[10] M. Khasin, "Acoustic and Visual Data for Incipient Boiling at Local Heat Leaks in a Water Tank (v1.0)," NASA Intelligent Systems Division Datasets, NASA Ames Research Center, last updated Nov. 25, 2025. Available: https://www.nasa.gov/intelligent-systems-division/datasets/. Accessed: January 15, 2026.

A Applet Screens

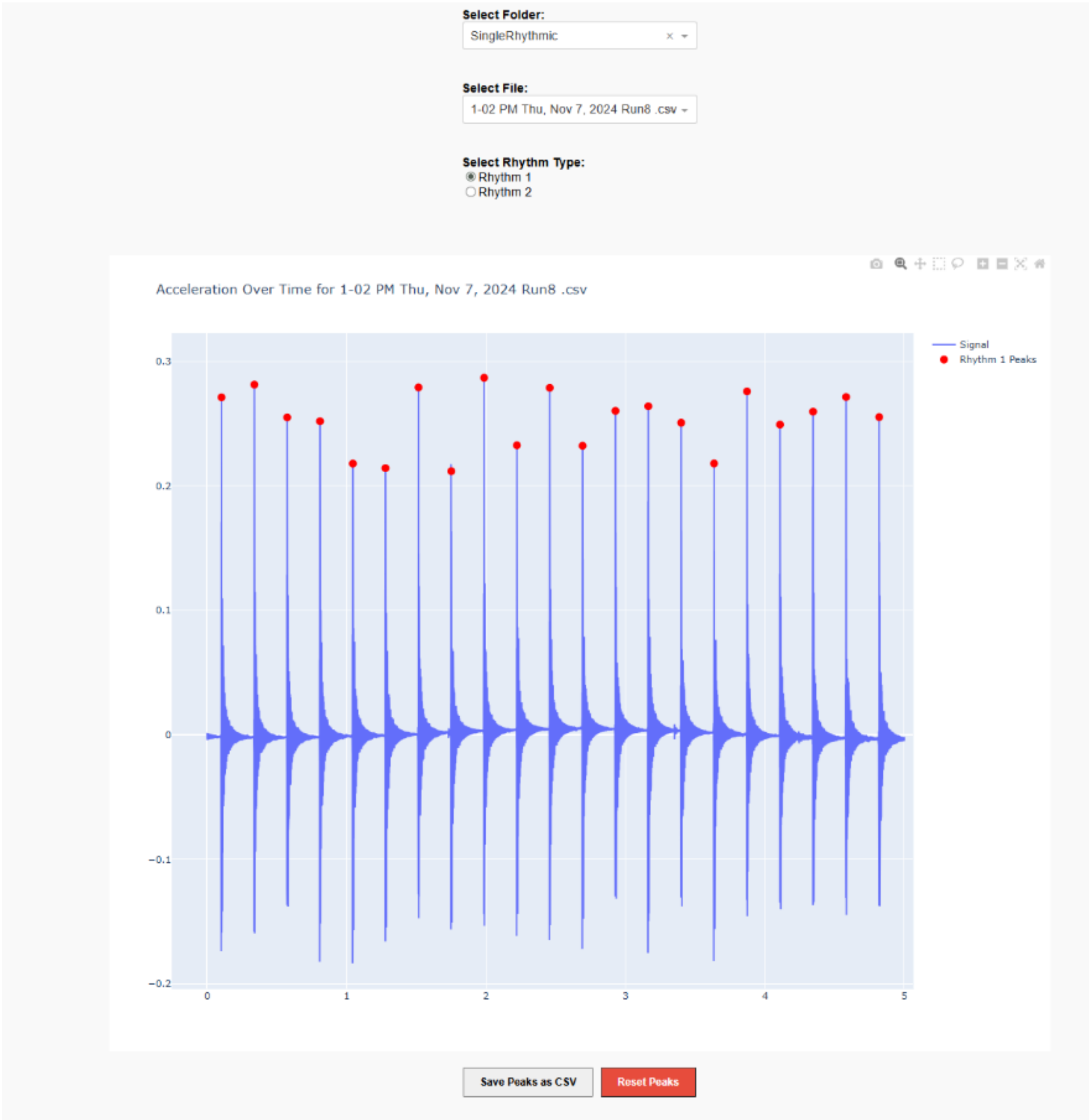

Figure 5: Full landing page of the Python Dash app used for manual labeling.

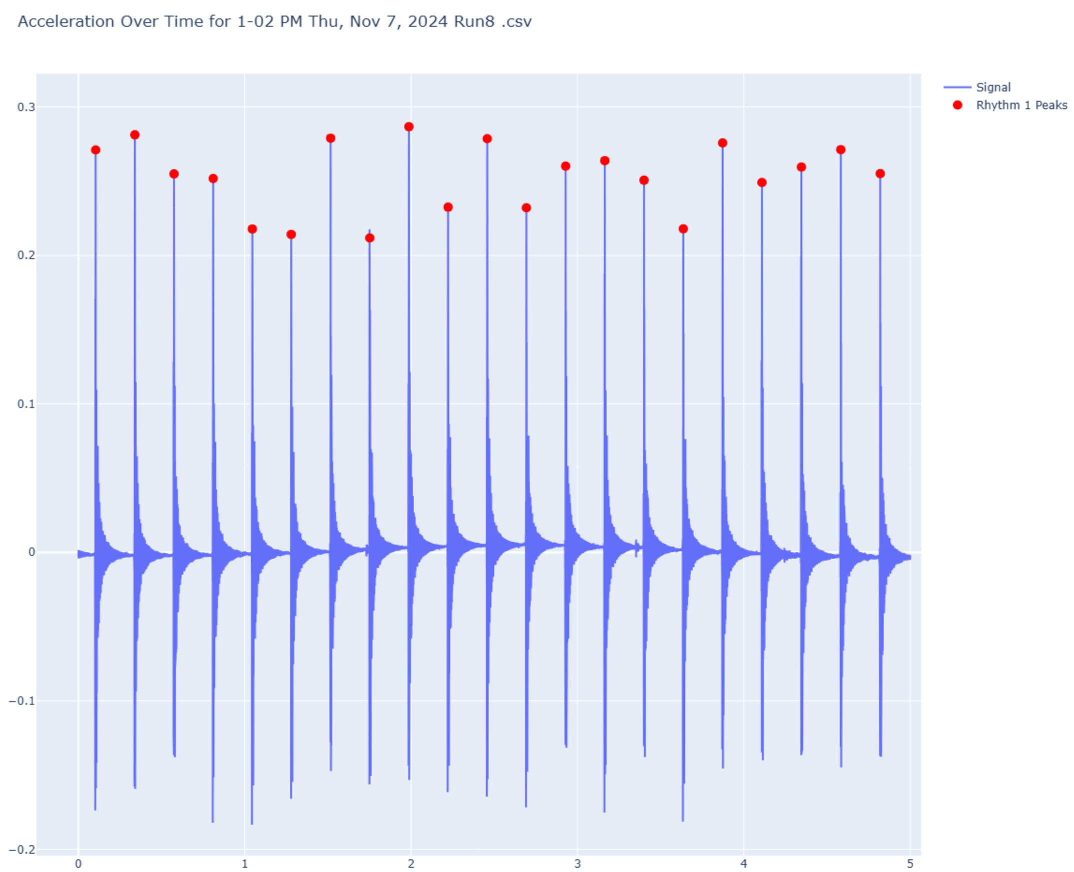

Figure 6: Example of a manually labeled acoustic run. Peaks are identified and validated by visual inspection.