Keeping the Spikes at the Edge: Wavelet Approximation of Machine Waveforms at a Tenth of Lossless Size

Version 0.3 — method, frozen study and confirmation

Zin Nam

ABSTRACT

A sensor at a machine produces far more waveform than an edge device can send or keep. Lossless storage needs about 1.5 bytes per sample; a cellular or low-power uplink affords a tenth of that or less. We ask which approximation keeps the most of what diagnosis needs at 0.05–0.3 bytes per sample, and whether a small device can produce it in real time. Our candidate is a plain transform coder: an orthogonal wavelet transform, one absolute threshold for all levels, uniform quantisation, varint and deflate coding, rate control by bisection on the threshold, and per-level noise levels that let the receiver decode either plainly or with noise filling. In a study frozen before its confirmation data were used, on five held-out datasets (two bearing test rigs, screwdriving, seismic stations and CNC milling), the coder had lower error than SZ3 and ZFP at every budget on four or five of the five datasets, detected run-to-failure bearing faults no later than on the original data (SZ3 was late on five of seven bearings), and encoded a 64 kHz channel at three times real time on one core of a Raspberry Pi 3B+. A heavy-tailed Bayesian prior for selecting coefficients did not help. Diagnosis did not survive everywhere: with noise filling, seismic STA/LTA triggered on 94.8% of the arrivals the original found (the frozen bar was 95%), and on short screwdriving cycles a weak shape statistic weakened further. The approximation is a strong code for vibration at the edge, not a general substitute for exact data.

1 INTRODUCTION

Condition monitoring at the edge faces a budget at every step. A 64 kHz vibration channel produces 5.5 billion samples a day; lossless storage in MorletDB needs 1.5–1.7 bytes per sample [35], about 9 GB a day, and a cellular or LoRa uplink carries a small fraction of that. What diagnosis needs from the signal is often short: an impact train in the envelope spectrum [4, 37], the onset of a seismic arrival [3, 47], the shape of a tightening curve. Downsampling removes the high-frequency content that carries impacts; statistics remove the shape. Error-bounded lossy compressors [16, 27, 28] bound the error per sample, which for broadband vibration spends the budget on noise.

Wavelets concentrate transients in few large coefficients and spread stationary background over many small ones [15, 29]; keeping the largest coefficients is the best approximation for their number in an orthogonal basis, and wavelet compression of vibration data has been used offline [41] and on sensor nodes [48]. What is missing is a measurement of what such an approximation preserves for diagnosis, against today's error-bounded compressors, at the budgets an edge uplink affords, and on a device of the size that would run it.

Contributions. (1) A small transform coder with per-window rate control that an edge device can run. (2) Four falsifiable hypotheses with refutation criteria, fixed before confirmation data are used. (3) Development findings on bearing, seismic and screwdriving data, including the mistakes that changed the method. (4) A place in a waveform database: the approximation as a level between exact samples and summaries in a graded-forgetting ladder [34].

2 METHOD

A window x (one second of vibration, one screwdriving cycle, 2^{15} seismic samples) is decomposed with an orthogonal discrete wavelet transform (sym8; 6, 4 and 8 levels). Every detail coefficient with $|d| > T$, $T = c \cdot \mathrm{std}(x)$, is kept; kept coefficients and the coarsest band are quantised with step qT. The positions of kept coefficients (as gaps) and their values (zigzag) are written as LEB128 varints, preceded by the length, the number of levels and the step, and the stream is compressed with deflate. Decoding inverts each step. *Rate control*: for a budget B in bytes per sample, the device bisects on c (eight tries, $q = 0.5$, the transform computed once) and sends the smallest threshold whose code fits B. The development harness used a fixed grid of (c, q) instead, choosing the smallest error within B; on the device, bisection was both faster and slightly better (Section 4).

Comparisons, each with the same rate control: SZ3 [27] and ZFP [28] at absolute error bounds; decimation (FIR, zero phase) with quantised, delta-coded values and the same coder; and the same coder with coefficients selected by the posterior of an empirical-Bayes spike-and-slab model with a heavy-tailed slab [1, 22, 38].

3 HYPOTHESES

H1, error per byte. At 0.05, 0.1, 0.2 and 0.3 bytes per sample the wavelet coder has lower median relative error than SZ3, ZFP and decimation on every dataset. *Refuted* at a budget if SZ3 or ZFP is as good or better on two of three datasets.

H2, diagnosis survives. At 0.1 bytes per sample, analysis on the approximation finds what it finds on the original: a run-to-failure bearing fault is detected in the envelope spectrum no more than 1% of the run's files later; STA/LTA triggers on at least 95% of the catalogued arrivals the original triggers on, with at most 10% more triggers elsewhere; kurtosis separates failed screwdriving cycles, and a label-free score of RMS and kurtosis separates bad CNC milling runs, with an AUROC within 0.02 of the original. *Refuted* if any fails.

H3, real time on a small device. On a Raspberry Pi 3B+ the encoder keeps up with a 64 kHz channel on one core with less than 64 MB of memory. *Refuted* if not.

H4, no Bayesian machinery needed. Posterior selection does not lower the error by more than 5% against the plain threshold at any budget. *Refuted*, and the Bayesian arm recommended, if it does on two of three datasets.

Table 1: Development, round 3: median relative error at 0.1 bytes per sample, and diagnosis on the approximation at 0.1 (original in brackets).

	Wavelet	SZ3	Decimation
Error, CWRU	**0.56**	0.93	0.86
Error, IMS	**0.70**	0.95	0.74
Error, seismic	**0.038**	0.063	0.079
Error, PyScrew	0.090	–	**0.086**
Error, CNC milling	**0.42**	0.79	0.47
IMS fault detected at file (532)	**532**	549	539
Seismic arrivals triggered (28 of 30)	29	26	26
Seismic triggers elsewhere (400)	955	**393**	556
PyScrew kurtosis AUROC (0.961)	0.957	–	0.954
CNC bad-run AUROC (0.765)	0.745	0.717	0.737
CNC RMS deviation	**4.2%**	5.4%	9.3%

4 DEVELOPMENT

Development used CWRU bearings, IMS set 2 (a bearing run to an outer-race failure), PyScrew scenario s01, two weeks of four seismic stations with a public catalogue and one of three machines of the Bosch CNC milling data [42], all separate from the confirmation data [32]. The first version of the study also proposed a label-free relevance score from the same posterior; it was not better than kurtosis on development data and was dropped before any freeze.

What changed the method. A threshold set from a noise estimate per level, the usual choice for denoising, treated the seismic microseism as noise and removed most of the signal; a coder needs one absolute threshold for all levels. STA/LTA on the raw trace triggered thousands of times on the original itself and needs a band-pass (2–20 Hz). Broadband vibration needs thresholds well above the signal's standard deviation to reach 0.05 bytes per sample. Fidelity of a diagnostic measure is two-sided: reconstructions can inflate a fault peak. And the posterior mean of a Bayesian model, a good denoiser, is a source of error in an archive.

Table 1 summarises the third round. The wavelet coder has the lowest error on bearing and seismic data at every budget from 0.1 up (and at 0.05 except on IMS, where decimation is marginally better), and ties decimation on short screwdriving cycles; SZ3 is never better below 0.3 bytes per sample, and ZFP's one-dimensional error-bound mode is not competitive. The bearing fault is detected on the same file as on the original at every budget down to 0.05; SZ3 detects it 172 files later at 0.05 and 17 at 0.1. The weak point is seismic false triggers: thresholding leaves artefacts that STA/LTA reads as onsets, 2.4 times as many triggers away from catalogued arrivals at 0.1 and 12% more at 0.3, so H2 as stated fails on development data at 0.1. Posterior selection was worse than the plain threshold (0.57 against 0.56 at 0.1 and 0.45 against 0.40 at 0.2 on CWRU), so H4 holds on development data. On CNC milling the coder again has the lowest error, bad runs remain about as separable as on the original (AUROC 0.745 against 0.765, at the edge of H2's 0.02), and per-window RMS, the usual wear proxy, deviates by 4% at 0.1 and 1.7% at 0.2; on a machine tool the approximation should travel with exact per-window statistics, which MorletDB already keeps at about 1% of the lossless size. On the Raspberry Pi 3B+ of our earlier edge study, one core of the Python prototype encodes at 13 times real time for 64 kHz with one fixed setting and at 2.1 times with bisection rate control, within 56 MB; the development harness's grid search (56 tries) would not keep up (0.12 times), and bisection, with a continuous threshold, reached lower error (0.67 against 0.71) at the same budget. A fourth round added noise filling and the device's bisection rate control: with noise filling, false seismic triggers fell from 927 to 380 (original 400) and the RMS deviation on CNC data from 9% to 0.05%, at a higher relative error; bisection was faster than a grid search and reached lower error. The protocol then named the decode per measure (plain for waveform error and spectral diagnosis, noise-filled for energy-based detection and trends) and was frozen.

Table 2: Confirmation: median relative error at 0.1 bytes per sample (plain decode), and the decisions.

	Wavelet	SZ3	ZFP	Decim.
XJTU-SY	**0.70**	0.97	1.00	0.76
Paderborn	**0.55**	0.72	1.01	0.63
PyScrew s02	0.078	–	–	**0.070**
Seismic	**0.056**	0.093	–	0.080
Bosch CNC	**0.39**	0.76	–	0.49

H1 holds at 0.05, 0.1, 0.2, 0.3 (4, 5, 5, 4 of 5 datasets)
H2 refuted: (b) seismic 94.8% < 95%; (c) PyScrew AUROC 0.529 vs 0.566
H3 holds: 3.06× real time at 64 kHz, 55.5 MB
H4 holds: Bayesian selection better only once

5 CONFIRMATION

The protocol was frozen in the repository before any confirmation data were loaded [32]. Confirmation data: XJTU-SY run-to-failure bearings with a single outer- or inner-race fault (11 bearings, 25.6 kHz) [45]; Paderborn bearings with real damage and healthy references (320 files, 64 kHz) [26]; PyScrew scenario s02 (12,499 cycles) [46]; 16 station-days of a held-out seismic gallery with a public catalogue [35]; and two further machines of the Bosch CNC data [42]. One run per dataset; two deviations, a rerun after a crash in aggregation (no result had been written) and non-unique recording labels (no measure affected).

H1, error per byte: holds at every budget. The coder beat SZ3 and ZFP on four of five datasets at 0.05 and 0.3 and on all five at 0.1 and 0.2 (Table 2). It lost to SZ3 only on XJTU-SY at 0.3 (0.400 against 0.387) and could not reach 0.05 bytes per sample on most screwdriving cycles, where decimation was slightly better at 0.05 and 0.1.

H2, diagnosis: refuted. On the seven XJTU-SY bearings on which the original data reveal the fault, the approximation detected it no later than allowed, never late at all; SZ3 was late on five, raised a false detection on one and missed another. The approximation was, however, early on four bearings and raised more isolated alarms before the fault (50 against 1 on the longest run): thresholding makes envelope spectra peakier. Bad CNC runs (AUROC 0.671 against 0.682) and damaged Paderborn bearings (kurtosis AUROC 0.220 against 0.201) stayed as separable as on the original. Two criteria failed. With noise filling, seismic

STA/LTA triggered on 91 of the 96 catalogued arrivals the original found (94.8%; the bar was 95%), with 2,813 triggers elsewhere against 2,990; SZ3 and decimation found 90. On PyScrew s02, where kurtosis separates failed cycles only weakly even on the original (0.566), the approximation weakened it to 0.529.

H3, the device: holds. On one core of the Raspberry Pi 3B+, the Python prototype encoded Paderborn's 64 kHz vibration at 3.06 times real time with bisection rate control at 0.099 bytes per sample, peaking at 55.5 MB; the coder with its per-level noise levels ran at 2.53 times.

H4: holds. Posterior selection with a heavy-tailed slab was worse than the plain threshold at every budget on XJTU-SY and better by more than 5% only once on Paderborn.

What this means. As a code per byte for vibration and seismic data at the edge, a plain wavelet threshold with a simple stream beats today's error-bounded compressors, keeps a developing bearing fault visible no later than the original data, and fits a small device. It should travel with exact per-window statistics or be decoded with noise filling for trends, and alarm thresholds should be calibrated on approximated data. It is not a substitute for exact data in trigger-based event detection or on short cycles.

6 RELATED WORK

Wavelet thresholding and its Bayesian forms. Thresholding rules [17, 18, 21] and their Bayesian counterparts with spike-and-slab priors [1, 13, 14, 43, 44], including empirical Bayes with level-wise weights [22–24], heavy-tailed and Pareto-based priors [2, 8, 19, 38, 39] and shrinkage priors outside wavelets [5, 9, 10, 40], target denoising; BayesShrink links shrinkage to compression [12]. In our confirmation the plain threshold was as good or better for an archive.

Compression. Error-bounded lossy compressors [16, 27, 28] and compressors for time series [11, 33] bound the error per value; wavelet compression of vibration data was studied offline [41] and on wireless nodes [48], and compressive sensing [7, 20] moves the work to the decoder.

Edge data reduction. Send-on-delta and event-triggered sampling [31], edge reduction for condition monitoring [30, 36] and tiny models on microcontrollers [6, 25] decide what to send; we measure what an approximation keeps when everything is sent at a fixed budget.

7 LIMITATIONS

One run per dataset on one machine, and one device of one kind. The protocol was frozen by the author's assistant while the author was away, at the author's request. Relative error and five diagnostic proxies stand in for an engineer's judgement; two of the proxies (kurtosis on screwdriving cycles and on Paderborn) separate the classes only weakly even on the original data. The encoder is a Python prototype; a native implementation would be faster.

USE OF AI ASSISTANCE AND RESPONSIBILITY

Code, experiment harnesses, analyses and text were developed with Claude (Anthropic), an AI assistant, working under the author's direction. The author reviewed the content and is responsible for it.

LICENCE

© 2026 Zin Nam. The text and figures of this report are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/); the MorletDB source code is licensed under Apache-2.0.

REFERENCES

[1] F. Abramovich, T. Sapatinas, and B. W. Silverman. 1998. Wavelet Thresholding via A Bayesian Approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 60, 4 (1998), 725–749. https://doi.org/10.1111/1467-9868.00151

[2] A. Achim, P. Tsakalides, and A. Bezerianos. 2003. SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling. *IEEE Transactions on Geoscience and Remote Sensing* 41, 8 (2003), 1773–1784. https://doi.org/10.1109/TGRS.2003.813488

[3] Rex V. Allen. 1978. Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America* 68, 5 (1978), 1521–1532. https://doi.org/10.1785/BSSA0680051521

[4] Jérôme Antoni. 2006. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing* 20, 2 (2006), 282–307. https://doi.org/10.1016/j.ymssp.2004.09.001

[5] Artin Armagan, David B. Dunson, and Jaeyong Lee. 2013. Generalized double Pareto shrinkage. *Statistica Sinica* 23, 1 (2013), 119–143. https://doi.org/10.5705/ss.2011.048

[6] Colby Banbury, Vijay Janapa Reddi, Peter Torelli, Jeremy Holleman, et al. 2021. MLPerf Tiny Benchmark. arXiv:2106.07597 https://arxiv.org/abs/2106.07597

[7] Yuequan Bao, James L. Beck, and Hui Li. 2011. Compressive sampling for accelerometer signals in structural health monitoring. *Structural Health Monitoring* 10, 3 (2011), 235–246. https://doi.org/10.1177/1475921710373287

[8] Larbi Boubchir and Jalal M. Fadili. 2006. A closed-form nonparametric Bayesian estimator in the wavelet domain of images using an approximate α-stable prior. *Pattern Recognition Letters* 27, 12 (2006), 1370–1382. https://doi.org/10.1016/j.patrec.2006.01.008

[9] C. M. Carvalho, N. G. Polson, and J. G. Scott. 2010. The horseshoe estimator for sparse signals. *Biometrika* 97, 2 (2010), 465–480. https://doi.org/10.1093/biomet/asq017

[10] Volkan Cevher. 2009. Learning with Compressible Priors. In *Advances in Neural Information Processing Systems 22 (NIPS 2009)*. https://proceedings.neurips.cc/paper_files/paper/2009/hash/5c936263f3428a40227908d5a3847c0b-Abstract.html

[11] Shubham Chandak, Kedar Tatwawadi, Chengtao Wen, Lingyun Wang, Juan Aparicio Ojea, and Tsachy Weissman. 2020. LFZip: Lossy Compression of Multivariate Floating-Point Time Series Data via Improved Prediction. In *2020 Data Compression Conference (DCC)*. 342–351. https://doi.org/10.1109/DCC47342.2020.00042

[12] S. Grace Chang, Bin Yu, and Martin Vetterli. 2000. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing* 9, 9 (2000), 1532–1546. https://doi.org/10.1109/83.862633

[13] Hugh A. Chipman, Eric D. Kolaczyk, and Robert E. McCulloch. 1997. Adaptive Bayesian Wavelet Shrinkage. *J. Amer. Statist. Assoc.* 92, 440 (1997), 1413–1421. https://doi.org/10.1080/01621459.1997.10473662

[14] Merlise Clyde and Edward I. George. 2000. Flexible Empirical Bayes Estimation for Wavelets. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 62, 4 (2000), 681–698. https://doi.org/10.1111/1467-9868.00257

[15] Ingrid Daubechies. 1992. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611970104

[16] Sheng Di and Franck Cappello. 2016. Fast Error-Bounded Lossy HPC Data Compression with SZ. In *2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. 730–739. https://doi.org/10.1109/IPDPS.2016.11

[17] David L Donoho and Iain M Johnstone. 1994. Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81, 3 (1994), 425–455. https://doi.org/10.1093/biomet/81.3.425

[18] David L. Donoho and Iain M. Johnstone. 1995. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Amer. Statist. Assoc.* 90, 432 (1995), 1200–1224. https://doi.org/10.1080/01621459.1995.10476626

[19] M.A.T. Figueiredo and R.D. Nowak. 2001. Wavelet-based image estimation: an empirical Bayes approach using Jeffrey's noninformative prior. *IEEE Transactions on Image Processing* 10, 9 (2001), 1322–1331. https://doi.org/10.1109/83.941856

[20] Shihao Ji, Ya Xue, and Lawrence Carin. 2008. Bayesian Compressive Sensing. *IEEE Transactions on Signal Processing* 56, 6 (2008), 2346–2356. https://doi.org/10.1109/TSP.2007.914345

[21] Iain M. Johnstone and Bernard W. Silverman. 1997. Wavelet Threshold Estimators for Data with Correlated Noise. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 59, 2 (1997), 319–351. https://doi.org/10.1111/1467-9868.00071

[22] Iain M. Johnstone and Bernard W. Silverman. 2004. Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse sequences. *The Annals of Statistics* 32, 4 (2004), 1594–1649. https://doi.org/10.1214/009053604000000030

[23] Iain M. Johnstone and Bernard W. Silverman. 2005. EbayesThresh: R Programs for Empirical Bayes Thresholding. *Journal of Statistical Software* 12, 8 (2005). https://doi.org/10.18637/jss.v012.i08

[24] Iain M. Johnstone and Bernard W. Silverman. 2005. Empirical Bayes selection of wavelet thresholds. *The Annals of Statistics* 33, 4 (2005), 1700–1752. https://doi.org/10.1214/009053605000000345

[25] Yuma Koizumi, Yohei Kawaguchi, Keisuke Imoto, Toshiki Nakamura, et al. 2020. Description and Discussion on DCASE2020 Challenge Task2: Unsupervised Anomalous Sound Detection for Machine Condition Monitoring. arXiv:2006.05822 https://arxiv.org/abs/2006.05822

[26] Christian Lessmeier, James Kuria Kimotho, Detmar Zimmer, and Walter Sextro. 2016. Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification. In *PHM Society European Conference*, Vol. 3. https://doi.org/10.36001/phme.2016.v3i1.1577

[27] Xin Liang, Kai Zhao, Sheng Di, Sihuan Li, Robert Underwood, Ali M. Gok, Jiannan Tian, Junjing Deng, Jon C. Calhoun, Dingwen Tao, Zizhong Chen, and Franck Cappello. 2023. SZ3: A Modular Framework for Composing Prediction-Based Error-Bounded Lossy Compressors. *IEEE Transactions on Big Data* 9, 2 (2023), 485–498. https://doi.org/10.1109/TBDATA.2022.3201176

[28] Peter Lindstrom. 2014. Fixed-Rate Compressed Floating-Point Arrays. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2674–2683. https://doi.org/10.1109/TVCG.2014.2346458

[29] Stéphane Mallat. 2009. *A Wavelet Tour of Signal Processing* (third ed.). Academic Press. https://doi.org/10.1016/B978-0-12-374370-1.X0001-8

[30] Anna Lito Michala, Ioannis Vourganas, and Andrea Coraddu. 2022. Vibration Edge Computing in Maritime IoT. *ACM Transactions on Internet of Things* 3, 1 (2022), 1–18. https://doi.org/10.1145/3484717

[31] Marek Miskowicz. 2006. Send-On-Delta Concept: An Event-Based Data Reporting Strategy. *Sensors* 6, 1 (2006), 49–63. https://doi.org/10.3390/s6010049

[32] MorletDB contributors. 2026. F2: keeping the spikes — edge approximation and label-free relevance. https://github.com/merkhet-io/morletdb/tree/main/evaluation/edge_relevance_f2

[33] Carlos Enrique Muniz-Cuza, Søren Kejser Jensen, Jonas Brusokas, Nguyen Ho, and Torben Bach Pedersen. 2024. Evaluating the Impact of Error-Bounded Lossy Compression on Time Series Forecasting. In *Proceedings of the 27th International Conference on Extending Database Technology (EDBT)*. OpenProceedings.org, 650–663. https://doi.org/10.48786/edbt.2024.56

[34] Zin Nam. 2026. Machines with a Memory: Episodic Recall and Graded Compaction for Industrial Waveforms. https://github.com/merkhet-io/morletdb/tree/main/research/paper2 Technical report, version 0.2.

[35] Zin Nam. 2026. MorletDB: Recordings, Not Rows — Lag-Tolerant Waveform Retrieval in PostgreSQL. https://github.com/merkhet-io/morletdb/tree/main/research/paper Technical report, version 0.20.

[36] Rick Pandey, Felix Grimm, Dominik Nille, Christoph Böckenhoff, Jonathan Gamez, Sebastian Uziel, Albert Dorneich, Tino Hutschenreuther, and Silvia Krug. 2025. Case Study on Compression of Vibration Data for Distributed Wireless Condition Monitoring Systems. *Applied Sciences* 15, 22 (2025), 12346. https://doi.org/10.3390/app152212346

[37] Robert B. Randall and Jérôme Antoni. 2011. Rolling element bearing diagnostics—A tutorial. *Mechanical Systems and Signal Processing* 25, 2 (2011), 485–520. https://doi.org/10.1016/j.ymssp.2010.07.017

[38] Vincent Rivoirard. 2004. Thresholding procedure with priors based on Pareto distributions. *TEST* 13, 1 (2004), 213–246. https://doi.org/10.1007/BF02603007

[39] Vincent Rivoirard. 2005. Bayesian modelling of sparse sequences and maxisets for Bayes rules. *Mathematical Methods of Statistics* 14, 3 (2005). https://www.ceremade.dauphine.fr/~rivoirar/bayesian.pdf

[40] Veronika Ročková and Edward I. George. 2018. The Spike-and-Slab LASSO. *J. Amer. Statist. Assoc.* 113, 521 (2018), 431–444. https://doi.org/10.1080/01621459.2016.1260469

[41] W.J. Staszewski. 1998. Wavelet based compression and feature selection for vibration analysis. *Journal of Sound and Vibration* 211, 5 (1998), 735–760. https://doi.org/10.1006/jsvi.1997.1380

[42] Mohamed-Ali Tnani, Michael Feil, and Klaus Diepold. 2022. Smart Data Collection System for Brownfield CNC Milling Machines: A New Benchmark Dataset for Data-Driven Machine Monitoring. *Procedia CIRP* 107 (2022), 131–136. https://doi.org/10.1016/j.procir.2022.04.022

[43] Brani Vidakovic. 1998. Nonlinear Wavelet Shrinkage with Bayes Rules and Bayes Factors. *J. Amer. Statist. Assoc.* 93, 441 (1998), 173–179. https://doi.org/10.1080/01621459.1998.10474099

[44] Brani Vidakovic. 1999. *Statistical Modeling by Wavelets*. Wiley. https://doi.org/10.1002/9780470317020

[45] Biao Wang, Yaguo Lei, Naipeng Li, and Ningbo Li. 2020. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Transactions on Reliability* 69, 1 (2020), 401–412. https://doi.org/10.1109/TR.2018.2882682

[46] Nikolai West and Jochen Deuse. 2025. Industrial screw driving dataset collection: Time series data for process monitoring and anomaly detection. https://doi.org/10.5281/zenodo.15393134 Version 1.2.2, CC BY 4.0.

[47] Mitchell Withers, Richard Aster, Christopher Young, Judy Beiriger, Mark Harris, Susan Moore, and Julian Trujillo. 1998. A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America* 88, 1 (1998), 95–106. https://doi.org/10.1785/BSSA0880010095

[48] Ning Xu, Sumit Rangwala, Krishna Kant Chintalapudi, Deepak Ganesan, Alan Broad, Ramesh Govindan, and Deborah Estrin. 2004. A wireless sensor network For structural monitoring. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*. 13–24. https://doi.org/10.1145/1031495.1031498