Journal of Manufacturing Processes 22 (2016) 161–168

Contents lists available at ScienceDirect

Journal of Manufacturing Processes

journal homepage: www.elsevier.com/locate/manpro

Technical Paper

Tool life predictions in milling using spindle power with the neural network technique

Cyril Drouillet[a,b], Jaydeep Karandikar[a,c], Chandra Nath[a,1], Anne-Claire Journeaux[a,b], Mohamed El Mansori[b], Thomas Kurfess[a,*]

[a] George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30309, USA
[b] Arts et Métiers ParisTech, Aix-en-Provence 13617, France
[c] GE Global Research, Niskayuna, NY, USA

ARTICLE INFO

Article history:
Received 22 July 2015
Received in revised form 4 February 2016
Accepted 12 March 2016
Available online 31 March 2016

Keywords:
Tool life
Tool condition monitoring
Neural network
End milling
Spindle power signal
Uncertainty

ABSTRACT

Tool wear is an important limitation to machining productivity and part quality. In this paper, remaining useful life (RUL) prediction of tools is demonstrated based on the machine spindle power values using the neural network (NN) technique. End milling tests were performed on a stainless steel workpiece at different spindle speeds and spindle power was recorded. The NN curve fitting approach with different MATLAB™ training functions was applied to the root mean square power (P_{rms}) values. Sample P_{rms} growth curves were generated to take into account uncertainty. The P_{rms} value in the time domain was found to be sensitive to tool wear. Results show a good agreement between the predicted and true RUL of tools. The proposed method takes into account the uncertainty in tool life and the percentage increase in nominal P_{rms} value during the RUL prediction. Using MATLAB™ on an Intel i7 processor, the computation takes 0.5 s Thus, the method is computationally inexpensive and can be incorporated for real time RUL predictions during machining.

1. Introduction

In the manufacturing industry, controlling the process to machine a workpiece is vital for enhancing both productivity and quality. There are limitations that can influence the manufacturing process, such tool wear and life, surface roughness, surface location error, and machining stability. Tool wear that occurs due to its interaction with the workpiece is a major limitation to machining processes. Tool failure in the machine can lead to unscheduled downtime and damage to the workpiece. It is estimated that the average machine tool downtime due to tool wear is 7–20% [1,2], which can result in a significant loss of productivity.

Tool wear is considered stochastic and in general, difficult to predict. There exist many empirical/mechanistic models to predict tool life based on the Taylor tool life equation [3]. However, the models are deterministic and do not consider the underlying uncertainty in tool wear. Furthermore, many experiments are necessary to determine the model constants. Manufacturing enterprises always demand for higher productivity, and the automation of the monitoring process is an effective way to meet this demand. A generic methodology described in [4] for using an intelligent automatic process monitoring system is simplified as shown in Fig. 1, and that can be applied in predicting the tool condition. Tool condition monitoring (TCM) can significantly reduce downtime, and improve productivity and quality by using sensors such as dynamometer, accelerometer, acoustic emission (AE) sensors, and current/power sensors. The use of artificial intelligence (AI) algorithms to process the sensor data is widely implemented in the TCM [5], and the most commonly used techniques include neural networks, neuro-fuzzy models, fuzzy logic, and Bayesian networks. The idea is to train a system to learn from its past behavior and use it to make predictions on tool life.

Modern machine tools are loaded with power sensor in order to track the machine and tool conditions, among others. Taking the advantage of such sensor technology, current demand of (near) real-time data processing can be enhanced for advancing digital manufacturing. However, a literature review presented in the next section clearly reveals that knowledge of using power sensor for TCM and tool life prediction is still very limited. Thus, power sensor data processing using neural networks for predicting remaining useful tool life is the main focus of this research.

* Corresponding author.
E-mail address: kurfess@gatech.edu (T. Kurfess).
[1] Currently works at Hitachi America Ltd., Research and Development Division, Farmington Hills, 34500 Grand River Ave, MI 48335, USA.

http://dx.doi.org/10.1016/j.jmapro.2016.03.010
1526-6125/

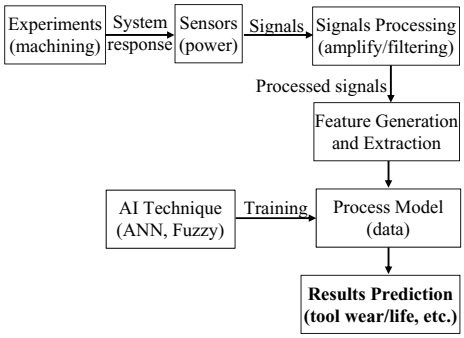

Fig. 1. Methodology for an intelligent process monitoring system.

2. Literature review: data sensors

The selection of the sensors, sensory features, and modeling approach is a crucial step for accurate monitoring of the process. Sensors can be separated in two types:

1) direct measurements such as lasers, optical microscopes, and optical and ultra-sonic sensors which measure the actual dimensions of the worn area on the tool [6];
2) indirect measurements as dynamometers, accelerometers, and current sensors which measure the signal that can be correlated with tool condition [7].

2.1. Direct sensors

The use of direct sensors provides a direct estimation of the tool wear because of the higher reflective properties of the worn part of the tool. The benefit of these sensors is the measurement of any kind of wear: crater, flank, and notch. Since it is a direct measurement, an accurate assessment of the tool state can be computed directly from the data generated by these sensors [8–11]. Similar systems have been used using the reflectance of the cutting chip surface [12]. In [8], a combination between a vision sensor and a neural network provides the learning of the tool state to develop an unsupervised TCM. However, direct sensors have two major limitations. First, direct sensors cannot be typically deployed for on-line TCM because of the environment (*e.g.*, coolant and machining chip). Second, TCM using direct sensors is not real-time because the measurements can only be done in the tool magazine and measured data have to be keyed in processing unit manually for further processing.

2.2. Indirect sensors

2.2.1. Dynamometer

Force dynamometer is commonly used for both off-line and on-line TCM [13–15]. This device provides a direct acquisition of cutting force, which is the best variable to describe the cutting process [16]. The force data are used to correlate the amplitude increase over machining data, force coefficients, and friction coefficients with tool wear. Dynamometer could also be used to detect critical tool breakage appearing in the signal features as a peak. The use of neural network combining with dynamometer has been done, and provides a fast decision making method for tool wear [13]. An accuracy of 90.7% was reported in term of breakage detection [15]. Although promising results were reported using dynamometers, it is intrusive and not feasible in a production environment [17]. Investment and maintenance of force dynamometer are also burdens in machining productivity.

2.2.2. Accelerometer

The frictional forces between the tool and the workpiece, and internal fracture of the inserts cause vibration. When tool wear rises above a certain value, there is an increase in the cutting forces, which in turn, increase the vibration amplitudes. For this reason, vibration signals analysis provides good tool breakage detection [18–20]. The drawbacks of using accelerometer are as follows [21]:

- machining speed should remain within a specific range,
- amplitude of the signal decreases with an increase between the sensor and the cutting edge,
- the mounting position of the accelerometer can change values, and
- the environment is not good for this kind of sensor (coolant, and chip strike).

2.2.3. Acoustic emission (AE) sensors

During the machining of a workpiece, noises come from the machine due to the stresses between the tool and the workpiece, chip fracture, chip friction, and tool breakage. The advantage of AE sensor is that the range of frequency is much higher than the environmental frequencies, making it a less-intrusive sensor. AE sensors have been successfully used for tool wear monitoring [22] and for tool breakage detection [23]. But despite their low cost and the facilities for the installation, this device is not recommended for a production environment since AE calibration is very important to avoid overload and non-voluntary noises. In addition, the mounting and location of the sensor affects the values of the sensor.

2.2.4. Power sensors

The power consumption of an AC motor is directly proportional to it torque resulting in a correlation between the power and the cutting forces exist. Subsequently, power sensor has shown to have a correlation with tool wear [24,25]. In [26], a threshold updating strategy for tool condition monitoring is presented which can deal with variable cutting operations. For turning and drilling, spindle power was used for detection of critical amount of tool wear (VB_B max), and the successful classification rates were 96% and 93%, respectively. Power sensor is mainly used as a complementary sensor for systems as sensor fusion combined with neural network, to improve signal reliability, and help TCM to reach good results [27].

This paper focuses on TCM to predict remaining useful tool life (RUL) using an approach that considers power sensor data to be processed and trained by the neutral network technique. The power sensor is cheap and non-intrusive making it suitable for tool condition monitoring. Moreover, many modern machine tools have power sensors implemented. This review of literature showed that there is a lack of studies using power sensor for TCM and tool life prediction. Since the future of machining seems to be a TCM unsupervised with intelligent systems, the power sensor along with the neural network method is used in this report to demonstrate the efficacy for tool wear diagnosis and RUL prediction.

3. Neural networks

3.1. Principle and overview

The neural network (NN) technique is based on the same function as a neuron in the brain. A simple neuron can be described as shown in Fig. 2. Each input X_n can be set as a vector with an information x_n and a weight w_n. The information passing through a transfer function is considered as an output [28]. As illustrated in Fig. 3, the Multi-Layer Perceptron (MLP) consists of an input layer of neurons, one or more hidden layers of neurons, and an output layer of neurons.

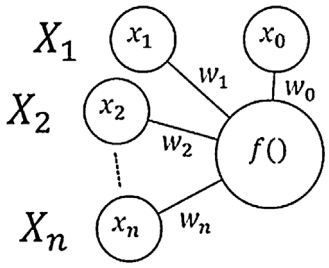

Fig. 2. A simple perception of the NN technique.

The supervised back-propagation learning algorithm using gradient descent search is a method commonly chosen. Training error of this network, or the mean squared error (MSE), is defined between the actual output and the target value.

$$E = \frac{1}{2}\sum (t_k - y_k)^2 \quad (1)$$

where t_k and y_k are the target and actual outputs of the kth neuron of the output layer. To optimize the performance of the neural network, a performance factor of the back-propagation neural network is required. The MSE between the network's output and the desired target is easy to calculate and relevant for the performance:

$$\text{MSE} = \frac{1}{m_1 q_{out}} \sum_{m=1}^{m_1} \sum_{k=1}^{q_{out}} e_k^2(m) \quad (2)$$

where q_{out} is the number of neurons of the output layer, and $e_k^2(m)$ is the error of the kth output neuron for the mth weight. To maximize the performance, minimization of MSE with respect to weights can be performed according to training algorithms in batch or incremental styles.

Artificial neural networks (ANNs) can be used to solve curve-fitting problems because functions can be stored [29]. The objective is to identify and create a relationship between the input and the output data. When the relation is modeled with the necessary accuracy, the neural network can be used as a function approximation. Instead of experimental data used in non-physics-based fitting method such as polynomial fitting (that offers model trend), ANNs work based on experience from these data. During machining, especially metallic alloys and superalloys, statistical models seem to be not that promising, because the tool can fail at any time under the same cutting condition. For example, one tool can fail at 10 min while another at 18 min. So, instant forecasting or knowing the trend based on the machining environment is not possible using such empirical or statistical model. Hill et al. [30], in their review of these two counterparts, stated that both the methods could be working based on the experimental data, but the NNs can forecast the result with a minimum time-lapse. A trained NN can be assumed to be an 'expert' for data analysis with the gained information and project the new situation [31]. Also, the NNs can be adapted and changed easily if the problem changes, whereas the statistical models are less flexible. Finally, some machining

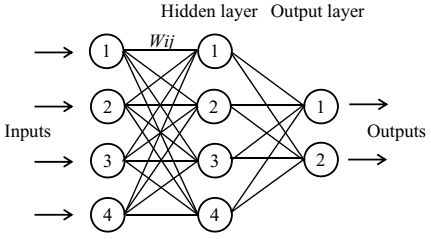

Fig. 3. Architecture of a neural network.

phenomena become different at different size scales with the input parameters (*e.g.*, polynomial at low feeds, exponential at higher speeds) and such mathematical models cannot handle the entire spectrum. Neural networks can implicitly detect complex nonlinear relationships between dependent and independent variables and their interactions [32]. Once at least one data set is trained, the NNs should be able to predict the future output during machining.

Because of these significant advantages, among others, the NNs are thought to be used in the production line in digital manufacturing where real-time data analysis is a vital key to productivity. It acts as a black box with an input-output relation equal to the relationship found during the training. With an arbitrary input, an output can be found with the best fitting possible corresponding to the trained relation. However, since NN techniques assume the forecasting results based on the received data characteristics and size, the predicted error may be larger in the beginning of the process due to a smaller data size received. Continuous feeding of the data helps better training themselves for a better prediction of the future output (error converging trend). Explicit limitations include its "black box" nature, computational burden with software requirement like MATLAB, proneness to over-fitting, and the empirical nature of model development [32]. Also, if the environment changes like noise or new coolant during machining, data should be reproduced and loaded for the training, at least one set to start with.

3.2. Best training function and network configuration

Based on the previous example, a study was conducted to find the best training function and configuration in the hidden layer for the neural network in the case of a curve fitting problem. Using the toolbox of MATLAB, the five following training functions were used:

- *Traingdm* based on the additive momentum algorithm,
- *Traingda* based on the self-adaptive learning rate algorithm,
- *Traincgb* based on the conjugate gradient backpropagation algorithm, and
- *Trainbfg* and *Trainlm* both based on the quasi-Newton algorithm.

For each function, to study the impact of the size of the hidden layer on the neural network, tests with [5810] neurons were computed. After analysis of these results, some observations are made as follows:

- In term of accuracy, the Levenberg–Marquardt (lm) is the best training function followed by the BFGS quasi-Newton backpropagation (bfg) training function, and the additive momentum algorithm (gdm) is the worst (see Fig. 4).
- The Gradient descent with momentum backpropagation algorithm is not fitted to this kind of problem, because the computation is stopped by the discovery of a local minimum.
- The *Trainlm* function is time expensive. Therefore, on an on-line TCM, it can cause some issues.
- The *Trainbfg* offers the best compromise between computation time and accuracy, and that could be used for on-line TCM (see Fig. 5).
- The size of the hidden layer does not matter for the accuracy with at least 5 neurons, but the less number of neurons there reduces the computation time [33].

In summary, to reach the best accuracy, the *Trainlm* function is used with the neural network. Moreover, to reduce the computation time, a hidden layer with 5 neurons is used first since it does not significantly reduce the accuracy.

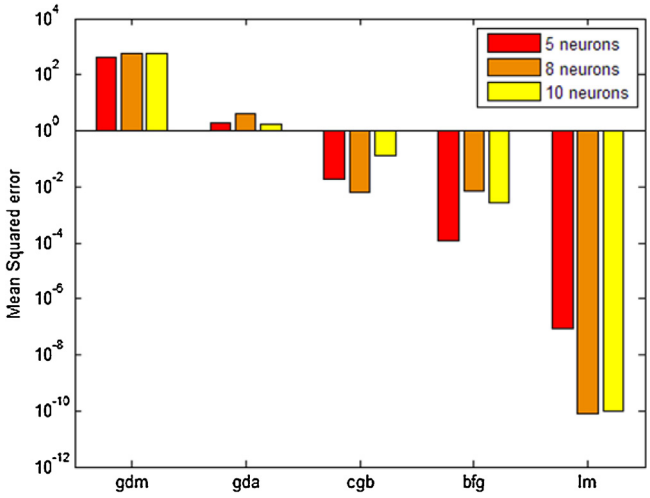

Fig. 4. Mean squared error (MSE) level using the training function and the hidden layer size.

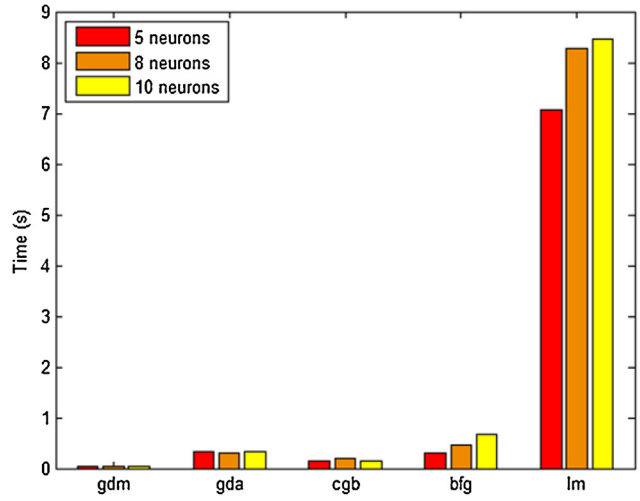

Fig. 5. Computation time of a training using the training function and the hidden layer size.

4. Machining experiments and results

4.1. Machining setup and procedure

In this section, the experimental steps followed to collect tool wear data for a 25.4 mm diameter inserted end mill, Stellram C7792VXP06CA1.0Z4R6.1 (high feed milling cutter) are described. Fig. 6 shows the machining setup on the Okuma Millac 44V vertical machining canter with the Kennametal Stellram tool holder. Four milling carbide inserts (product XPLT060308ER-D41) are inserted in the holder. The workpiece material was a block of stainless steel SS403. The down milling tests were completed at four spindle speeds, Ω, of 1380, 1800, 2700, 4100 rpm (cutting speeds 110, 144, 215, 327 m/min, respectively) with a 0.5 mm axial depth of cut, 5 mm radial depth of cut (~20% radial immersion), and 0.6 mm/tooth feed rate. A through-tool flood coolant at 50 bar pressure was used for cooling and lubricating of the tool-workpiece contact zone. Spindle speeds were selected based on stable cutting conditions at the selected axial and radial depth of cut.

The spindle power was monitored during cutting using a Universal Power Cell connected to the National Instruments cDAQ data

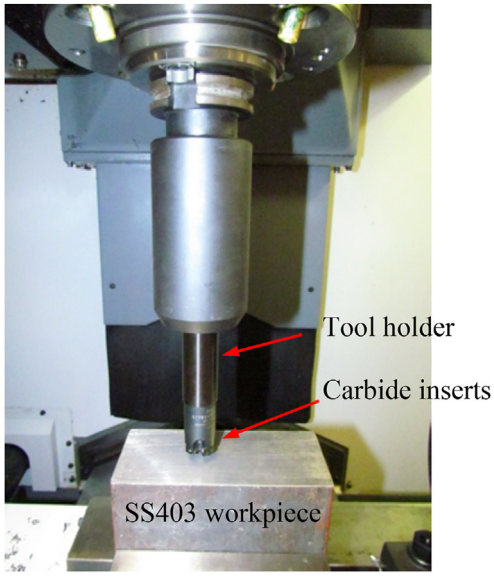

Fig. 6. Machining set up on the Okuma machining center.

acquisition module. The RMS power, P_{rms}, in the time domain was sampled for each layer cut. The insert wear profile was measured after regular intervals. Tool life was defined as the time required for the maximum flank wear width, *FWW*, to reach 0.3 mm. Calibrated digital images were used to identify the *FWW*. The first measurement was performed after the first pass to determine the initial rapid rise in *FWW*.

4.2. Machining results

The sensor P_{rms} values (reported in volts) were normalized by the initial value such that all subsequent sensor values were shown as a percentage of the nominal value. The purpose for this method is to enable the same scale to be used across all cutting processes and machining parameters [34]. This method is also used in commercial systems such as ARTIS. Figs. 7 and 8 show the normalized P_{rms} values against machining time and FWW at different spindle speeds.

With these data, the true remaining useful life (RUL) of the tool was deduced for each of these tests. Then, to verify the accuracy, the neural network was used at two speeds: 1800 and 4100 rpm, respectively, with a curve fitting between the power curve in input and the true RUL in target. After the training of each of these neural networks, the RUL at 2700 rpm was predicted. Figs. 9 and 10 show these results and compare them to the true RUL.

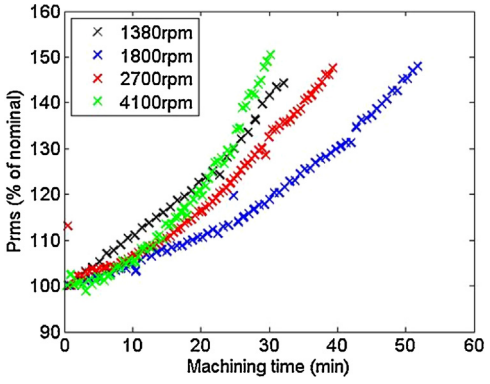

Fig. 7. Plot of normalized power against machining time.

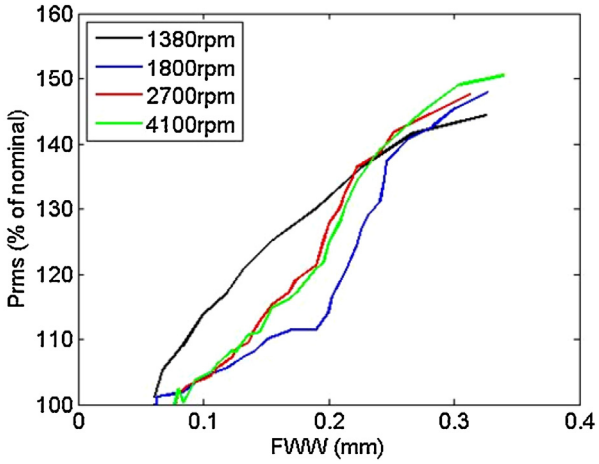

Fig. 8. Plot of normalized power against FWW values.

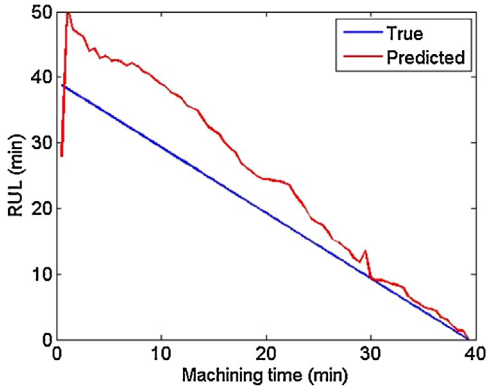

Fig. 9. Prediction of the RUL at 2700 rpm with the training at 1800 rpm.

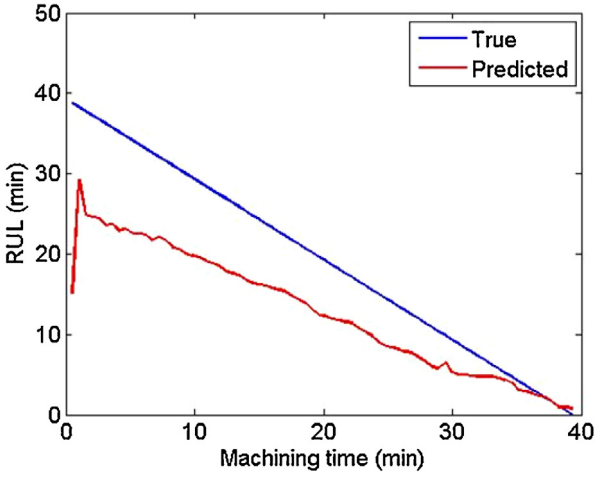

Fig. 10. Prediction of the RUL at 2700 rpm with the training at 4100 rpm.

5. Uncertainty generation

The growth of P_{rms} is not perfect and there is uncertainty as a function of machining time and tool life, resulting from different cutting conditions, measurement errors and error in true machining parameters values given by the machine during cutting. According to the trend of the P_{rms}, the function was approximated using a second order polynomial. The uncertainty was assumed to happen in the following three stages:

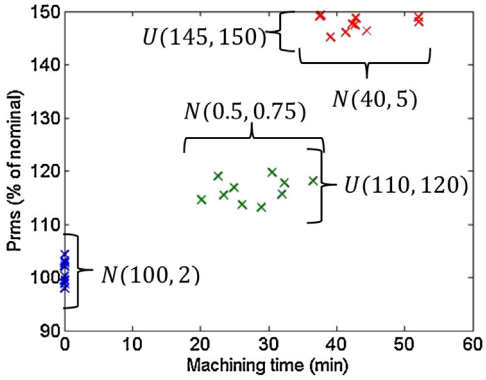

Fig. 11. Uncertainty distribution with a sample of 10 points for three cutting stages.

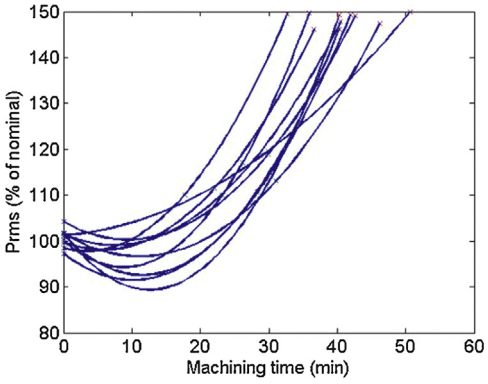

Fig. 12. Sample curves of 10 points for three cutting stages.

- the beginning of cutting (*i.e.*, no wear yet)
- the end of the initial wear stage
- the end of tool life

At the beginning of cutting, a normal distribution was assumed with a standard deviation of 2% at the value of 100% of nominal power, $N(100, 2)$, where N is the normal distribution function and the numbers in parenthesis denote the mean and the standard deviations, respectively. The time for the initial P_{rms} was 0 min. The P_{rms} at the end of the initial wear stage was assumed to be between 110% and 120% with equal probability, $U(110, 120)$, where U is the uniform distribution function and the numbers in parenthesis denote the minimum and the maximum values, respectively. The time at the end of the initial wear stage was assumed to follow a normal distribution between 0.5 and 0.75 of the tool life, $N(0.5, 0.75)$. Third, the P_{rms} at the end of the tool life was considered as a uniform distribution between 145% and 150% of the nominal value, $U(145, 150)$. Finally, the tool life was assumed to follow a normal distribution with mean equal to 40 min and standard deviation equal to 5 min, $N(40, 5)$. This distribution was based on the training data shown in Fig. 7. Fig. 11 illustrates the distribution described above with a sample of 10 points for each zone.

Sample path curves for the power were built as the consecution of a point in each zone generated by the above distributions. To generate the sample curve a second order least squares line was fit to the three values of P_{rms} and time. This procedure was repeated for 1000 curves to integrate a maximum of uncertainty in the sample curves. Fig. 12 shows this procedure for a sample of 10 curves.

6. Predicted RUL results and discussions

This section presents results of the RUL prediction for tests at 1380 and 2700 rpm. After the uncertainty generation, a sample lot

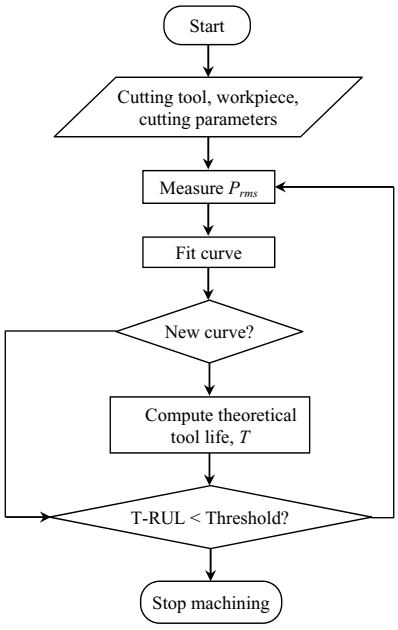

Fig. 13. Algorithm for predicting remaining useful life of tool inserts during machining.

of 1000 curves were used to predict the RUL. Each curve was compared with the power signal to select the best fitted curve. Once the best fitted curve was extracted, it was used to predict the RUL. Fig. 13 illustrates the algorithm describing this procedure.

To verify the accuracy of the results, the mean absolute error and the maximum error between the true RUL and the predicted RUL were compared. Fig. 14(a) and (b) shows the predicted and true tool life results with 5 neurons for 2700 rpm and 1380 rpm, respectively. These results show the accuracy of this method incorporating uncertainty. As observed in these figures, the noise presents with the previous training in Figs. 9 and 10 was eliminated. Only a little uncertainty remains with a mean absolute error of about 1–1.5 min. To avoid issues during the tool condition monitoring or the RUL prediction, this error has to be taken into account during the choice of the threshold.

Based on the study presented in Section 3.2, a performance comparison is performed between the best three training functions including *Trainbfg*, *Trainlm*, and *Traincgb*. The method used was the one presented in Fig. 13 and was applied to the 2700 rpm data. Once the best fitted curve selected, the RUL prediction was computed 50 times with the three training functions and three sizes of hidden layer [5810]. Once results obtained, the mean absolute error and maximum error values are computed to do a comparison in term of accuracy. Finally, the computation time was extracted to compare the speed. Results are presented in Table 1 for three different fitted curves. To confirm the trend of the results, a last prediction of the RUL was performed at 1380 rpm (see Table 2).

From the predicted RUL results, the *trainlm* function seems to be the most stable function, but it is found to be the most time consumer. To reduce the computation time of the *trainlm* function and keep a good accuracy, an augmentation of the hidden layer size seems to be suitable. To illustrate these tables, Figs. 15–17, respectively, show the prediction at 2700 rpm with the same fitted

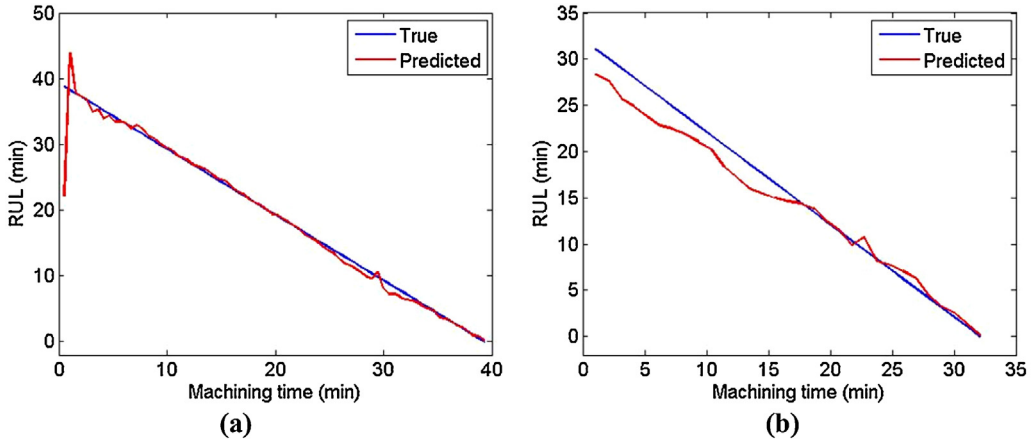

Fig. 14. Prediction of the RUL at (a) 2700 rpm, and (b) 1380 rpm using *Trainlm* training function with 5 neurons.

Table 1
Results comparison between training functions at 2700 rpm for three different fitted curves.

	Training function	*Trainbfg*			*Trainlm*			*Traincgb*		
	Number of neurons	5	8	10	5	8	10	5	8	10
Curve fitting #1	Mean absolute error (min)	0.98	1.10	1.03	0.90	0.90	0.91	1.29	1.09	1.02
	Maximum error (min)	16.98	16.89	16.69	16.85	16.85	16.84	17.80	16.50	16.96
	End tool life error (min)	0.86	0.88	0.92	0.69	0.68	0.72	0.81	0.65	0.64
	Computation time (s)	0.12	0.09	0.21	0.51	0.51	0.42	0.08	0.07	0.09
Curve fitting #2	Mean absolute error (min)	1.08	1.14	1.20	1.02	1.03	1.03	1.34	1.15	1.04
	Maximum error (min)	17.15	16.96	17.19	17.04	17.02	17.02	17.76	16.61	17.24
	End tool life error (min)	0.79	0.95	0.94	0.71	0.73	0.78	0.94	0.71	0.91
	Computation time (s)	0.12	0.10	0.45	0.50	0.48	0.37	0.07	0.08	0.08
Curve fitting #3	Mean absolute error (min)	1.07	1.06	1.23	1.04	1.04	1.04	1.33	1.06	1.07
	Maximum error (min)	15.69	15.68	15.88	15.59	15.59	15.59	16.41	15.16	15.76
	End tool life error (min)	0.92	0.97	0.56	0.87	0.88	0.89	1.32	1.14	1.16
	Computation time (s)	0.13	0.10	0.41	0.50	0.47	0.38	0.08	0.08	0.09

Table 2
Results comparison between training functions at 1380 rpm.

Train function	*Trainbfg*			*Trainlm*			*Traincgb*		
Number of neurons	5	8	10	5	8	10	5	8	10
Mean absolute error (min)	1.59	1.86	1.56	1.21	1.23	1.34	1.39	1.34	1.52
Maximum error (min)	4.93	6.61	5.38	3.72	3.88	5.10	4.78	5.10	5.60
End tool life error (min)	1.32	0.75	0.80	1.17	1.30	1.02	1.56	1.24	1.15
Computation time (s)	0.08	0.07	0.07	0.44	0.24	0.13	0.07	0.07	0.07

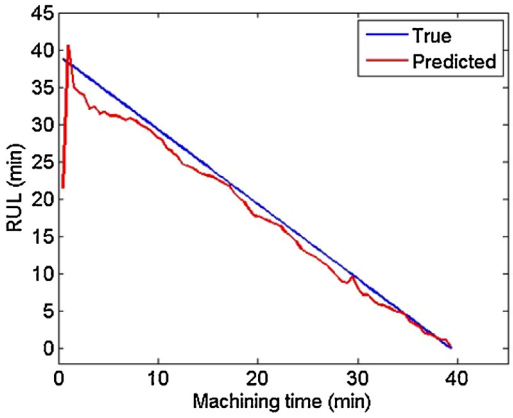

Fig. 15. Prediction of the RUL at 2700 rpm with *trainbfg*.

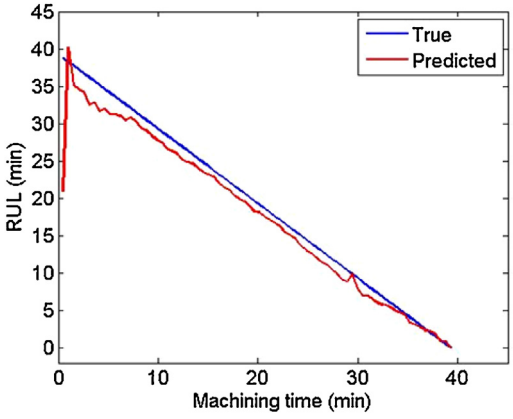

Fig. 16. Prediction of the RUL at 2700 rpm with *trainlm*.

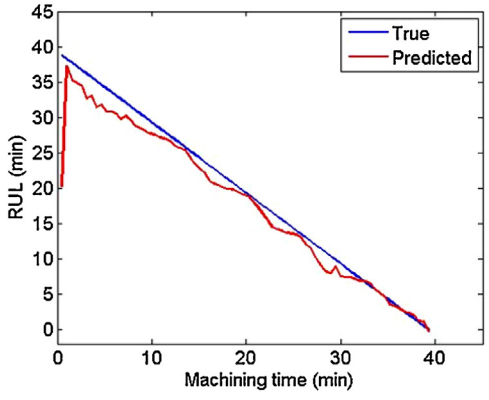

Fig. 17. Prediction of the RUL at 2700 rpm with *traincgb*.

curve and a neural network using the three training functions with a hidden layer of 8 neurons.

According to the results shown above, *trainlm* function is the most accurate with the minimum noise production. The computation time is higher than other training functions, but is still low (0.5 s maximum). This insignificant delay should not bring issues during (near) real time tool condition monitoring and RUL prediction.

7. Conclusions

Using the spindle power sensor data and a curve fitting method of Artificial Neural Network (ANN), remaining useful life (RUL) of the tools for milling operation was predicted. The root mean square power (P_{rms}) value in the time domain was found to be sensitive to tool wear. Random sample P_{rms} growth curves were generated representing the true curve. Each sample curve was attached to a true RUL curve. By combining these sample curves with the curve fitting method, the prediction of the RUL has been done for a set of two experiments, with a mean error of 1 min between the predicted RUL and the true RUL. The method offers many advantages as pointed out below:

- First, it incorporates uncertainties in tool life and the normalized power sensor value at the end of tool life. Therefore, the method is robust to outlier points and more reliable than methods that rely on a deterministic threshold sensor value.
- Second, this method can be applied to different spindle speeds since it includes the uncertainty coming from it, and works on a percentage of nominal power value.
- Third, the choice of working on RUL provides a better comprehension of the threshold and the value. While a power threshold informs the user about a power limit, a time threshold informs with a value that directly reflects the state of the tool and the production state (*e.g.*, how much time remains before a change of tool).
- Fourth, the method is computationally inexpensive and can be incorporated for real-time RUL predictions. To illustrate, using Matlab™ on an Intel i7 processor, the computation takes 0.5 s. Furthermore, the computation time can be decreased by using a suitable training function, but with a slightly compromised accuracy.
- Finally, the predicted RUL values/curves are found to be in a good agreement with the true RUL values/curves, with a mean error of about 1 min. This error can be taken into account in the threshold value to avoid any tool breakage issues.

Future work will focus on validation of this approach at different machining conditions by varying feeds, and axial and radial depths of cut.

References

[1] Kegg RL. One-line machine and process diagnostics. CIRP Ann 1984;33:469–73.
[2] Kurada S, Bradley C. A review of machine vision sensors for tool condition monitoring. Comput Ind 1997;34(1):55–72.

[3] Marksberry PW, Jawahir IS. A comprehensive tool-wear/tool-life performance model in the evaluation of NDM (near dry machining) for sustainable manufacturing. Int J Mach Tools Manuf 2008;48(7–8):878–86, http://dx.doi.org/10.1016/j.ijmachtools.2007.11.006.
[4] Abellan-Nebot JV, Romero Subirón F. A review of machining monitoring systems based on artificial intelligence process models. Int J Adv Manuf Technol 2009;47(1–4):237–57, http://dx.doi.org/10.1007/s00170-009-2191-8.
[5] Cho D-W, Lee SJ, Chu CN. The state of machining process monitoring research in Korea. Int J Mach Tools Manuf 1999;39(11):1697–715, http://dx.doi.org/10.1016/S0890-6955(99)00026-7.
[6] Kurada S, Bradley C. A review of machine vision sensors for tool condition monitoring. Comput Ind 1997;34(1):55–72, http://dx.doi.org/10.1016/S0166-3615(96)00075-9.
[7] Rehorn AG, Jiang J, Orban PE. State-of-the-art methods and results in tool condition monitoring: a review. Int J Adv Manuf Technol 2004;26(7–8):693–710, http://dx.doi.org/10.1007/s00170-004-2038-2.
[8] Giardini C, Ceretti E, Maccarini G. A neural network architecture for tool wear detection through digital camera observations. In: Kuljanic E, editor. Advanced manufacturing systems and technology SE – 14, vol. 372. Vienna: Springer; 1996. p. 137–44, http://dx.doi.org/10.1007/978-3-7091-2678-3_14.
[9] Shahabi HH, Ratnam MM. In-cycle monitoring of tool nose wear and surface roughness of turned parts using machine vision. Int J Adv Manuf Technol 2008;40(11–12):1148–57, http://dx.doi.org/10.1007/s00170-008-1430-8.
[10] Xiong G, Liu J, Avila A. Cutting tool wear measurement by using active contour model based image processing. In: 2011 IEEE international conference on mechatronics and automation, 2. 2011. p. 670–5, http://dx.doi.org/10.1109/ICMA.2011.5985741.
[11] Cuppini D, Errico GD. Tool image processing with applications to unmanned metal-cutting. A computer vision system for wear sensing and failure detection; 1986. p. 416.
[12] Yeo SH, Khoo LP, Neo SS. Tool condition monitoring using reflectance of chip surface and neural network. J Intell Manuf 2000:507–14.
[13] Obikawa T, Shinozuka J. Monitoring of flank wear of coated tools in high speed machining with a neural network ART2. Int J Mach Tools Manuf 2004;44(12–13):1311–8, http://dx.doi.org/10.1016/j.ijmachtools.2004.04.021.
[14] Dong J, Subrahmanyam KVR, Wong YS, Hong GS, Mohanty AR. Bayesian-inference-based neural networks for tool wear estimation. Int J Adv Manuf Technol 2005;30(9–10):797–807, http://dx.doi.org/10.1007/s00170-005-0124-8.
[15] Huang BP, Chen JC. Neural network-based tool breakage monitoring system for end milling operations. J Ind Technol 2000;16(2):1–7.
[16] Lin SC, Lin RJ. Tool wear monitoring in face milling using force signals. Wear 1996;198:136–42.
[17] Jemielniak K, Kosmol J, n.d. Tool and process monitoring – state of art and future prospects.
[18] Yesilyurt I, Ozturk H. Tool condition monitoring in milling using vibration analysis. Int J Prod Res 2007;45(4):1013–28, http://dx.doi.org/10.1080/00207540600677781.
[19] Fu P, Li WL, Zhu LQ. Cutting tool wear monitoring based on wavelet denoising and fractal theory. Appl Mech Mater 2011;48–49:349–52, http://dx.doi.org/10.4028/www.scientific.net/AMM.48-49.349.
[20] Zhang JZ, Chen JC. Tool condition monitoring in an end-milling operation based on the vibration signal collected through a microcontroller-based data acquisition system. Int J Adv Manuf Technol 2007;39(1–2):118–28, http://dx.doi.org/10.1007/s00170-007-1186-6.
[21] Bahr B, Motavalli S, Arfi T. Sensor fusion for monitoring machine tool conditions. Int J Comput Integr Manuf 1997;10(5):314–23, http://dx.doi.org/10.1080/095119297131066.
[22] Li XL, Yuan ZJ. Tool wear monitoring with wavelet packet transform fuzzy clustering method. Wear 1998;219:145–54.
[23] Li XL, Dong S, Yuan ZJ. Discrete wavelet transform for tool breakage monitoring. Int J Mach Tools Manuf 1999;39:1935–44.
[24] Lee B, Tarng YS. Application of the discrete wavelet transform to the monitoring of tool failure in end milling using the spindle motor current. Int J Adv Manuf Technol 1999;15(4):238–43.
[25] Du R, Liu Y, Xu Y. Tool condition monitoring using transition fuzzy probability. In: 3rd international conference on metal cutting and high speed machining. 2001. p. 375–93.
[26] Shao H, Wang HL, Zhao XM. A cutting power model for tool wear monitoring in milling. Int J Mach Tools Manuf 2004;44(14):1503–9, http://dx.doi.org/10.1016/j.ijmachtools.2004.05.003.
[27] Ghosh N, Ravi YB, Patra A, Mukhopadhyay S, Paul S, Mohanty AR, et al. Estimation of tool wear during CNC milling using neural network-based sensor fusion. Mech Syst Signal Process 2007;21:466–79.
[28] Rooij AJF, Van Johnson RP, Jain LC. Neural network training using genetic algorithms. River Edge, NJ, USA: World Scientific Publishing Co., Inc.; 1996.
[29] Theodosis P. Neural networks: interpolation and curve fitting; 2007. p. 1–12.
[30] Hill T, Marquez L, O'Connor M, Remus W. Artificial neural network models for forecasting and decision making. Int J Forecast 1994;10(1):5–15.
[31] Stergiou C. Why use a neural network? http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol1/cs11/article1.html [accessed 15.12.15].
[32] Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49(11):1225–31.
[33] Stathakis D. How many hidden layers and nodes? Int J Remote Sens 2009;30(8):2133–47.
[34] Karandikar J, McLeay T, Turner S, Schmitz T. Remaining useful tool life predictions using Bayesian inference. Volume 2: Systems; micro and nano technologies; sustainable manufacturing. ASME; 2013, June 10, http://dx.doi.org/10.1115/MSEC2013-1152.