Chatter Detection Based on ARMAX Model-Based Monitoring Method in Thin Wall Turning Operation

Yang Liu and Zhenhua Xiong(✉)

School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
mexiong@sjtu.edu.cn

Abstract. The ARMAX model-based monitoring method is proposed for chatter detection in thin wall turning operation. ARMAX modelling is deduced to fit the cutting force for the time varying dynamic process caused by the stiffness variation of thin wall workpiece. Residuals closely connected with chatter are extracted and monitored by control charts in real time. Cutting experiments for two different depth of cuts are performed for verification, from which it is found that the ARMAX model-based monitoring process has lower false alarm rate than the typical ARMA modelling. In addition, the model parameters affected by the varying stiffness is also time dependent, so the RELS parameter estimation algorithm is employed. The forgetting factor in the RELS algorithm is optimized through experiments to further reduce the false alarm rate. It is observed that the RELS ARMAX model-based algorithm with forgetting factors between 0.85 and 0.9 has the best monitoring performance with zero false alarms.

Keywords: Chatter · Thin wall turning · Stiffness variation · ARMAX · RELS

1 Introduction

Three different types of mechanical vibrations, known as free vibrations, forced vibrations and self-excited vibrations, are generally found in metal cutting processes. Researchers have developed several effective techniques to reduce or eliminate the occurrence of the free and forced vibrations. The self-excited vibration (i.e. chatter) is mainly caused by the regenerative effect between the cutting tool and the workpiece and brings the machining process to instability, which is the most threatening and uncontrollable type of vibration. The external reason for regenerative chatter is the lack of stiffness or damping of the machine tool, the tool holder, the cutting tool and the workpiece material, in case of the machining of thin wall workpieces, deep holes and long overhang workpieces. For this reason, the mechanism analysis, detection and suppression techniques of the regenerative chatter have been a popular topic for academic and industrial research.

The automatic detection of regenerative chatter is of vital significance to avoid damage to workpieces or machine tools. A number of studies have been reported about the development of tool condition monitoring (TCM) techniques, which are focused on

Z. Chen et al. (Eds.): ICIRA 2018, LNAI 10984, pp. 319–330, 2018.
https://doi.org/10.1007/978-3-319-97586-3_29

the abnormal state monitoring and diagnosis caused by tool wear, tool breakage as well as regenerative chatter [1–8]. One of the TCM approaches is time series analysis, in which a specific time series modelling is employed to the measured signals, such as cutting forces, vibration, or acoustic emission. A popular solution for linear time series monitoring involves autoregressive (AR) or autoregressive moving average (ARMA) modelling.

The AR(1) modelling was used to the acoustic emission signal for online tool wear monitoring. The AR parameters and power of the AR residual signals were selected as features and found to be effective in tool condition monitoring [9]. A time series based tooth period modelling technique was proposed for detecting tool breakage by monitoring a cutting force or torque signal, which can also be used for chatter detection or tool wear evaluation as stated [10]. A tool breakage detection algorithm was developed after removing the transient process by AR(1) filtering [11]. The AR(5) model was fitted to the cutting forces and it was found that the power of the residual signal was most effective and consistent. The AR parameters classified by pattern recognition technique had showed distinct response towards tool wear [12]. The AR(1) model was deduced to the amplitudes of the relevant frequencies of the drilling torque signals, while chatter vibration was detected with respect to the residual signals based on a ranked EWMA control chart [13]. The relationship between the mechanical model of cutting process and its corresponding time series model was discussed, in which the ARMA(4,3) model was derived to deal with the acceleration vibration signals [14]. The ARMA modelling was employed to the tool vibration signals in a turning process. The ARMA distance calculated based on eigenvalues of the tool/holder system, was applied in an online tool wear estimation algorithm [15].

The time series modelling had been widely investigated for TCM, but few researches have focused the time dependent property of the model used in case of flexible cutting tool or workpiece. In this work, the ARMAX modelling is developed based on the time varying depth of cut caused by stiffness variation and dynamic vibration of the thin wall workpiece. This model is more appropriate and more accurate to describe the cutting process after considering the exogenous input of stiffness. In addition, the RELS parameter estimation algorithm is employed to further increase the accuracy for online model identification. At last, the ARMAX modelling with residual control chart is performed for thresholding or decision-making online.

2 Time Series Modelling of Thin Wall Turning Operation

2.1 Stiffness Variation

The experimental setup is shown in Fig. 1, where the turning experiments are performed on a CK 6150A lathe and the cutting force signal is recorded with a Kistler 9257B dynamometer sampled at 20 kHz. The thin wall disc workpiece with two-stepped shape is shown in Fig. 2. As the cutting tool moves along the feed direction from inner to outer of the free-end, the axial stiffness is debilitated due to the cantilever structure of thin wall workpiece. In addition, the removal of workpiece material causes the stiffness variation slightly.

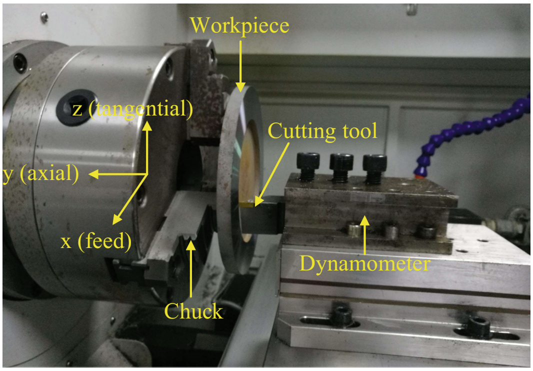

Fig. 1. Turning experimental setup.

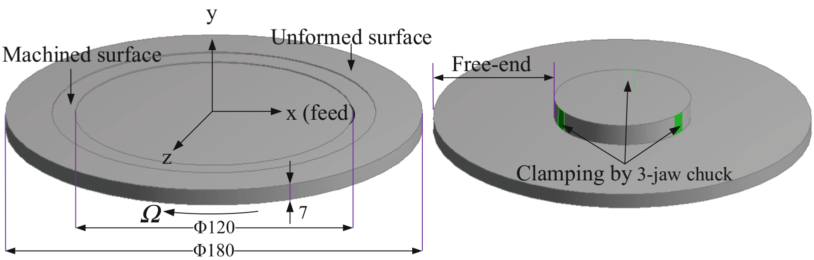

Fig. 2. Thin wall disc workpiece (All dimensions are in mm).

In this work, the stiffness variation of workpiece is obtained offline by finite element simulation before machining. An axial moving unit static force is imposed at the tool nose as the cutting goes on. The axial static deformation is simulated, then the axial stiffness is calculated. However, only integral locations on the machined surface are considered due to the time-consuming process of simulation, and the whole stiffness curve is obtained by spline interpolation. The axial stiffness for different depth of cuts

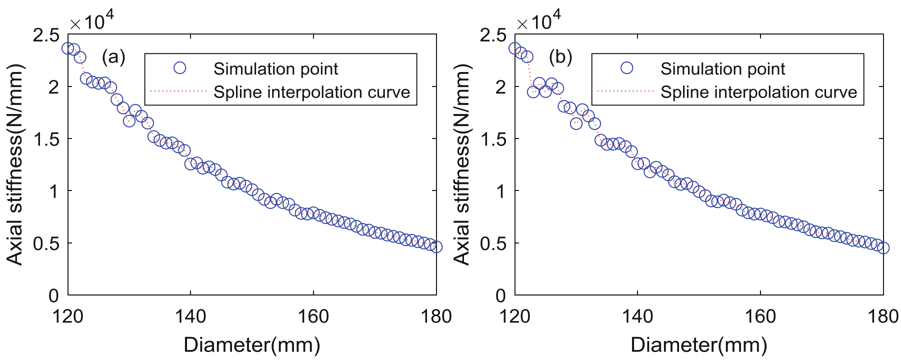

Fig. 3. Axial stiffness for different depth of cuts (a) 0.05 mm; (b) 0.1 mm.

is shown in Fig. 3, from which it is observed that the axial stiffness is reduced severely as the cutting tool moves along the feed direction.

2.2 ARMAX Modelling of the Cutting Force

It was reviewed that the cutting force was the best signal for chatter detection compared with acoustic emission and acceleration signals, and the characteristic patterns of cutting force variation make it possible to clearly distinguish chatter [16]. Apart from chatter, the cutting forces are also sensitive to other cutting conditions such as cutting parameters, temperature, hardness of material and dynamics behavior of cutting tool, making correlation with chatter more complicated. In this work, time varying process dynamics of the thin wall turning operation due to stiffness variation of the workpiece is mainly considered, in order to obtain the key information which is closely connected with chatter.

A simple and widely accepted equation of the cutting force is expressed as [17]

$$F_{\mathrm{y}} = K_{\mathrm{y}} a h \tag{1}$$

where F_{y} and K_{y} are the cutting force and the cutting coefficient in the axial direction, respectively. a and h are the depth of cut and the feed rate, respectively.

Chatter will occur even under low cutting parameters at the free-end of the thin wall disc. Thus, a small axial depth of cut, which is far less than the tip radius, is assumed to analysis the tool displacement. The cutting tool is regarded as rigid and the workpiece is flexible in the axial direction. It is noticed that the actual depth of cut is varying during cutting process. When the second revolution starts, there are undulations, caused by chatter vibration and weak stiffness, both in the inner machined surface of current revolution and the outside surface of previous revolution. As shown in Fig. 4, the cutting area ah is changed to the gray part, which can be expressed as

$$ah \approx \sum \Delta S_i = h_0 \left[a_0 - \frac{\left(y_{t-\tau} + 2y_{t-\tau+1} + \cdots + 2y_{t-\tau+1} + y_t \right)}{2n} \right] \tag{2}$$

where a_0 and h_0 are the selected depth of cut and feed rate, respectively. n is the number of sampling points in one revolution period τ. y_t is the feed displacement at time t.

The axial workpiece displacement is assumed to be divided into two parts, i.e. the static deflection and the dynamic vibration.

$$y_t = y_t^{\mathrm{d}} + y_t^{\mathrm{s}} \tag{3}$$

here

$$y_t^{\mathrm{s}} = \frac{F_t}{k_t} \tag{4}$$

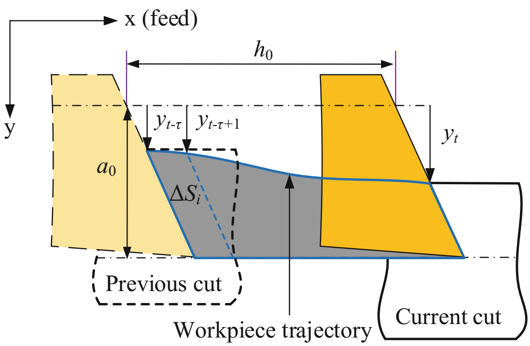

Fig. 4. Dynamic cutting process of thin wall turning in one revolution period.

where the superscripts, s and d represent the static deflection and the dynamic vibration. k_t is the axial workpiece stiffness at current time.

The workpiece trajectory from A to B is nonlinear due to the time varying dynamics of thin wall turning operation. For simplification, the sum of tool displacements in Eq. (2) may be represented only by the displacements at present time and one spindle revolution period before, which can be expressed as

$$ah \approx h_0 \left[a_0 - \frac{\left(p_1 y^{\mathrm{s}}_{t-\tau} + p_2 y^{\mathrm{d}}_{t-\tau} + p_3 y^{\mathrm{s}}_t + p_4 y^{\mathrm{d}}_t \right)}{2} \right] \tag{5}$$

where p_i are the unsolved parameters. Substituting Eqs. (3–5) into Eq. (1) and rearranging the result as

$$\left(\frac{2}{K_{\mathrm{c}} h_0} + \frac{p_3}{k_t} \right) F_t + \left(\frac{p_1}{k_{t-\tau}} \right) F_{t-\tau} = 2a_0 - p_2 y^{\mathrm{d}}_{t-\tau} - p_4 y^{\mathrm{d}}_t \tag{6}$$

As mentioned, the cutting force signals usually contain some detrimental components due to external cutting conditions such as stiffness variation in this case. It can surely influence the performance of common used chatter detection methods. In contrast, the model in Eq. (6) can extract the significative information hiding in the complex cutting force signals. In the stable state, the dynamic vibration is assumed to be normally distributed as white noise, thus time series modelling can be performed [18]. Previous observations are substituted by current data using the shaft operator q as

$$F_{t-\tau} = q^{-1} F_t,\ y^{\mathrm{d}}_{t-\tau} = q^{-1} y^{\mathrm{d}}_t \tag{7}$$

Substitution Eq. (7) into Eq. (6) yields

$$\left[1 + \frac{p_1}{P_t k_{t-\tau}} q^{-1} \right] F_t = \frac{2a_0 K_{\mathrm{c}} h_0 k_t}{2k_t + p_3 K_{\mathrm{c}} h_0} + \left(-\frac{p_2}{P_t} q^{-1} - \frac{p_4}{P_t} \right) e_t \tag{8}$$

here

$$P_t = \frac{2}{K_c h_0} + \frac{p_3}{k_t},\ e_t = y_t^d \tag{9}$$

where $-p_4/P_t$ is assumed to be 1 and other parameters will be estimated by model fitting.

Then it is observed that Eq. (8) has the same structures to the ARMAX model (AutoRegressive Moving Average with eXogenous input) whose general form is

$$\begin{aligned} &A(q)y_t = B(q)u_t + C(q)e_t \\ &A(q) = 1 + a_1 q^{-1} \cdots + a_{na} q^{-na} \\ &B(q) = b_1 + b_2 q^{-1} \cdots + b_{nb} q^{-nb+1} \\ &C(q) = 1 + c_1 q^{-1} \cdots + c_{nc} q^{-nc} \end{aligned} \tag{10}$$

The output y_t is the measured axial cutting force, while the exogenous input u_t is proportional to the cutting coefficient, the depth of cut, the feed rate and the current axial stiffness as follows.

$$u_t = 2K_c a_0 h_0 k_t \tag{11}$$

In addition, the orders for polynomials $A(q)$, $B(q)$ and $C(q)$ are 1, 0 and 1, respectively. The model parameters are also time varying with respect to the stiffness variation as follows.

$$a_1 = \frac{p_1}{P_t k_{t-\tau}},\ b_0 = \frac{1}{2k_t + p_3 K_c h_0},\ c_1 = -\frac{p_2}{P_t} \tag{12}$$

3 Chatter Detection

3.1 Online Model Identification and Forecast

For the ARMAX model, all parameters to be estimated are arranged in a vector ϑ, while the relevant input-output data are collected into the regressor vector φ as

$$\vartheta_t = \left[a_1 \cdots a_{na}\, b_1 \cdots b_{nb-1}\, c_1 \cdots c_{nc}\right]^T,\ \varphi_t = \left[-y_{t-1} \cdots - y_{t-na}\, u_t \cdots u_{t-nb+1}\, \bar{\varepsilon}_{t-1} \cdots \bar{\varepsilon}_{t-nc}\right]^T \tag{13}$$

where $\bar{\varepsilon}_t = y_t - \varphi_t^T \hat{\vartheta}_t$ represents the residual.

The parameters need to be identified using an appropriate estimation algorithm. Due to the time varying properties of the process, the adaptive recursive extended least square (RELS) algorithm may be applied to perform online model identification [19]. The RELS approach allows the estimate to be updated at each step, by correcting it by a term proportional to the innovation brought about by the last collected observation. In other words, the estimate at time t can be obtained from the estimate at time t-1, without the need to explicitly reconsider all of the past data. The RELS algorithm is expressed as

$$
\begin{aligned}
\hat{\vartheta}_t &= \hat{\vartheta}_{t-1} + \gamma_t R_t^{-1} \varphi_t \varepsilon_t \\
R_t &= R_{t-1} + \gamma_t \left(\varphi_t \varphi_t^T - R_{t-1} \right) \\
\varepsilon_t &= y_t - \varphi_t^T \hat{\vartheta}_{t-1}
\end{aligned}
\tag{14}
$$

where γ_t is the gain which has a constant value γ in the following algorithm. ε_t is the prediction error.

For a time dependent system that changes gradually, the recent measurements are usually considered more reliable and the older ones are no longer representative of the process dynamics [20]. In order for the estimation algorithm to progressively "forget" older data, the most common choice is to take a constant forgetting factor λ. In this way, the most recent data are given unitary weight in the minimization process, while an exponentially time decreasing weight is assigned to past measurements. The parameters will be estimated by minimizing the sum of the exponential weighted least squares of the prediction errors as follows.

$$
\vartheta_t = \arg \min_{\vartheta} \sum_{k=1}^{t} \lambda^{t-k} \varepsilon_t^2 \tag{15}
$$

The relationship between the loss factor and the constant gain is $\gamma + \lambda = 1$. With a predefined forgetting factor, all of the autoregressive, moving average and exogenous parameters can be estimated with the measured cutting force and simulated stiffness. Then the forecast for residuals is calculated by

$$
\bar{\varepsilon}_t = F_t + \hat{a}_1 F_{t-1} - \hat{b}_1 u_t - \hat{c}_1 \bar{\varepsilon}_{t-1} \tag{16}
$$

3.2 Thresholding Based on Residual Control Chart

It is crucial to set up an appropriate threshold for decision-making of chatter detection, while the traditional determination of thresholds is to design preliminary cutting experiments and find a constant threshold between stable state and chatter state empirically. However, this method has limitation in choosing the threshold and is less reliable and less adaptive for variable cutting conditions.

Statistical process control or control charts, which is focused on the inherent characteristic of data series, are usually used to monitor the statistical state and alarm out-of-control in quality engineering. The advantage of the statistical process control is to determine the threshold non-empirically and update it automatically. It should also be noted that control charts are built based on the basic assumption that the process is independently and normally distributed. However, this assumption rarely holds in practice. In thin wall tuning, any sensor signal acquired during one spindle revolution is expected to be characterized by some level of autocorrelation due to the time varying process dynamics. The autocorrelation in axial cutting forces is mainly caused by stiffness variation, vibration and other reasons include cutting impact, tool wear and temperature. Control charts should not be used directly, because more false alarms will occur for data series with high level autocorrelation. An approach that has proved useful in dealing with autocorrelated data is residual control chart [21]. Firstly, an appropriate time series model is fitted to the data series to remove the autocorrelation from the data. Then the control charts like Shewhart charts can be applied to the residuals, where the control limits or thresholds are calculated as

$$\left.\begin{array}{l}\text{Upper control limit: } \mu_t + L_t\sigma_t \\ \text{Center line: } \mu_t \\ \text{Lower control limit: } \mu_t - L_t\sigma_t\end{array}\right\} \tag{17}$$

where μ_t and σ_t are the mean and standard deviation of the residuals. The width of control limits L_t is generally constant as 3, which is also well known as the 'three-sigma' criterion.

4 Experimental Validation

The measured axial cutting forces for different depth of cuts are shown in Fig. 5. Larger forces are noted at the entry point probably due to the impact when the tool cuts into the workpiece, then in the stable state it appears to be distributed as a white noise. It has a considerable oscillation around zero when the cutting state turns into chatter. The cutting forces change dramatically between stable and chatter states. This in-between state is generally called the onset state when the cutting forces have slight fluctuations and no visible chatter marks are generated on the workpiece before chatter has fully developed. These three cutting states can be identified artificially from experiments based on the noise and the surface appearance. It is observed that chatter occurs about 10 s earlier at 0.1 mm than that at 0.5 mm.

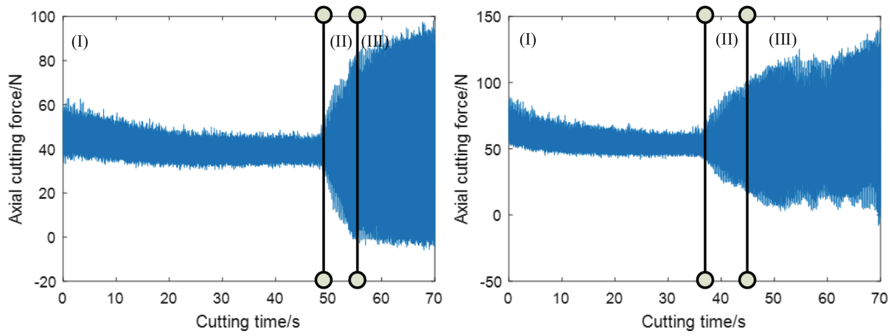

Fig. 5. Axial cutting forces in real cutting experiments under spindle speed of 220 rpm, feed rate of 0.1 mm/r, depth of cuts of 0.05 mm (left) and 0.1 mm (right) (I) Stable state; (II) Onset state; (III) Chatter state.

4.1 ARMAX Model-Based Monitoring

The outputs of the ARMAX model, i.e. measured cutting forces, are extracted in a period of one spindle revolution about 0.27 s, which means from all sampling points there are

5455 available groups of data. Different groups make little influence, so the first group is chosen uniformly for the following algorithm. In contrast to the previous research, the ARMA model-based monitoring is also performed.

Both the ARMA model and the ARMAX model are identified online by least squares (LS) algorithm every time a new data is collected. Then the residuals are obtained and control limits are calculated by the Shewhart chart with the time histories of residuals from the beginning to the current time. Without other criterions, the process is regarded as in-control (i.e. stable state) when the observations remain inside the control limits and out-of-control (i.e. onset state or chatter state) when they fall outside the upper or lower control limits. The numbers of out-of-control for the Shewhart charts with ARMAX or ARMA modelling and without modelling are shown in Fig. 6.

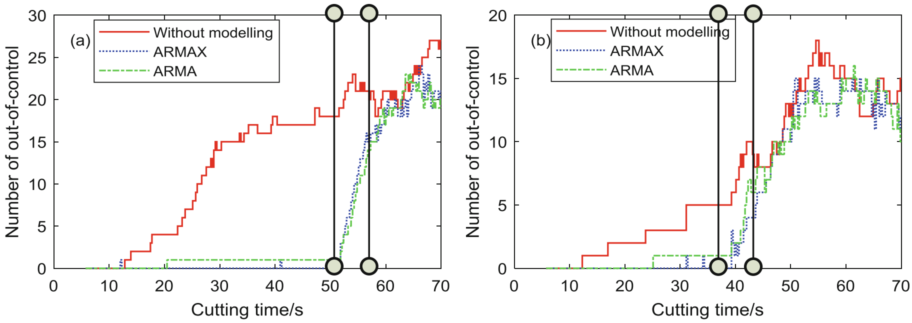

Fig. 6. Monitoring processes with and without modelling (a) 0.05 mm; (b) 0.1 mm.

It is found that there are too many false alarms for monitoring process without modelling, so direct monitoring of the cutting force signal is infeasible in this case. Although a steep increase of the out-of-control number occurs in the onset state both with ARMA and ARMAX modelling, the least out-of-control number is observed in the stable state for the ARMAX model-based monitoring processes.

The false alarm rates are calculated as ratios between the times of out-of-control number larger than 0 in the stable state and the sampling times of the whole stable state. The results are shown in Table 1, where the false alarm rates of ARMAX model-based monitoring processes are reduced substantially to small values under 2%.

Table 1. False alarm rate (%) with and without model-based monitoring processes.

Time series modelling	Depth of cut/mm	
	0.05	0.1
No model fitting	84.34	80.33
ARMAX	1.87	1.64
ARMA	67.47	41.8

4.2 RELS ARMAX Model-Based Monitoring

As mentioned in Eq. (12), the ARMAX model parameters are also time dependent due to the stiffness variation, so the RELS algorithm with 'forget' capacity may be more appropriate for parameter estimation. The key coefficient of the RELS algorithm need to be defined is the forgetting factor. The residuals will be forecasted recursively by the RELS ARMAX modelling with different forgetting factors, then the cutting process is monitored by the residual control chart. The compassion of false alarm rates between RELS ARMAX model-based monitoring processes with different forgetting factors is shown in Fig. 7.

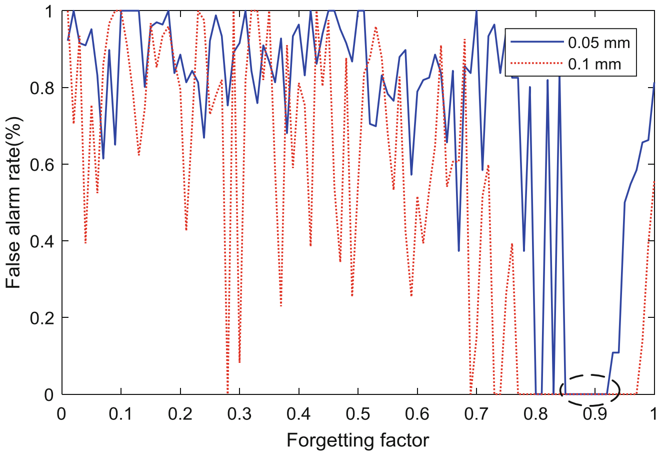

Fig. 7. False alarm rate (%) of RELS ARMAX model-based monitoring processes with different forgetting factors.

Enormous differences between the false alarm rates with different forgetting factors are observed. Zero false alarm rate can be achieved, while large false alarm rates are also noticed, which means the forgetting factor should be carefully selected. A zone of optimized forgetting factor, i.e. 0.85 to 0.9, is found through the experimental results of two depth of cuts. The monitoring performance of the RELS ARMAX model-based method with optimized forgetting factor is further improved than that of the LS ARMAX modelling.

5 Conclusion

In the case of thin wall turning operation, an ARMAX model for the cutting force is deduced by analyzing the dynamic cutting process caused by stiffness variation and dynamic vibration. The model residuals are monitored and the chatter state is identified by the residual control charts. The main contribution of this work is summarized as

1. The residuals from ARMAX modelling of the cutting force are determinant information closely connected with chatter, which are monitored by the residual control charts. This ARMAX model-based monitoring process shows better performance than the typical ARMA modelling in real cutting experiments.
2. Base on the time varying properties of the model, time dependent parameters of ARMAX model are estimated online by the RELS parameter estimation algorithm. The false alarm rate of the RELS ARMAX model-based monitoring process with an experimental optimized forgetting factor is further decreased to zero, which may be an effective chatter detection method for engineering application.

References

1. Tansel, I.N., Wagiman, A., Tziranis, A.: Recognition of chatter with neural networks. Int. J. Mach. Tools Manuf **31**(4), 539–552 (1991)
2. Wu, Y., Du, R.: Feature extraction and assessment using wavelet packets for monitoring of machining processes. Mech. Syst. Signal Process. **10**(1), 29–53 (1996)
3. Govekar, E., et al.: A new method for chatter detection in grinding. CIRP Ann. Manuf. Technol. **51**(1), 267–270 (2002)
4. Tangjitsitcharoen, S., Pongsathornwiwat, N.: Development of chatter detection in milling processes. Int. J. Adv. Manuf. Technol. **65**(5–8), 919–927 (2013)
5. Sun, Y., Zhuang, C., Xiong, Z.: Real-time chatter detection using the weighted wavelet packet entropy. In: IEEE/ASME International Conference on IEEE, pp. 1652–1657 (2014)
6. Nouri, M., et al.: Real-time tool wear monitoring in milling using a cutting condition independent method. Int. J. Mach. Tools Manuf **89**, 1–13 (2015)
7. Sun, Y., Xiong, Z.: An optimal weighted wavelet packet entropy method with application to real-time chatter detection. IEEE/ASME Trans. Mechatron. **21**(4), 2004–2014 (2016)
8. Fu, Y., et al.: Timely online chatter detection in end milling process. Mech. Syst. Signal Process. **75**, 668–688 (2016)
9. Ravindra, H.V., Srinivasa, Y.G., Krishnamurthy, R.: Acoustic emission for tool condition monitoring in metal cutting. Wear **212**(1), 78–84 (1997)
10. Tansel, I.N., McLaughlin, C.: Detection of tool breakage in milling operations—I. The time series analysis approach. Int. J. Mach. Tools Manuf **33**(4), 531–544 (1993)
11. Altintas, Y.: In-process detection of tool breakages using time series monitoring of cutting forces. Int. J. Mach. Tools Manuf **28**(2), 157–172 (1988)
12. Kumar, S.A., Ravindra, H.V., Srinivasa, Y.G.: In-process tool wear monitoring through time series modelling and pattern recognition. Int. J. Produc. Res. **35**(3), 739–751 (1997)
13. Messaoud, A., Weihs, C., Hering, F.: Detection of chatter vibration in a drilling process using multivariate control charts. Comput. Stat. Data Anal. **52**(6), 3208–3219 (2008)
14. Song, D.Y., et al.: A new approach to cutting state monitoring in end-mill machining. Int. J. Mach. Tools Manuf **45**(7–8), 909–921 (2005)
15. Aghdam, B.H., Vahdati, M., Sadeghi, M.H.: Vibration-based estimation of tool major flank wear in a turning process using ARMA models. Int. J. Adv. Manuf. Technol. **76**(9–12), 1631–1642 (2015)
16. Tlusty, J., Andrews, G.C.: A critical review of sensors for unmanned machining. CIRP Ann. Manuf. Technol. **32**(2), 563–572 (1983)
17. Altintas, Y.: Manufacturing Automation: Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design, 2nd edn. Cambridge University Press, New York (2012)

18. Brockwell, P.J., Davis, R.A.: Introduction to Time Series and Forecasting. Springer, New York (2016). https://doi.org/10.1007/b97391
19. Paleologu, C., Benesty, J., Ciochina, S.: A robust variable forgetting factor recursive least-squares algorithm for system identification. IEEE Signal Process. Lett. **15**, 597–600 (2008)
20. Fung, E.H.K., et al.: Modelling and prediction of machining errors using ARMAX and NARMAX structures. Appl. Math. Modell. **27**(8), 611–627 (2003)
21. Messaoud, A., Weihs, C.: Monitoring a deep hole drilling process by nonlinear time series modeling. J. Sound Vib. **321**(3–5), 620–630 (2009)