



Comparative Evaluation of Chatter Detection Indicators under Simulated Machining Conditions

Enrique MIRELES HERNANDEZ^(a, b), Mikhail GUSKOV^(a), Philippe LORONG^(a), Théo Dorlin^(b), Habib KARAOUNI^(b)

(a) Laboratoire PIMM, Arts et Metiers Institute of Technology, 151 boulevard de l'Hopital, 75013 Paris, France

(b) Safran S.A., Research & Technology Center, F-78772 Magny-les-Hameaux, France

Mail : enrique.mireles_hernandez@ensam.eu

Abstract: Early detection of chatter would potentially enable reducing surface degradation, tool wear, and machine tool degradation. While numerous time- and frequency-domain indicators exist, their relative sensitivity and robustness are rarely assessed under a unified framework. This work presents a systematic benchmark of chatter detection indicators using controlled time-domain simulations of machining dynamics. The study compares representative approaches based on vibration energy trends, phase space geometry, statistical signal descriptors, and other time-domain features, including a novel energy based indicator. All methods are tested within identical simulation conditions and signal processing settings. The evaluation focuses on three aspects: the earliest detectable sign of instability, robustness to measurement noise, and consistency across varying cutting parameters. The resulting performance map provides insight into how different indicators respond to chatter onset and under what conditions they remain reliable. The study aims to clarify the trade-offs among existing methods and to guide the selection of effective, real-time or offline chatter detection strategies for stable and productive machining operations.

Keywords: Chatter detection, Early detection, Time-domain simulation, Indicator benchmarking, Signal processing, Machining stability,

1 Introduction

Chatter is a self-excited vibration that arises when the tool–workpiece interaction enters an unstable regime dominated by regenerative effects and machine structural dynamics (Quintana 2011; Altintas 2012). It increases forces and displacements, degrades accuracy and surface finish, and can lead to tool breakage or spindle damage (Zhu 2020). Chatter occurs in conventional milling and turning as well as in low-stiffness configurations such as robotic machining, where structural flexibility amplifies instability (Wang 2016). Despite extensive work on dynamic models and stability diagrams (Altintas 1995), the time-varying and nonlinear nature of cutting limits purely analytical prediction under realistic conditions, motivating online monitoring within TCM and intelligent-machine frameworks (Li 2017).

Several signal sources have been exploited for chatter detection, including structural vibration, cutting force, and radiated sound (Delio 1992). Acoustic sensors are non-intrusive and low-cost, whereas cutting force directly reflects the tool–workpiece interaction. The goal is not only to identify fully developed chatter but to detect as early as possible the transition from stable cutting to incipient slight chatter, where regenerative oscillations start to grow while remaining masked by noise and forced excitation (Yang 2019). Such early detection is especially critical in high-speed machining and low-stiffness systems (Zhu 2020) and motivates algorithms able to classify the process state online under noisy, non-stationary conditions (Chen 2019). Existing methods are usually grouped into four families: time-domain indicators (e.g. RMS, kurtosis, cycle energy), frequency-domain approaches based on FFT or spectral density (Tangjitsitcharoen 2013), time–frequency techniques such as STFT, wavelets, or Hilbert–Huang, and machine-learning approaches using features from vibration, force, or sound (Cao 2017b).

However, strong heterogeneity in machines, structural stiffness, tools, materials, cutting parameters, noise levels, and signal-processing pipelines hinders fair comparison and generalization across studies (Nath 2020), and reliable detection of slight chatter embedded in noise remains an open problem (Tansel 2012). In this context, there is still no unified comparative framework assessing, under controlled and reproducible conditions, the ability of different indicators to detect chatter onset. The present work addresses this gap through a systematic benchmark based on time-domain simulations. Unlike studies focused on a single method and configuration (Zhou 2018), multiple indicators are compared here on identical signals, cutting conditions, and noise levels. The set includes representative time-domain, statistical, time–frequency, and energy-based metrics, together with a novel phase-space-area indicator tailored to slight chatter. All indicators are evaluated with a common protocol that quantifies their sensitivity to early instability, robustness to measurement noise, and consistency under varying cutting parameters, yielding a comparative performance map of their behaviour around chatter onset.

2 Time-domain cutting model and simulation campaign

Clarification: Objectif: Expliquer d'où on obtient les données

2.1 Dynamic cutting model

The simulated signals are obtained from a classical regenerative cutting model (Altintas 2012). The cutting setup is shown in (Figure 1), and the theoretical stability lobes in (Figure 2b). The

simulation parameters are listed in (Table 1).

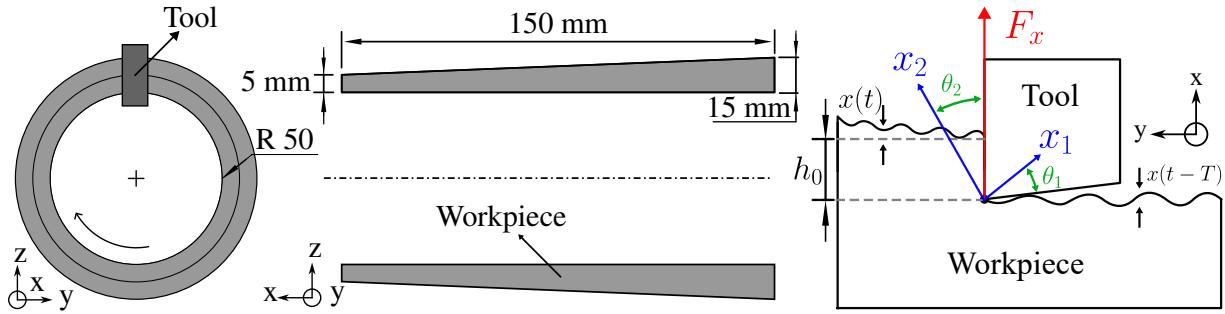


Figure 1 Two-degrees-of-freedom (2-DOF) system for dynamic simulations, used to generate Stability Lobe Diagrams (SLD).

The tool-workpiece dynamics are modelled with a single vibration mode and a linear cutting-force law acting on the instantaneous chip thickness. The system is therefore governed by

$$m\ddot{x}(t) + c\dot{x}(t) + kx(t) = K_f a(t) [h_0 + x(t - T) - x(t)], \quad (1)$$

Table 1 Modal parameters used in the simulation.

Mode	ω_n (Hz)	ζ (%)	k (N/m)	m (kg)	θ (degree)
x_1	250	1.2	2.26×10^8	9.57	30°
x_2	150	1.0	2.13×10^8	15.48	60°

2.2 Numerical implementation

All signals are generated with the machining simulator *Nessy2m*, which provides tool-tip displacement, velocity, and force for the cutting setup shown in (Figure 2a). A single scenario is used: a depth-of-cut ramp from 5 mm to 15 mm at 120,000 rpm and 0.05 mm/tooth (feed rate 6000 mm/min). The sampling rate ensures 200 samples per revolution (400 kHz), and the total duration is 15 s.

Only the dominant 150 Hz mode is activated, and no measurement noise is added. Most indicators use the velocity signal, while the proposed phase-space energy indicator requires both displacement and velocity.

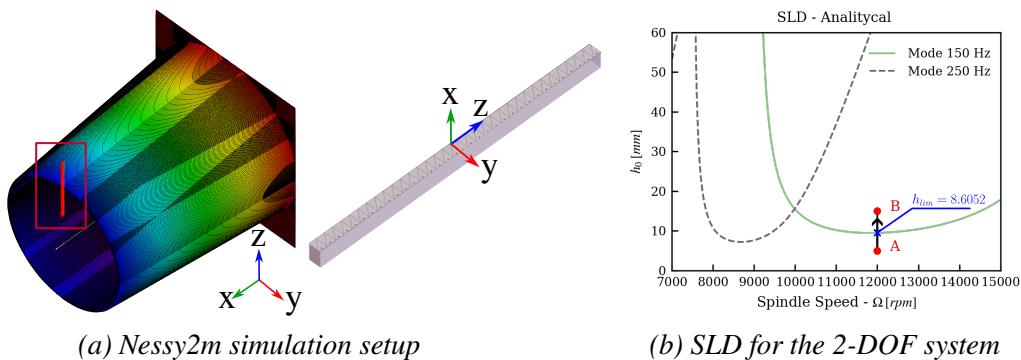


Figure 2 Nessy2m simulation setup and corresponding Stability Lobe Diagram (SLD).

2.3 Simulation campaign

A single high-resolution time-domain simulation is used for all indicators to ensure a fair and fully controlled comparison. The depth-of-cut ramp provides a smooth transition from stable cutting to regenerative instability, with the stability limit crossed at approximately $h_{lim} = 8.6052$ mm (Figure 2b) around $t_E = 7.32$ s (Section 4.1). This allows all indicators to be evaluated on the same, well-defined chatter onset. Each method is therefore tested under identical conditions using exactly the same vibration signal.

Clarification: Objectif: Décrire les indicateurs qu'on va tester

3 Chatter indicators considered in this benchmark

A minimal but representative set of chatter indicators is selected to cover the main mechanisms linked to early instability. All methods are applied to the same tool-tip velocity signal $v(t)$ under a unified comparison protocol. Only the essential functional definitions are reported here; full algorithmic details can be found in the original references.

3.1 EMD–Hilbert indicator (force-based surrogate)

This indicator isolates high-frequency content that becomes excited as chatter develops. First, empirical mode decomposition (EMD) extracts IMF_1 , the intrinsic mode concentrating the highest frequencies in $v(t)$. Regenerative chatter injects energy into a specific frequency band, causing IMF_1 to grow. The Hilbert transform produces the analytic signal, from which the instantaneous amplitude and Hilbert spectrogram are derived. Chatter manifests as an increasing number of instantaneous-amplitude points within the predefined chatter-sensitive band. The progression is: raw signal → IMF isolation → Hilbert analytic signal → instantaneous amplitude in chatter band (Wang 2018).

$$\underbrace{\text{IMF}_1(t) = \text{EMD}\{v(t)\}_1}_{\text{Mode extraction}} \longrightarrow \underbrace{z_1(t) = \text{IMF}_1(t) + i \mathcal{H}\{\text{IMF}_1(t)\}}_{\text{Hilbert analytic signal}} \longrightarrow \underbrace{I_{\text{EMD}}(t) = |z_1(t)|}_{\text{Instantaneous amplitude in chatter band}}.$$

3.2 MaxEnt–SPRT indicator (cycle-based probabilistic)

This indicator transforms the vibration into progressively richer statistical representations. A once-per-revolution (OPR) value v_k captures cycle-to-cycle behavior directly influenced by regenerative dynamics. The maximum entropy (MaxEnt) of short OPR segments quantifies the spread of the distribution, which increases during slight chatter. The sequential probability ratio test (SPRT) then accumulates deviations relative to a baseline stable model. An essential offline calibration phase determines the MaxEnt probability densities for stable $p_0(\text{ME})$ and incipient-chatter $p_1(\text{ME})$ conditions (Zhao 2019).

$$\underbrace{v_k = v(t_k)}_{\text{OPR sample}} \longrightarrow \underbrace{\text{ME}(k) = - \int f_k(v) \ln f_k(v) dv}_{\text{MaxEnt (cycle dispersion)}} \longrightarrow \underbrace{I_{\text{SPRT}}(n) = \sum_{k=1}^n \ln \left[\frac{p_1(\text{ME}(k))}{p_0(\text{ME}(k))} \right]}_{\text{Sequential indicator (cumulative evidence)}}.$$

3.3 SST–SVD indicator (*synchrosqueezed time–frequency energy*)

This indicator reveals chatter by progressively refining the spectral structure of the vibration signal. First, the Synchrosqueezing Transform (SST) sharpens the time–frequency (TF) representation. Second, spindle- and tooth-passing harmonics are removed to expose energy growth near structural frequencies. Third, the filtered TF map is decomposed via SVD: (1) construct the TF window, (2) compute its singular value spectrum, and (3) track the largest singular value. In the original formulation, detection relies on a 3σ threshold, requiring a stable segment to estimate mean and variance (Cao 2017a).

$$\underbrace{T_v(f, t) = \text{SST}\{v(t)\}}_{\text{Sharpened TF map}} \longrightarrow \underbrace{T_v^{\text{filt}}(f, t)}_{\substack{\text{TPF} \\ \text{harmonic removal}}} \longrightarrow \underbrace{(U, D, V) = \text{SVD}(T_v^{\text{filt}}(\cdot, t_i))}_{\text{TF decomposition}} \longrightarrow \underbrace{I_{\text{SST}}(t_i) = D_{11}}_{\substack{\text{Dominant TF energy} \\ (\text{chatter growth})}}.$$

3.4 Coefficient of Variation (CV)

This indicator measures normalized amplitude variability. The velocity signal is transformed into a sliding-window RMS sequence, capturing envelope fluctuations. As chatter develops, the RMS values become increasingly irregular. The CV normalizes this variability and provides a simple and robust instability metric. Each new RMS value is computed by discarding the oldest sample and adding the newest one (Ye 2018).

$$\underbrace{v_{\text{RMS}}(k) = \sqrt{\frac{1}{N} \sum_{i=1}^N v_i^2}}_{\text{Envelope extraction (sliding window)}} \longrightarrow I_{\text{CV}} = \underbrace{\frac{\text{std}(v_{\text{RMS}})}{\text{mean}(v_{\text{RMS}})}}_{\text{Normalized variability}}.$$

3.5 Phase-space spiral-area indicator (*proposed*)

This indicator evaluates the mechanical energy flow in the phase-space trajectory $(x(t), v(t))$. A sliding observation window containing five complete oscillation cycles is used. For each cycle, the spiral area quantifies net energy exchange: areas shrink when damping dominates (stable cutting) and expand when regenerative feedback injects energy (incipient chatter). The logarithmic decrement between consecutive areas captures the energy-growth rate (Mireles 2025).

$$\underbrace{A_j = \frac{1}{2} \sum_{\ell} (x_{\ell} v_{\ell+1} - x_{\ell+1} v_{\ell})}_{\text{Cycle energy (spiral area)}} \longrightarrow \underbrace{\delta_j = -\ln\left(\frac{A_{j+1}}{A_j}\right)}_{\text{Energy-growth rate}} \longrightarrow \underbrace{I_{\text{phase}}(t_k) = \delta^{(k)}}_{\text{Instability indicator over five cycles}}.$$

4 Benchmark protocol and performance metrics

Clarification: Objectif: Expliquer comment on va évaluer les indicateurs

4.1 Ground truth and reference regions

4.1.1 Linear stability boundary and t_E

The ground truth is obtained from the linearised regenerative 1-DOF model (Eq. 1). Linear perturbations $x(t) = X e^{\lambda t}$ lead to the characteristic equation

$$m\lambda^2 + c\lambda + k + K_f a(t)(1 - e^{-\lambda T}) = 0. \quad (2)$$

At each instant, the dominant root $\lambda_{\max}(a(t))$ is extracted and its real part

$$\sigma(t) = \Re\{\lambda_{\max}(a(t))\} \quad (3)$$

is evaluated numerically.

The time at which the system crosses the stability boundary is defined as

$$t_E : \quad \sigma(t_E) = 0, \quad \sigma(t) < 0 \quad \forall t < t_E. \quad (4)$$

4.1.2 Amplitude growth and $t_{\text{vis}}(R)$

After t_E , the vibration envelope satisfies

$$A(t) \approx A(t_E) \exp(G(t)), \quad G(t) = \int_{t_E}^t \sigma(\tau) d\tau. \quad (5)$$

The visibility time $t_{\text{vis}}(R)$ is defined as the first instant where the accumulated growth reaches

$$G(t_{\text{vis}}(R)) = \ln R, \quad (6)$$

providing a noise-free, dynamics-based visibility threshold.

Under the depth-of-cut ramp $a(t) = a_{p,0} + \dot{a}_p t$, the times t_E and $t_{\text{vis}}(R)$ correspond to depths a_E and a_{vis} on the stability diagram. Both t_E and $t_{\text{vis}}(R)$ are model-based references common to all indicators.

4.1.3 Visibility levels R

Three levels are used:

- $R = 3$ (+9.5 dB): very early chatter,
- $R = 5$ (+14 dB): intermediate growth,
- $R = 10$ (+20 dB): clearly visible chatter.

The dependence of $t_{\text{vis}}(R)$ on R remains moderate due to the slow variation of $\sigma(t)$.

4.1.4 Temporal regions

Using t_E and $t_{\text{vis}}(R)$, the benchmark timeline is partitioned into

$$\mathcal{R}_{\text{st}} = [t_{\text{st},1}, t_{\text{st},2}], \quad (\text{strictly stable: } t < t_E), \quad (7a)$$

$$\mathcal{R}_{\text{EC}}(R) = [t_E, t_{\text{vis}}(R)], \quad (\text{early chatter: } t_E \leq t \leq t_{\text{vis}}(R)), \quad (7b)$$

$$\mathcal{R}_{\text{ch}}(R) = [t_{\text{vis}}(R), t_{\text{end}}], \quad (\text{developed chatter: } t > t_{\text{vis}}(R)). \quad (7c)$$

False alarms are evaluated in \mathcal{R}_{st} and early detections in $\mathcal{R}_{\text{EC}}(R)$.

4.1.5 Dimensionless time and normalised detection index

A unified time scale is introduced as

$$\xi(t) = \frac{t - t_E}{t_{\text{vis}}(R) - t_E}, \quad (8)$$

where $\xi = 0$ corresponds to the stability boundary and $\xi = 1$ to the chosen visibility level. If an indicator triggers at t_d , its normalised index is

$$\eta = \xi(t_d). \quad (9)$$

Interpretation: $\eta < 0$ (premature), $0 < \eta < 1$ (early chatter), $\eta > 1$ (late). This normalisation removes scale, smoothing and unit dependence, enabling direct comparison across heterogeneous indicators.

4.2 Unified comparison layer and detection time

Each indicator retains its native formulation, but all decision rules are anchored to the strictly stable region \mathcal{R}_{st} , which provides a method-specific stable baseline. Indicators whose original design already relies on this notion (SST–SVD via its 3σ envelope, MaxEnt–SPRT via p_0) are used as published. For CV, EMD–Hilbert and the spiral-area ratio, the values observed in \mathcal{R}_{st} define respectively a reproducible CV threshold, a stable energy level, and a zero-drift trend for the logarithmic area ratio. Chatter is declared when the indicator departs persistently from its baseline according to its native decision logic. This yields comparable detection times t_d without modifying the intrinsic behaviour of any method.

4.3 Performance metrics

Using the normalised time $\xi(t)$ and detection index η , four quantities characterise each indicator.

Detection delay. η provides the primary sensitivity measure ($\eta < 0$ premature, $0 < \eta < 1$ early, $\eta > 1$ late). Absolute delay: $\Delta t = t_d - t_E$.

False alarms. With activations \mathcal{A} in the strictly stable region $\mathcal{R}_{\text{st}} = [t_{\text{st},1}, t_{\text{st},2}]$,

$$\text{FAR} = \frac{|\mathcal{A} \cap \mathcal{R}_{\text{st}}|}{|\mathcal{R}_{\text{st}}|}. \quad (10)$$

Stable-region variability. For the raw indicator $I(t)$,

$$\sigma_{\text{st}}^2 = \text{Var}(I(t); t \in \mathcal{R}_{\text{st}}), \quad \delta_{\text{st}} = \max_{t \in \mathcal{R}_{\text{st}}} |I(t) - \bar{I}_{\text{st}}|. \quad (11)$$

Early-chatter slope. Responsiveness across $\mathcal{R}_{\text{EC}} = [t_E, t_{\text{vis}}(R)]$ is

$$\gamma = \frac{I(t_{\text{vis}}) - I(t_E)}{t_{\text{vis}} - t_E}. \quad (12)$$

These metrics compactly quantify sensitivity, false alarms, and baseline stability under a unified time scale.

5 Results

6 Discussion

7 Conclusion and outlook

Acknowledgments

The authors would like to thank their collaborators and sponsors for supporting this research effort.

References

- Altintas, Y. (2012). Manufacturing Automation: Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design. 2nd ed. Cambridge: Cambridge University Press. doi: [10.1017/CBO9780511843723](https://doi.org/10.1017/CBO9780511843723).
- Altintas, Y. and E. Budak (1995). “Analytical prediction of stability lobes in milling”. In: CIRP Annals–Manufacturing Technology 44.1, pp. 357–362. doi: [10.1016/S0007-8506\(07\)62342-7](https://doi.org/10.1016/S0007-8506(07)62342-7).
- Cao, H., Y. Yue, X. Chen, and X. Zhang (2017a). “Chatter detection in milling process based on synchrosqueezing transform of sound signals”. In: International Journal of Advanced Manufacturing Technology 89.9–12, pp. 2747–2755. doi: [10.1007/s00170-016-9660-7](https://doi.org/10.1007/s00170-016-9660-7).
- Cao, H., K. Zhou, X. Chen, and X. Zhang (2017b). “Early chatter detection in end milling based on multi-feature fusion and 3σ criterion”. In: International Journal of Advanced Manufacturing Technology 92.9–12, pp. 4387–4397.

- Chen, J., P. Hu, H. Zhou, J. Yang, J. Xie, Y. Jiang, Z. Gao, and C. Zhang (2019). "Toward intelligent machine tool". In: Engineering 5, pp. 679–690. doi: [10.1016/j.eng.2019.07.018](https://doi.org/10.1016/j.eng.2019.07.018).
- Delio, T., J. Tlusty, and S. Smith (1992). "Use of audio signals for chatter detection and control". In: ASME Journal of Engineering for Industry 114.2, pp. 146–157.
- Li, Z., Y. Wang, and K. S. Wang (2017). "Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario". In: Advanced Manufacturing 5, pp. 377–387. doi: [10.1007/s40436-017-0203-8](https://doi.org/10.1007/s40436-017-0203-8).
- Mireles, E., M. Guskov, and P. Lorong (2025). "Early Chatter Detection Through Real-Time Supervision and Time-Domain Simulation-Based Indicators". In: MM Science Journal. doi: [10.17973/MMSJ.2025_11_2025147](https://doi.org/10.17973/MMSJ.2025_11_2025147).
- Nath, C. (2020). "Integrated tool condition monitoring systems and their applications: a comprehensive review". In: Procedia Manufacturing 48, pp. 852–863. doi: [10.1016/j.promfg.2020.05.123](https://doi.org/10.1016/j.promfg.2020.05.123).
- Quintana, G. and J. Ciurana (2011). "Chatter in machining processes: a review". In: International Journal of Machine Tools and Manufacture 51, pp. 363–376. doi: [10.1016/j.ijmachtools.2011.01.001](https://doi.org/10.1016/j.ijmachtools.2011.01.001).
- Tangjitsitcharoen, S., T. Saksri, and S. Ratanakuakangwan (2013). "Advance in chatter detection in ball end milling process by utilizing wavelet transform". In: Journal of Intelligent Manufacturing 26.3, pp. 485–499. doi: [10.1007/s10845-013-0805-3](https://doi.org/10.1007/s10845-013-0805-3).
- Tansel, I. N., M. Li, M. Demetgul, K. Bickraj, B. Kaya, and B. Ozcelik (2012). "Detecting chatter and estimating wear from the torque of end milling signals by using Index Based Reasoner (IBR)". In: International Journal of Advanced Manufacturing Technology 58.1–4, pp. 109–118.
- Wang, G., H. Dong, Y. Guo, and Y. Ke (2016). "Dynamic cutting force modeling and experimental study of industrial robotic boring". In: International Journal of Advanced Manufacturing Technology 86, pp. 179–190.
- (2018). "Early Chatter Identification of Robotic Boring Process Using Measured Force of Dynamometer". In: International Journal of Advanced Manufacturing Technology 94.1–4, pp. 1243–1252. doi: [10.1007/s00170-017-0941-6](https://doi.org/10.1007/s00170-017-0941-6).
- Yang, K., G. Wang, Y. Dong, Q. Zhang, and L. Sang (2019). "Early chatter identification based on an optimized variational mode decomposition". In: Mechanical Systems and Signal Processing 115, pp. 238–254.
- Ye, J., P. Feng, C. Xu, Y. Ma, and S. Huang (2018). "A Novel Approach for Chatter Online Monitoring Using Coefficient of Variation in Machining Process". In: International Journal of Advanced Manufacturing Technology 96.1–4, pp. 287–297. doi: [10.1007/s00170-017-1544-y](https://doi.org/10.1007/s00170-017-1544-y).
- Zhao, Y., K. H. Adjallah, A. Sava, and Z. Wang (2019). "Early Chatter Detection Using MaxEnt and SPRT". In: 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT). Paris, France: IEEE, pp. 1550–1555. doi: [10.1109/CoDIT.2019.8820670](https://doi.org/10.1109/CoDIT.2019.8820670).
- Zhou, Y. and W. Xue (2018). "Review of tool condition monitoring methods in milling processes". In: International Journal of Advanced Manufacturing Technology 96, pp. 2509–2523. doi: [10.1007/s00170-018-1768-5](https://doi.org/10.1007/s00170-018-1768-5).
- Zhu, L. and C. Liu (2020). "Recent progress of chatter prediction, detection and suppression in milling". In: Mechanical Systems and Signal Processing 143, p. 106840. doi: [10.1016/j.ymssp.2020.106840](https://doi.org/10.1016/j.ymssp.2020.106840).