

Tool wear monitoring using naïve Bayes classifiers

Jaydeep Karandikar · Tom McLeay · Sam Turner ·
Tony Schmitz

Received: 5 June 2014 / Accepted: 27 October 2014 / Published online: 14 November 2014
© Springer-Verlag London 2014

Abstract A naïve Bayes classifier method for tool condition monitoring is described. End-milling tests were performed at different spindle speeds and the cutting force was measured using a table-mounted dynamometer. The effect of tool wear on force features in the time and frequency domains was evaluated and used for training the classifier. The amount of tool wear was predicted using the naïve Bayes classifier method. Two cases are presented. First, the tool wear is divided into discrete states based on the amount of flank wear and the probability of the tool wear being in any state is updated using force data. Second, a continuous case is considered and the probability density function of the tool flank wear width is updated. The results are discussed.

Keywords Tool condition monitoring · Naïve Bayes classifier · Flank wear · Uncertainty · Cutting force

1 Introduction

Tool wear is an important limitation to machining productivity. Tool wear or breakages can result in unscheduled machine downtime in an industrial production environment, poor quality, or scrapping of the part resulting in a significant economic loss. The loss in productivity can be minimized by changing

tools frequently; however, this results in an increase in tooling costs. It has been reported that only 50–80 % of the expected tool life is typically used [1]. Estimates state that the amount of downtime due to cutter breakage on an average machine tool is on the order of 7–20 % [2, 3]. Therefore, an intelligent and robust tool condition monitoring system is desirable; an accurate tool condition monitoring system may result in increase in spindle speed by 10–50 %, a reduction in downtime resulting in a 10–40 % reduction in machining cost [4].

A tool condition monitoring system is generally composed of three parts: 1) identifying and extracting relevant features correlated to tool wear; 2) training the system using tool wear experiments; and 3) developing an intelligent inference technique for predicting tool wear [5, 6]. A variety of methods for inferring the tool condition have been discussed in the literature, including neural networks [7–11], fuzzy logic [12–14], neuro-fuzzy networks [15–18], Bayesian networks [19], and Hidden Markov models (HMM) [20–23]. A wide range of measurement and sensory feature extraction techniques have also been demonstrated, where measurements of acoustic signals, vibration, temperature, force, and power are analyzed using time and frequency domain signal processing tools. However, there are several challenges to the implementation of existing methods in the literature. First, tool condition monitoring systems which rely on mathematical models generally require a significant amount of empirical data and, therefore, are challenging to apply in industrial applications [5]. Second, another important limitation to tool condition monitoring is the stochastic nature of the sensor signal due to large-scale variation and non-homogeneities in the work-piece [24]. Therefore, a fusion of multi-sensor data is necessary for a robust system which incorporates the underlying uncertainties in the sensor signals. Third, it is desired that the monitoring system be computationally efficient to be useful in real time. To address these challenges, the authors propose a naïve Bayes' classifier for tool condition monitoring.

J. Karandikar (✉)

The George W. Woodruff School of Mechanical Engineering,
Georgia Institute of Technology, Atlanta, GA, USA
e-mail: jaydeep.karandikar@me.gatech.edu

T. McLeay · S. Turner

Advanced Manufacturing Research Centre with Boeing, University
of Sheffield, Rotherham, UK

T. Schmitz

Mechanical Engineering and Engineering Science, University of
North Carolina at Charlotte, Charlotte, NC, USA

A naïve Bayes classifier is an efficient and effective algorithm for machine learning and data mining [25–27]. The naïve Bayes classifier has several advantages over alternative classification schemes such as neural networks or fuzzy logic. First, naïve Bayes classifier is computationally efficient because of the independence assumption; the need to calculate covariances between different sensor features is eliminated, and each feature can be estimated using a one-dimensional distribution. The theoretical background and the conditions for the efficacy of naïve Bayes have been presented in [21]. Second, the classifier requires a small amount of training data and incorporates the inherent uncertainty in the sensor signals. Third, the naïve Bayes is an attractive solution for multi-sensor fusion in tool condition monitoring. In spite of the advantages, the naïve Bayes approach has not yet been considered in the literature for tool condition monitoring. The main contribution of the paper is to demonstrate the methodology and the efficacy of the naïve Bayes for tool condition monitoring using discrete and continuous data.

In this paper, the application of the naïve Bayes classifier to tool condition monitoring using end-milling force data measured by a table-mounted cutting force dynamometer is demonstrated. Both discrete and continuous cases are presented. Two force data features are considered, the time domain mean cutting force and the sum of the frequency domain amplitudes at the tooth passing frequency and its harmonics. However, the method described here can also be applied to other sensor signals, such as power, acoustic emission, and vibration. The remainder of the paper is organized as follows. Section 2 introduces the naïve Bayes classifier and the corresponding assumptions. Section 3 develops the classifier for a tool condition monitoring system. Section 4 describes the training experiments and tool wear results. Sections 5 and 6 describe the discrete and continuous cases, respectively. A discussion on the results and the method is presented in Section 7 followed by conclusions in Section 8.

2 Naïve Bayes classifiers

Bayes' rule provides a framework for incorporating judgment (prior beliefs) with observational data. Let the prior distribution about an uncertain event, A , be $P(A)$, the likelihood of obtaining an experimental result B given that event A occurred be $P(B|A)$, and the probability of observing experimental result B (without knowing A has occurred) be $P(B)$. Bayes' rule is used to determine the posterior belief about event A after observing the experiment results, $P(A|B)$, as shown in Eq. 1. Using Bayes' rule, information gained through

observations can be combined with the prior prediction about the event to obtain a posterior distribution, which represents the updated belief.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

For n experimental results, B_1, B_2, \dots, B_n , Bayes' rule from Eq. (1) is expressed as shown in Eq. 2.

$$P(A|B_1, B_2, \dots, B_n) = \frac{P(B_1, B_2, \dots, B_n|A)P(A)}{P(B_1, B_2, \dots, B_n)} \quad (2)$$

The numerator of Eq. 2 is a joint probability of $P(A, B_1, B_2, \dots, B_n)$ which can be written from the rules of conditional probability as shown in Eq. 3.

$$\begin{aligned} P(A, B_1, B_2, \dots, B_n) &= P(A)P(B_1, B_2, \dots, B_n|A) \\ &= P(A)P(B_1|A)P(B_2, \dots, B_n|A, B_1) \\ &= P(A)P(B_1|A)P(B_2|A, B_1)P(B_3, \dots, B_n|A, B_1, B_2) \\ &= P(A)P(B_1|A)P(B_2|A, B_1)P(B_3|A, B_1, B_2) \\ &\quad \dots P(B_n|A, B_1, B_2, \dots, B_{n-1}) \end{aligned} \quad (3)$$

Note that for a large value of n , the joint probability $P(A, B_1, B_2, \dots, B_n)$ is difficult to compute. The naïve assumption states conditional independence, i.e., each experimental result is independent of the other; see Eq. 4.

$$P(B_i|A, B_j) = P(B_i|A) \quad 1 \leq i, j \leq n, i \neq j \quad (4)$$

The assumption of conditional independence simplifies Eq. 3 to be

$$\begin{aligned} P(A, B_1, B_2, \dots, B_n) \\ &= P(A)P(B_1|A)P(B_2|A)P(B_3|A) \dots P(B_n|A) \end{aligned} \quad (5)$$

Using the (naïve) conditional independence assumption, Bayes' rule shown in Eq. 2 can be written as

$$P(A|B_1, B_2, \dots, B_n) \propto P(A)P(B_1|A)P(B_2|A)P(B_3|A) \dots P(B_n|A). \quad (6)$$

Note that the denominator in Eq. 2 is a normalizing constant and is not shown in Eq. 6.

3 Naïve Bayes classifier for tool condition monitoring

In tool condition monitoring, the uncertain variable of interest is the state of tool wear, TW . The experimental observations are the sensor features, F . For tool wear monitoring, Bayes' equation with the naïve assumption is given as

$$P(TW|F_1, F_2, \dots, F_n) \propto P(TW)P(F_1|TW)P(F_2|TW)P(F_3|TW) \dots P(F_n|TW) \propto P(TW) \prod_1 P(F_i|TW), \quad (7)$$

where $P(TW|F_1, F_2, \dots, F_n)$ is the posterior probability distribution of tool wear given sensor features, $P(TW)$ is the prior probability distribution of tool wear and $P(F_i|TW)$ is the likelihood of observing the sensor feature values given the tool wear state. The value of the likelihood terms is not known and can be determined from training experiments. Again, using Bayes' rule:

$$P(F_i|TW) \propto P(F_i)P(TW|F_i), \quad (8)$$

where $P(F_i|TW)$ is now the posterior probability distribution of sensor feature given tool wear (note that in Eq. 8, it is now the posterior that needs to be determined from the training experiment), $P(F_i)$ is the prior probability distribution of the sensor feature and $P(TW|F_i)$ is the likelihood of observing the tool wear state given a sensor feature. The value of $P(TW|F_i)$ is known in training experiments since the state of tool wear (e.g., flank wear width) is also measured along with the sensor feature values. Note that in Eqs. 7 and 8 the normalizing constant is not shown. The procedure for tool condition monitoring using naïve Bayes classifiers is now described. First, training experiments are performed in which the tool wear is measured at regular intervals along with the sensor features (in this case, force metrics in the time and frequency domain). Second, the experimental training data is used to determine the values of $P(F_i|TW)$ using Eq. 8 since the tool wear state is known in the training experiments. Third, the values of $P(F_i|TW)$ are used in Eq. 7 for inference of the tool wear state given sensor features.

The remainder of the paper is organized as follows. Section 4 describes the training experiments and results. Section 5 shows a discrete case, where tool wear is divided into discrete states based on the value of flank wear width (F_{WW}). The probability of the tool being in any state is updated using sensor values. Section 6 uses a continuous method where the probability density function (pdf) of the tool's F_{WW} is updated using the sensor results. Conclusions are presented in Section 8.

4 Training experiments

The experimental steps followed to collect the tool wear data for a 19-mm diameter inserted endmill (one square uncoated Kennametal 107888126 C9 JC carbide insert; zero rake and helix angles, 15-deg relief angle, 9.53-mm square, 3.18-mm thick) are described in this section. The workpiece material was 1018 steel. Three tests were completed at a spindle speed, Ω , of 2500 rpm with a 3-mm axial depth of cut and 4.75-mm radial depth of cut (25 % radial immersion). The feed per tooth value was 0.06 mm/tooth. The cutting forces were monitored using a table-mounted force dynamometer (Kistler 9257B). The insert wear profile was also recorded at regular intervals. To avoid removing the insert/tool from the spindle, a handheld digital microscope ($\times 60$ magnification) was fixtured inside the machine enclosure and was used to measure the rake and flank surfaces. The calibrated digital images were used to identify the F_{WW} (no crater wear was observed). Figure 1 shows the dynamometer/workpiece (left) and microscope for intermittent F_{WW} measurement (right). Figure 2 shows the increase in maximum F_{WW} with cutting time for the three tests performed at 2500 rpm. Additionally, three tests were carried out at additional spindle speeds of 5000 and 7500 rpm. The procedure and parameters (other than spindle speed) were the same as described for the 2500 rpm experiments. The results are shown in Fig. 3; clearly, tool wear depends on spindle speed (surface speed).



Fig. 1 Dynamometer/workpiece (top) and microscope for in process measurement of F_{WW} (bottom)

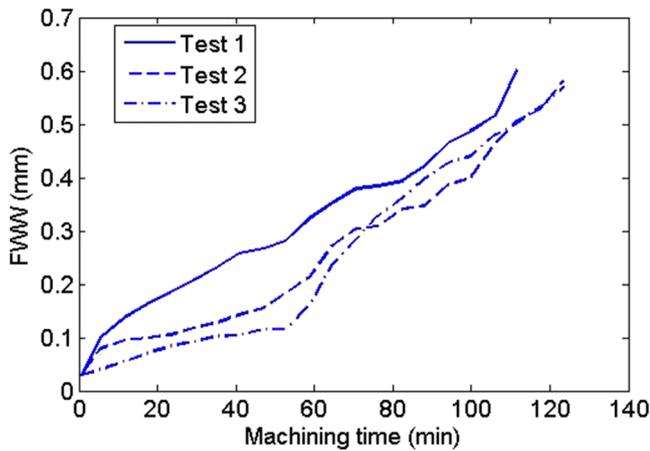


Fig. 2 Growth in FWW with machining time at 2500 rpm

The forces in the x and y directions (x is the feed direction, y is perpendicular to x within the plane of the cut) were measured at the same intervals as tool wear. Force metrics in the time and frequency domains were evaluated. The mean force, F_m , in the time domain and the sum of magnitudes of the tooth passing frequency and its first five harmonics, S_{tpf} , in the frequency domain were found to be most sensitive to tool wear. The selection of F_m and S_{tpf} was based on the R^2 value of the linear fit on the trends shown by the metrics as a function of FWW . To illustrate, the R^2 value for F_m and S_{tpf} was 0.55 and 0.72, respectively. For brevity, the graphs of additional metrics are not included in the manuscript.

Figure 4 shows the variation in mean forces as a function of FWW at different spindle speeds. Figure 5 shows the sum of the magnitudes of content at the tooth passing frequency and its harmonics in the frequency domain as a function of FWW . Although the mean forces and the sum of magnitudes of tooth

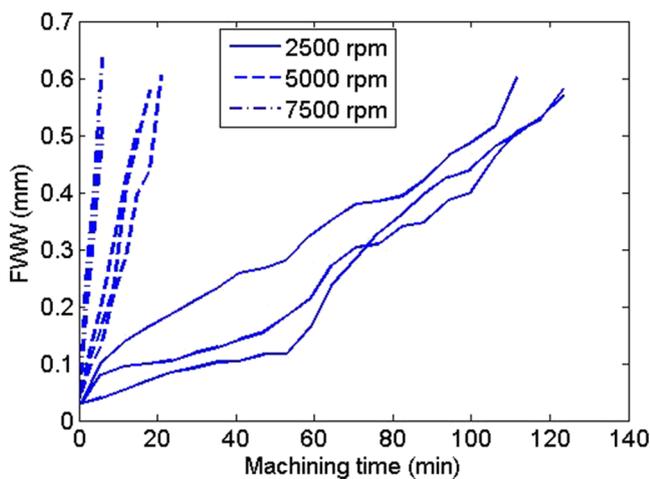


Fig. 3 Growth in FWW with machining time at 2500, 5000, and 7500 rpm

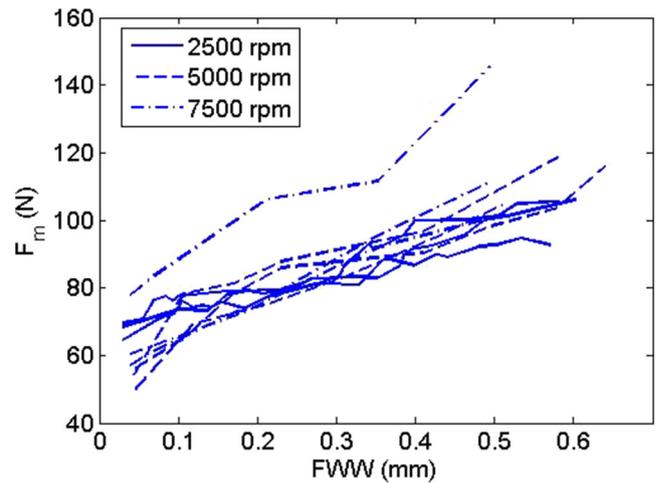


Fig. 4 Growth in mean force with FWW at spindle speeds, 2500, 5000, and 7500 rpm

passing frequency/harmonics show an increase with FWW , there is no clear dependence on spindle speed. Therefore, the dependence on spindle speed was not considered since no clear trend was observed. Figure 6 shows the values of F_m (left) and S_{tpf} (right), expressed as percent of the nominal value (measured during the first cut), denoted as $F_{m\%}$, and $S_{tpf\%}$, respectively. The primary reason for expressing the values as percent of the nominal was to reduce the uncertainty when the tool is new.

5 Tool condition monitoring using Bayes classifier (discrete case)

In this section, tool condition monitoring using a discrete Bayes classifier is described. Table 1 lists the discrete states of tool wear. The sensor features, F_m and S_{tpf} , were also

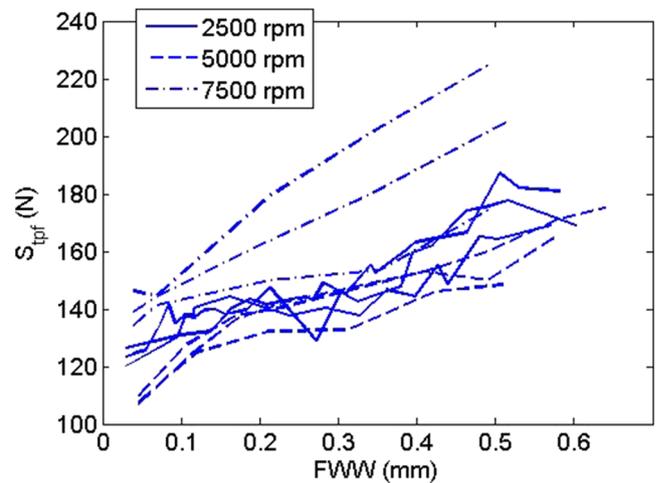


Fig. 5 Growth in sum of magnitude of tooth passing frequencies with FWW at spindle speeds, 2500, 5000, and 7500 rpm

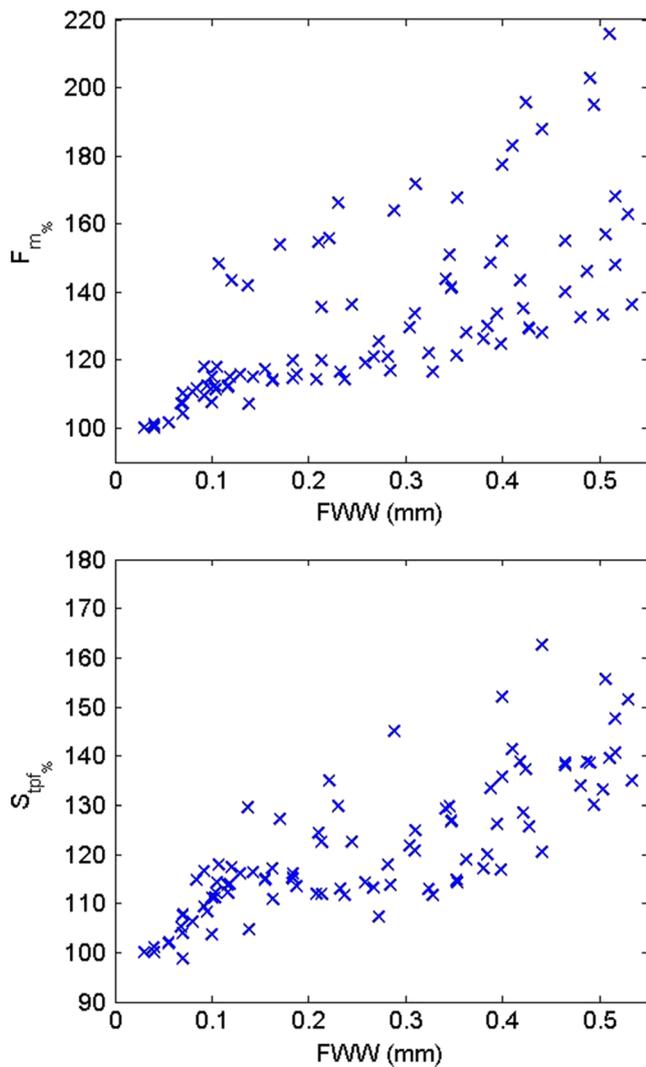


Fig 6 F_m and S_{spf} expressed as a percent of the nominal value, denoted by $F_{m\%}$ (top) and $S_{spf\%}$ (bottom), respectively

discretized and divided into five levels; see Table 2. The states for tool wear, F_m , and S_{spf} , are denoted as TW_S , F_{m_S} , and S_{spf_S} , respectively. The subscript S represents the discrete states of the variables. As seen from Table 2, the states for sensor features are divided into five levels. In general, the numbers of levels and states should be chosen such that representation of tool wear states in each sensor level and the corresponding uncertainties is incorporated.

Table 1 Discrete states for tool wear

State (TW_S)	Condition	FWW (mm)
1	New	≤ 0.1
2	Mildly worn	> 0.1 and ≤ 0.3
3	Worn	> 0.3

Table 2 Discrete states for F_m and S_{spf} , denoted by F_{m_S} and S_{spf_S} , respectively

State (F_{m_S}, S_{spf_S})	$F_{m\%}$	$S_{spf\%}$
1	< 112	< 112
2	> 112 and ≤ 124	> 112 and ≤ 124
3	> 124 and ≤ 136	> 124 and ≤ 136
4	> 136 and ≤ 148	> 136 and ≤ 148
5	> 148	> 148

5.1 Conditional probabilities using Bayes’ rule

The $F_{m\%}$ and $S_{spf\%}$ values from the tool wear experiments were used to determine the conditional probabilities between the two force metrics and tool wear. Note that there is uncertainty in the values of $F_{m\%}$ and $S_{spf\%}$ as a function of FWW ; see the spread of potential force metric values at a given FWW in Fig. 6. The conditional probabilities were determined using the learning experiments (described in Section 4.0) and Bayes’ rule (Eq. 8). For sensor features states F_{m_S} , and S_{spf_S} , and tool wear states, TW_S , Eq. 8 is written as

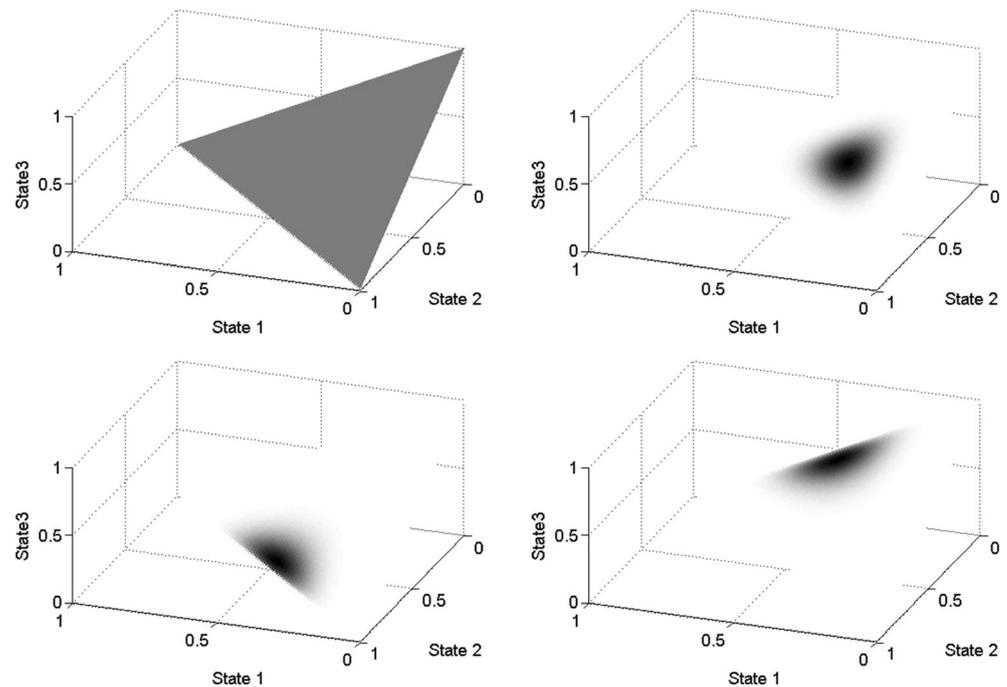
$$P(F_{m_S} | TW_S) \propto P(F_{m_S})P(TW_S | F_{m_S}) \tag{9}$$

$$P(S_{spf_S} | TW_S) \propto P(S_{spf_S})P(TW_S | S_{spf_S}) \tag{10}$$

As noted, the sensor features and tool wear were divided into discrete states. A Dirichlet-multinomial distribution conjugate pair was used for Bayes updating. For a Dirichlet-multinomial conjugate pair, the prior and the posterior are a distributed Dirichlet [28]. A Dirichlet distribution is characterized by the parameter, α , which defines the shape of the distribution. Note that a discrete Dirichlet distribution was used in this case. The Dirichlet distribution was a five-dimensional joint distribution because the force metrics were divided into five states. For a discrete Dirichlet, each state has a corresponding value of α . To illustrate, Fig. 7 shows the Dirichlet distribution for three states using different values of α . If the value of α is equal for all states, the distribution is uniform. However, note that both a Dirichlet{1,1,1} distribution and a Dirichlet{10,10,10} distribution represent beliefs where each state is equally likely. The former represents a belief where all states are assumed equally likely because no information is available otherwise, while the latter represents a belief where significant information is available which makes each state equally likely a priori. The probability of each state is the expected value of the state from the joint distribution which is obtained by normalizing the parameters to unity.

For the prior Dirichlet distribution of the force states, α has five values, one for each state. As noted, Dirichlet-

Fig. 7 Three-dimensional Dirichlet distribution for $\alpha\{1,1,1\}$ (top left), $\alpha\{10,10,10\}$ (top right), $\alpha\{10,10,2\}$ (bottom left), and $\alpha\{10,2,10\}$ (bottom right)



multinomial distributions form a conjugate pair. The posterior Dirichlet parameters are equal to the addition of the prior Dirichlet parameters for each state and the number of observations in that state. To illustrate, let the prior, $P(F_{m_s})$ be defined as Dirichlet $\{1, 1, 1, 1, 1\}$. The probability of each state is the expected value of the state from the joint five-dimensional distribution, which is obtained by normalizing the parameters to unity. Thus, the prior probability of each state is 0.2. Assume that five experimental measurements at a new tool condition ($TW_S=1$) were performed and the F_{m_s} measured values were in states $\{1, 2, 1, 2, \text{ and } 3\}$. Therefore, the number of observations in each state is $\{2, 2, 1, 0, \text{ and } 0\}$. The posterior distribution $P(F_{m_s}|TW_S)$ is given by Dirichlet $\{3, 3, 2, 1, 1\}$. The posterior probabilities of the five states are $\{0.3, 0.3, 0.2, 0.1, 0.1\}$.

The prior distributions of F_{m_s} and S_{tpf_s} were taken as a uniform Dirichlet $\{1, 1, 1, 1, 1\}$. As no prior information was available, the alpha values were assumed to be 1 for all states. The prior expected probability for each state is 0.2. The updating was performed using the experimental results and the Dirichlet-multinomial conjugate distribution. To illustrate, assume the number of observations for a new tool ($TW_S=1$) in each sensor state are $\{19, 3, 0, 0, \text{ and } 0\}$ and $\{20, 2, 0, 0, \text{ and } 0\}$ for F_{m_s} and S_{tpf_s} , respectively. The parameters of the posterior distribution of F_{m_s} and S_{tpf_s} given $TW_S=1$ are then Dirichlet $\{20, 4, 1, 1, 1\}$ and Dirichlet $\{21, 3, 1, 1, 1\}$, respectively. The posterior expected probabilities for each state of F_{m_s} and S_{tpf_s} , given $TW_S=1$, are shown in Fig. 8. The same procedure is followed to determine the expected probability for F_{m_s} and S_{tpf_s} states at all TW_S states. The conditional

probabilities for F_{m_s} and S_{tpf_s} are listed in Table 3 and Table 4, respectively. The posterior probabilities of the sensor states describe how likely it is that each sensor state will be observed given the tool wear state. To illustrate, given $TW_S=1$, it is very likely that $F_{m_s}=1$ and $S_{tpf_s}=1$, whereas it is less likely that $F_{m_s}=5$ and $S_{tpf_s}=5$. In general, the force metrics should be discretized to ensure some states are more likely than others for a given tool wear state.

5.2 Tool wear predictions using naïve Bayes' classifier

In Section 5.1, the likelihood of each force state given the tool wear state was calculated using Bayes' rule. The likelihood can then be used to determine the posterior probability of tool wear state given the force states; see Eq. 7. Additional tool wear testing was completed at 3750 rpm. Figure 9 shows the values of $F_{m\%}$ (left) and $S_{tpf\%}$ (right) as a function of FWW . The values are expressed as a percent of the nominal. The procedure for the tests was the same as described in Section 4. The FWW was measured at regular intervals along with the values of $F_{m\%}$ and $S_{tpf\%}$. The values of $F_{m\%}$ and $S_{tpf\%}$ were converted to states, defined in Table 2, and used to calculate the probability of tool wear state using Eq. 7, which was compared to the true measured FWW value.

The prior probabilities of TW_S were selected as $\{0.33, 0.33, 0.33\}$; it was assumed that the tool is equally likely to be in any state. The updating procedure required the following steps. The first measurement values were $F_{m_s}=1$ and $S_{tpf_s}=1$. The first column of Table 3 and Table 4 gives the likelihood

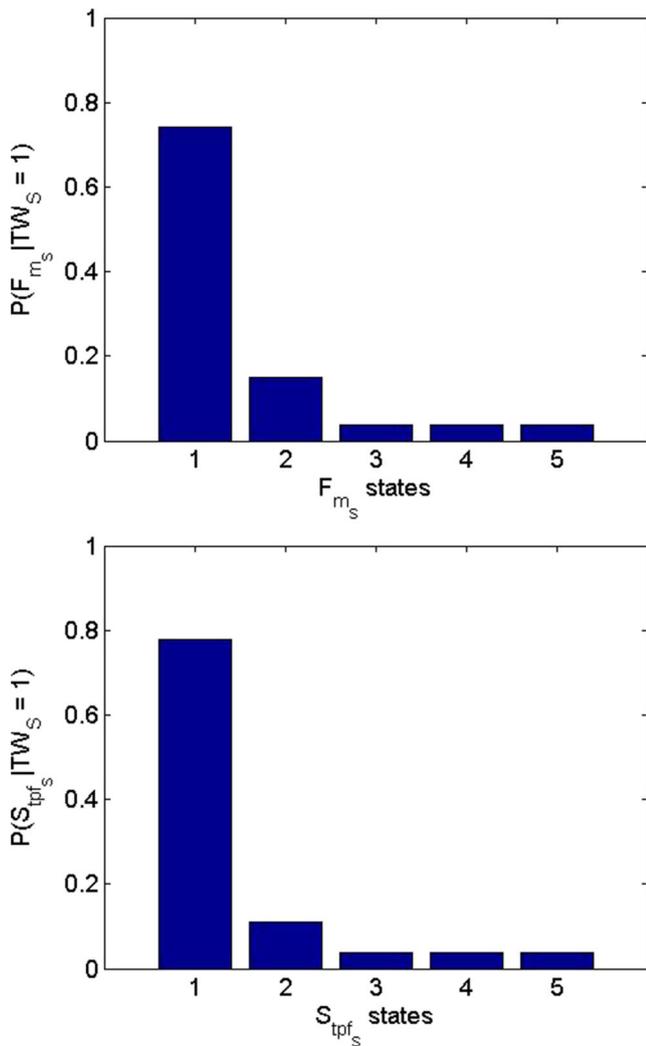


Fig. 8 Posterior probabilities for F_{m_s} (top) and S_{tpf_s} (bottom) given that the tool is new ($TW_S=1$)

of $F_{m_s} = 1$ and $S_{tpf_s} = 1$, given the tool wear states, respectively. The posterior probabilities of tool wear states were given by the product of the prior and the likelihoods of the sensor states normalized so that the sum is equal to unity. The calculations are summarized in Table 5. The posterior probabilities after the first update become the prior probabilities for the next and so on. Table 6 lists the measured sensor states and the posterior probabilities for each tool wear state. Note that the

Table 3 Posterior probabilities for F_{m_s} given tool wear states, $P(F_{m_s} | TW_S)$

	$F_{m_s} = 1$	$F_{m_s} = 2$	$F_{m_s} = 3$	$F_{m_s} = 4$	$F_{m_s} = 5$
$TW_S=1$	0.74	0.15	0.037	0.037	0.036
$TW_S=2$	0.08	0.56	0.08	0.10	0.18
$TW_S=3$	0.01	0.06	0.21	0.13	0.59

Table 4 Posterior probabilities for S_{tpf_s} given tool wear states $P(S_{tpf_s} | TW_S)$

	$S_{tpf_s} = 1$	$S_{tpf_s} = 2$	$S_{tpf_s} = 3$	$S_{tpf_s} = 4$	$S_{tpf_s} = 5$
$TW_S=1$	0.78	0.11	0.037	0.037	0.036
$TW_S=2$	0.21	0.56	0.15	0.05	0.03
$TW_S=3$	0.03	0.16	0.22	0.24	0.35

posterior probabilities are rounded to two significant digits. The maximum probability, or the most likely tool wear state, is shown in bold. Figure 9 displays the predicted posterior probabilities of the tool wear states as a function of true FWW . As shown in Fig. 9 and Table 6, good agreement between the predicted tool wear state and measured FWW was obtained. It can be seen from Table 6 that the tool wear state was correctly identified in 11 of the 13 tests. For tests 5 and 6, the tool wear state was underestimated.

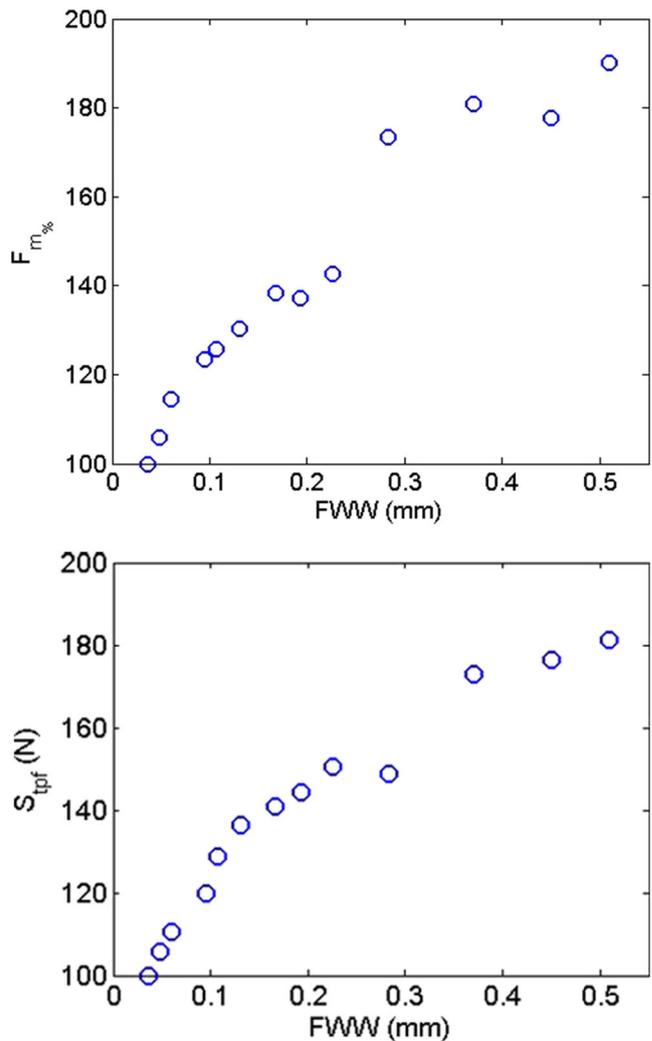


Fig. 9 $F_{m\%}$ (top) and $S_{tpf\%}$ (bottom) as a function of FWW at 3750 rpm

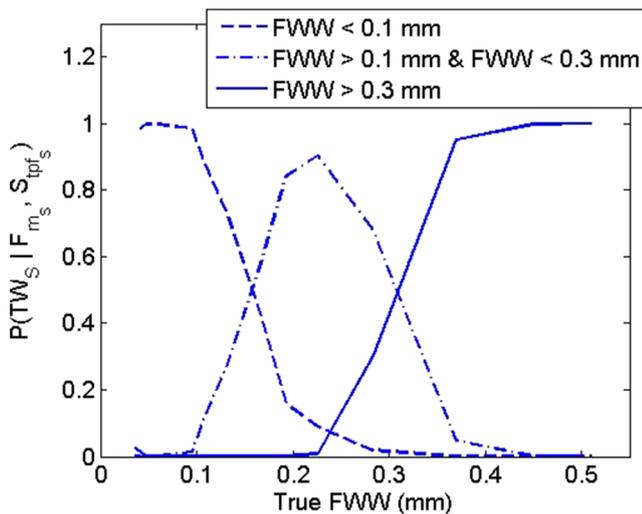


Fig. 10 Posterior probabilities of tool wear states as a function of true *FWW* at 3750 rpm

Additional tests were completed at 6250 rpm. Figure 11 shows the values of F_{m_s} (left) and S_{tpf_s} (right) as a function of *FWW* at 6250 rpm. The posterior probabilities were calculated using the same procedure described previously. Table 7 shows the measured states of F_{m_s} and S_{tpf_s} and posterior probabilities for tool wear states and the true *FWW* at 6250 rpm. Figure 12 shows the posterior probabilities of tool wear states as a function of true *FWW* at 6250 rpm.

6 Tool condition monitoring using Bayes classifier (continuous case)

In Section 5, tool wear predictions using Bayes classifiers were described for a discrete case. The tool condition was divided into three discrete states using the *FWW* and the probability of the tool being in each state was updated using force metrics. This section describes a continuous case where the probability of tool *FWW* is updated using the same force metrics. The $F_{m_{\%}}$ and $S_{tpf_{\%}}$ values from the tool wear experiments described in Section 4.0 were used to learn the network, or to determine the conditional probabilities between the force metrics and tool wear; see Eqs. 9 and 10.

Table 5 Posterior probabilities of TW_S given $F_{m_s} = 1$ and $S_{tpf_s} = 1$ (calculated using Eq. 7)

	Prior	Likelihood	Likelihood	Posterior	Posterior
	($F_{m_s} = 1$)	($S_{tpf_s} = 1$)	(non-normalized)	(normalized)	
$TW_S=1$	0.33	0.74	0.78	0.190	0.971
$TW_S=2$	0.33	0.08	0.21	0.005	0.028
$TW_S=3$	0.33	0.01	0.02	6.6×10^{-5}	3.37×10^{-4}

Table 6 Measured states of F_{m_s} and S_{tpf_s} and posterior probabilities for tool wear states and the true *FWW* at 3750 rpm

Test No.	F_{m_s}	S_{tpf_s}	$TW_S=1$	$TW_S=2$	$TW_S=3$	True <i>FWW</i>
1	1	1	<i>0.97</i>	0.03	0.00	0.04
2	1	1	<i>1.00</i>	0.00	0.00	0.05
3	2	1	<i>1.00</i>	0.00	0.00	0.06
4	2	2	<i>0.99</i>	0.01	0.00	0.09
5	3	3	<i>0.89</i>	0.11	0.00	0.11
6	3	4	<i>0.73</i>	0.27	0.00	0.13
7	4	4	<i>0.42</i>	<i>0.58</i>	0.00	0.17
8	4	4	<i>0.16</i>	<i>0.84</i>	0.00	0.19
9	4	5	<i>0.09</i>	<i>0.90</i>	0.01	0.22
10	5	5	<i>0.02</i>	<i>0.68</i>	0.30	0.28
11	5	5	0.00	0.05	<i>0.95</i>	0.37
12	5	5	0.00	0.00	<i>1.00</i>	0.45
13	5	5	0.00	0.00	<i>1.00</i>	0.51

The maximum probability values are showed in italics which denote the most likely tool wear state

6.1 Conditional probabilities using Bayes' rule

The conditional probabilities, $P(F_{m_{\%}}|TW)$ and $P(S_{tpf_{\%}}|TW)$, were determined using an ordinary Bayesian linear regression. For a continuous variable, *TW* is characterized by *FWW*. The observation errors were assumed to be independent with equal variance. The linear regression model was developed as

$$F_{m_{\%}} = \beta_1 + \beta_2 FWW + \epsilon \tag{11}$$

$$S_{tpf_{\%}} = \beta_3 + \beta_4 FWW + \epsilon \tag{12}$$

The posterior distribution for $F_{m_{\%}}$ and $S_{tpf_{\%}}$ are given by [29]:

$$F_{m_{\%}} | FWW \sim N(\beta_1 + \beta_2 FWW, \sigma_{F_{m_{\%}}}^2 I) \tag{13}$$

$$S_{tpf_{\%}} | FWW \sim N(\beta_3 + \beta_4 FWW, \sigma_{S_{tpf_{\%}}}^2 I) \tag{14}$$

where N denotes a normal distribution, $\beta_1, \beta_2, \beta_3,$ and β_4 are ordinary linear least squares values (intercept and slope), and $\sigma_{F_{m_{\%}}}^2$ and $\sigma_{S_{tpf_{\%}}}^2$ are the standard deviations in the measured $F_{m_{\%}}$ and $S_{tpf_{\%}}$ values, respectively. ϵ is defined as the errors in the regression which are random and normally distributed. Linear least squares fitting was applied to the measured $F_{m_{\%}}$ values and β_1 and β_2 were calculated as 100.1 % and 124.2 %/mm, respectively. Similarly, β_3 and β_4 were calculated by a linear least squares fit to the $S_{tpf_{\%}}$ values; they were 101 % and 75.8 %/mm. The values of $\sigma_{F_{m_{\%}}}^2$ and

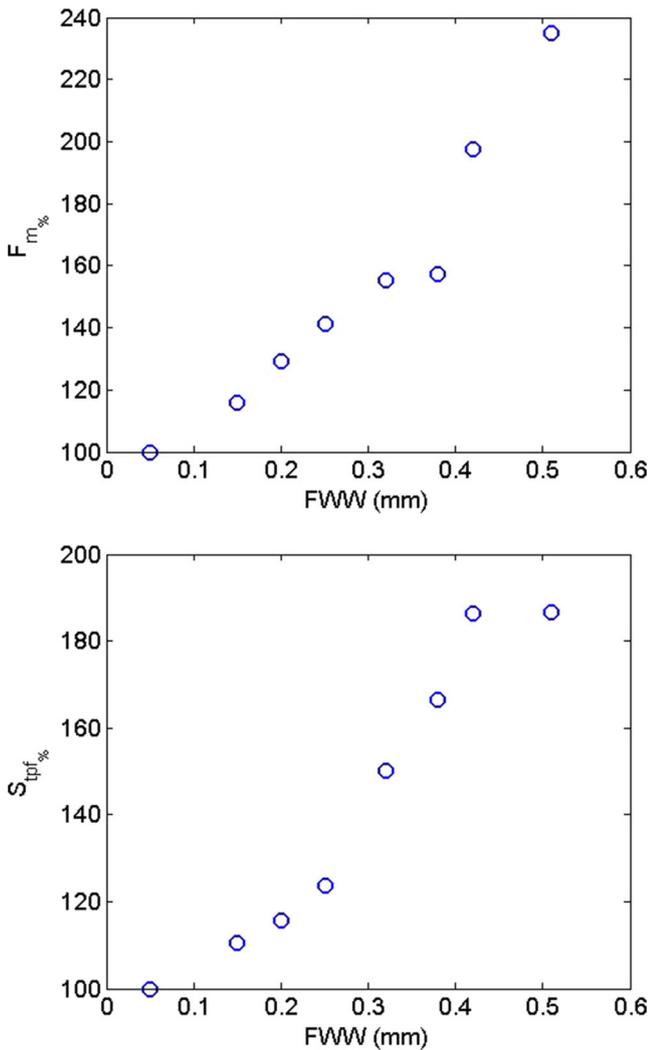


Fig. 11 $F_{m\%}$ (top) and $S_{tpf\%}$ (bottom) as a function of FWW at 6250 rpm

$\sigma^2_{S_{tpf\%}}$ were calculated from ordinary least squares analysis to be 20.15 and 8.36 %, respectively [29]. The distribution of $F_{m\%}$ and $S_{tpf\%}$ as a function of FWW was calculated using

Table 7 Measured states of F_{m_s} and S_{tpf_s} and posterior probabilities for tool wear states and the true FWW at 6250 rpm

Test no.	F_{m_s}	S_{tpf_s}	$TW_S=1$	$TW_S=2$	$TW_S=3$	True FWW
1	1	1	<i>0.97</i>	0.03	0.00	0.05
2	2	1	<i>0.97</i>	0.03	0.00	0.15
3	3	2	<i>0.78</i>	0.22	0.00	0.21
4	4	2	0.20	<i>0.80</i>	0.00	0.25
5	5	5	0.07	<i>0.93</i>	0.00	0.32
6	5	5	0.02	<i>0.79</i>	0.19	0.38
7	5	5	0.00	0.09	<i>0.91</i>	0.44
8	5	5	0.00	0.00	<i>1.0</i>	0.51

The maximum probability values are showed in italics which denote the most likely tool wear state

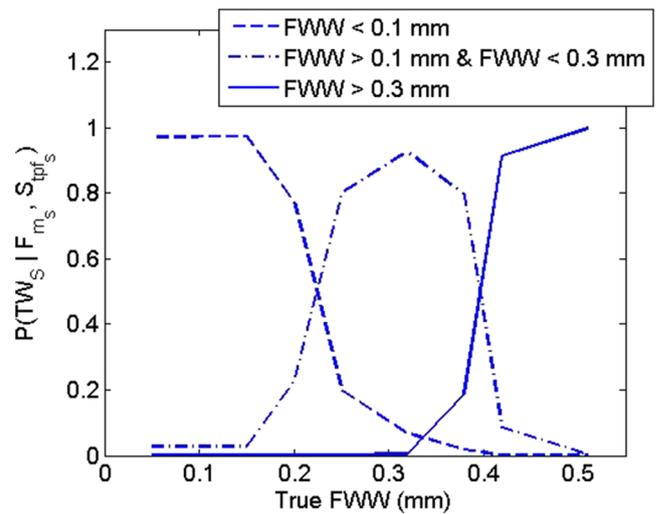


Fig. 12 Posterior probabilities of tool wear states as a function of true FWW at 6250 rpm

Eqs. 13 and 14. Figures 13 and 14 show the conditional probabilities for the sensor values, $P(F_{m\%}|FWW)$ and $P(S_{tpf\%}|FWW)$, respectively. The gray-scale color bar denotes the probability density. The measured values are denoted by “x”.

6.2 Tool wear predictions using naïve Bayes’ classifier

In Section 6.1, the likelihood of each sensor given FWW was calculated using ordinary Bayesian linear regression. The likelihood was used to determine the posterior probability of FWW given the sensor states; see Eq. 7. Tool wear tests were completed at 3750 rpm. Figure 9 shows the values of $F_{m\%}$ (left) and $S_{tpf\%}$ (right) as a function of FWW at 3750 rpm. The procedure for the tests was the same as described in Section 4. The FWW was measured at regular intervals along with the values of $F_{m\%}$ and $S_{tpf\%}$. The values of $F_{m\%}$ and $S_{tpf\%}$ were

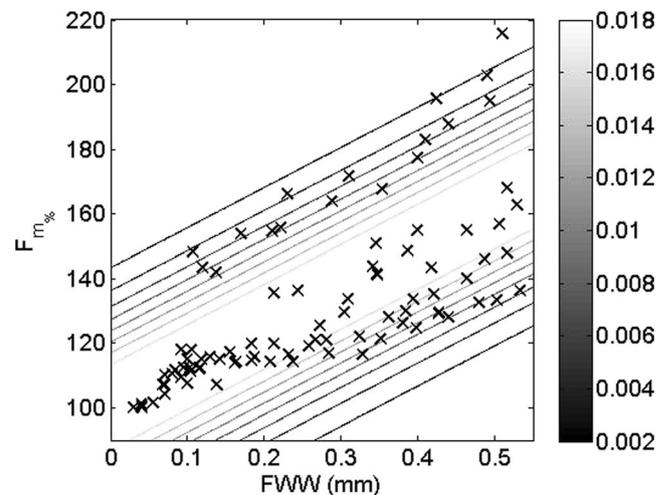


Fig. 13 Conditional probabilities, $P(F_{m\%}|FWW)$

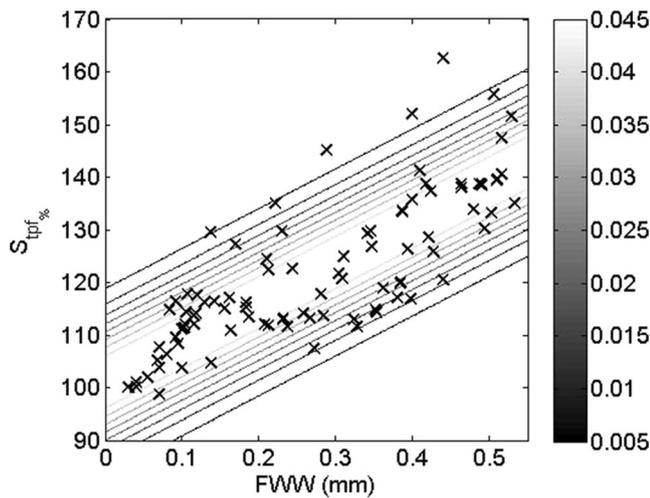


Fig. 14 Conditional probabilities, $P(S_{tpf\%} | FWW)$

used to calculate the posterior pdf of FWW using Eq. 7, which was compared to the true FWW value.

The prior probability of FWW was assumed to be uniform between 0 to 1 mm. The updating procedure is as follows. Consider a measurement of $F_{m\%} = 110.0\%$ and $S_{tpf\%} = 110\%$. The likelihood function gives how likely the measured sensor value is given FWW . The likelihood is given by the value of the probability density at the measured sensor value; see Figs. 13 and 14. Figure 15 shows the likelihood function for $F_{m\%} = 110.0\%$ (left) and $S_{tpf\%} = 110.0\%$ (right). Figure 15 shows that the $F_{m\%} = 110.0\%$ value is most likely for a FWW of 0.08 mm, whereas $S_{tpf\%} = 110\%$ is most likely for $FWW = 0.12$ mm. On the other hand, the measured values of $F_{m\%} = 110.0\%$ and $S_{tpf\%} = 110.0\%$ are unlikely for $FWW > 0.4$ mm. Figure 16 shows the likelihood function for $F_{m\%} = 150.0\%$ (left) and $S_{tpf\%} = 150.0\%$ (right). The most likely values of FWW for $F_{m\%} = 150.0\%$ and $S_{tpf\%} = 150.0\%$ are 0.4 and

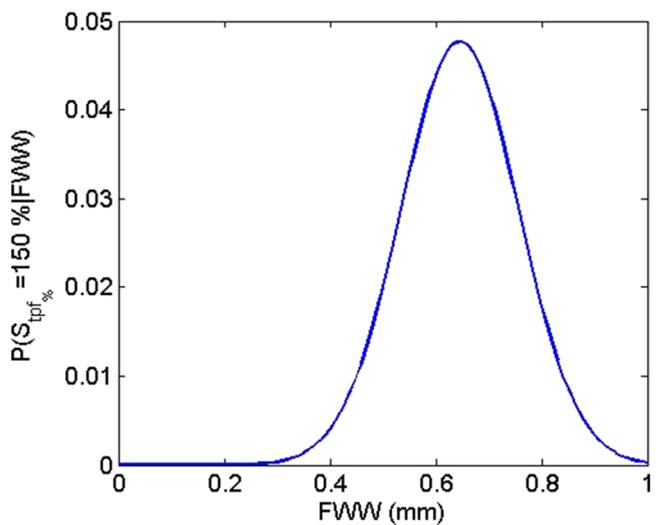
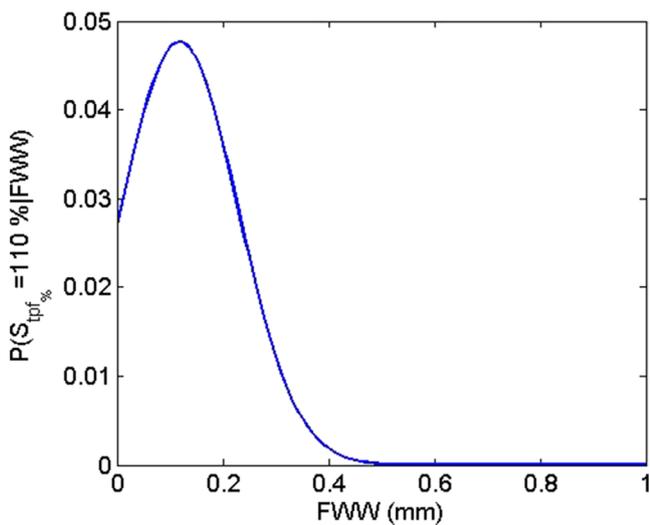
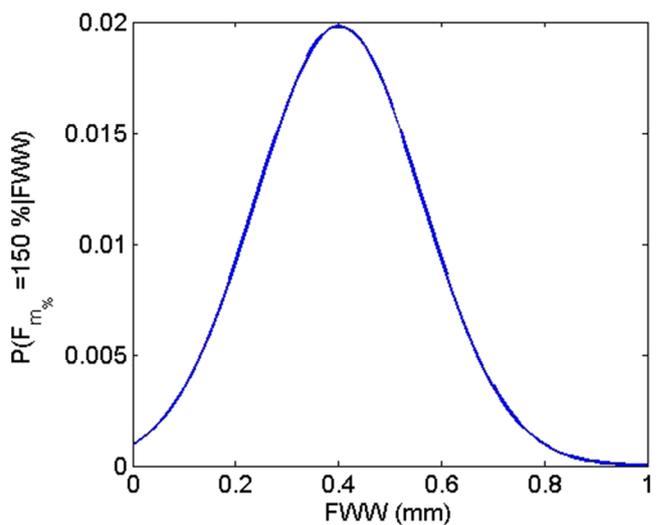
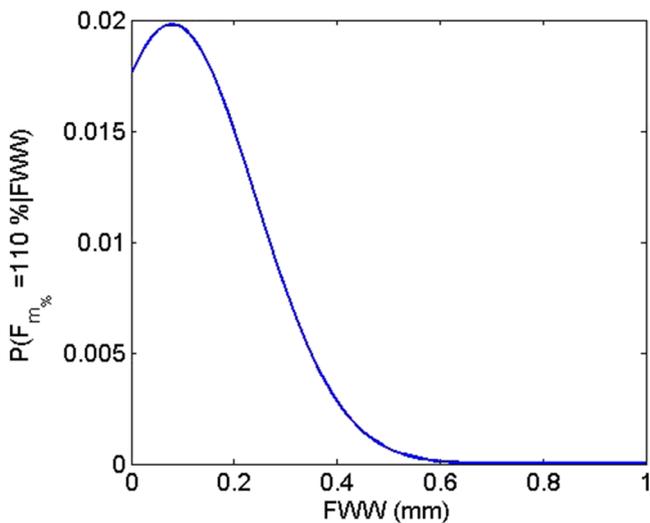


Fig. 15 Likelihood functions for $F_{m\%} = 110.0\%$ (top) and $S_{tpf\%} = 110\%$ (bottom)

Fig. 16 Likelihood functions for $F_{m\%} = 150.0\%$ (top) and $S_{tpf\%} = 150\%$ (bottom)

Table 8 Measured $F_{m\%}$ and $S_{tpf\%}$ values at 3750 rpm and the true FWW

Test no.	$F_{m\%}$	$S_{tpf\%}$	True FWW (mm)
1	100.0	100.0	0.04
2	106.0	105.8	0.05
3	114.5	110.4	0.06
4	123.5	120.0	0.09
5	125.8	128.8	0.11
6	130.4	136.5	0.13
7	138.4	140.8	0.17
8	137.2	144.4	0.19
9	142.7	150.6	0.22
10	173.3	148.8	0.28
11	181.0	173.0	0.37
12	177.8	176.4	0.45
13	190.0	181.2	0.51

0.65 mm, respectively. The posterior distribution is the product of the prior and the likelihood; see Eq. 7. The posterior distribution is normalized so that the area under the curve is equal to unity. The posterior after the first update becomes the prior for the second update and so on.

Table 8 shows the measured $F_{m\%}$ and $S_{tpf\%}$ values at 3750 rpm and the true FWW. Thirteen measurements were performed until the FWW reached 0.51 mm (see Table 6). Figure 17 shows the posterior pdf after a single update (top

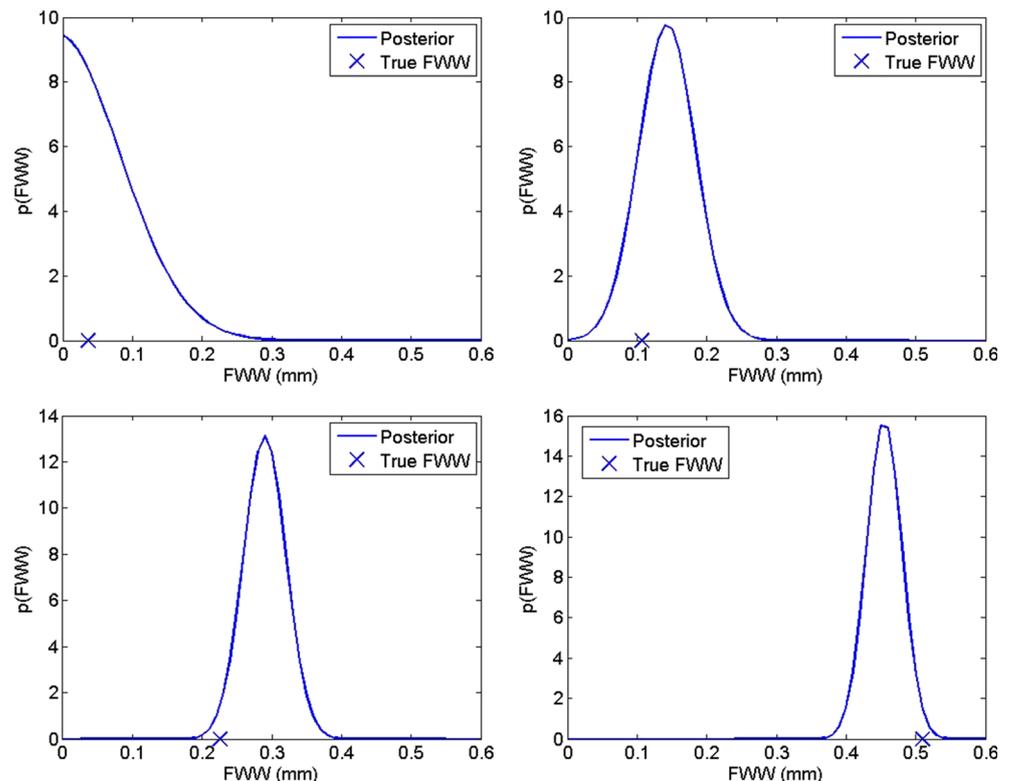
left), five updates (top right), nine updates (bottom left), and 13 updates (bottom right). The true FWW values are also shown, denoted by “x”. Figure 18 shows the posterior probability of FWW after each update using sensor values. The true FWW values are denoted by “x”. The gray-scale color bar denotes the probability density. Results show good agreement between the predicted and the true FWW. The probability of FWW exceeding the limiting FWW can be set as the criteria for detecting a worn tool.

Additional testing was completed at 6250 rpm. The posterior pdf of FWW was calculated using the same procedure described previously. Table 9 shows the measured $F_{m\%}$ and $S_{tpf\%}$ values at 6250 rpm and the true FWW. Figure 19 shows the posterior probability of FWW after each update using sensor values. The gray-scale color bar denotes the probability density. The true FWW values are also shown (denoted by “x”).

7 Discussion

A Bayesian classifier method for tool wear prediction was shown. The classifier method is an efficient method for inference about tool wear under uncertainty. This section presents some considerations in applying the method for tool condition monitoring. First, the prediction accuracy of the method

Fig. 17 Posterior pdf of FWW after one update (top left), five updates (top right), nine updates (bottom left), and 13 updates (bottom right). The true FWW is denoted by “x”



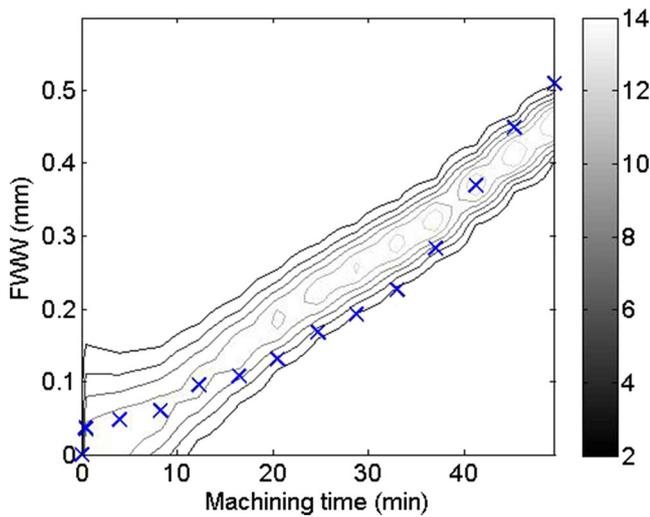


Fig. 18 Posterior FWW probability after each update at 3750 rpm

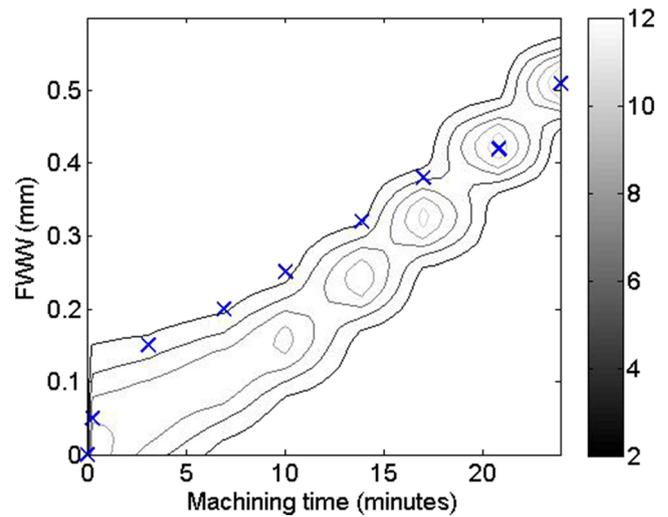


Fig. 19 Posterior FWW probability after each update at 6250 rpm

depends on the amount of training data and the extent of confirmation of the prediction data with the training data set. To illustrate, Fig. 6 shows the training data set using three tests each at 2500, 5000, and 7500 rpm for the prediction of tool wear. Figures 9 and 11 show the prediction data sets at 3750 and 6250 rpm, respectively. Figure 20 shows the comparison of $F_{m\%}$ at 3750 (left) and 6250 rpm (right), respectively, with the training data set. The training data is denoted as “x” and prediction as “o”. As shown in Fig. 20, the prediction data set agrees with the training data set, and therefore, the tool wear predictions are good. However, note that there still exists uncertainty in the tool wear and the values of the sensor signals as calculated in the likelihood functions. The classifier method allows for inference using the training data set. Subsequent prediction data can be incorporated in the training data set. Second, the prediction depends on the number of training experiments and number of sensor metrics. Future work will focus on optimizing the number of training experiments and sensor metrics. As noted, the likelihood function can be used as a criterion for selecting suitable sensor metrics. In addition, the dependence on machining parameters such as

Table 9 Measured $F_{m\%}$ and $S_{ipf\%}$ values at 6250 rpm and the true FWW

Test no.	$F_{m\%}$	$S_{ipf\%}$	True FWW (mm)
1	100.0	100.0	0.05
2	116.1	110.5	0.15
3	129.0	115.5	0.21
4	141.2	123.6	0.25
5	155.3	150.2	0.32
6	157.4	166.5	0.38
7	197.5	186.2	0.44
8	234.9	186.5	0.51

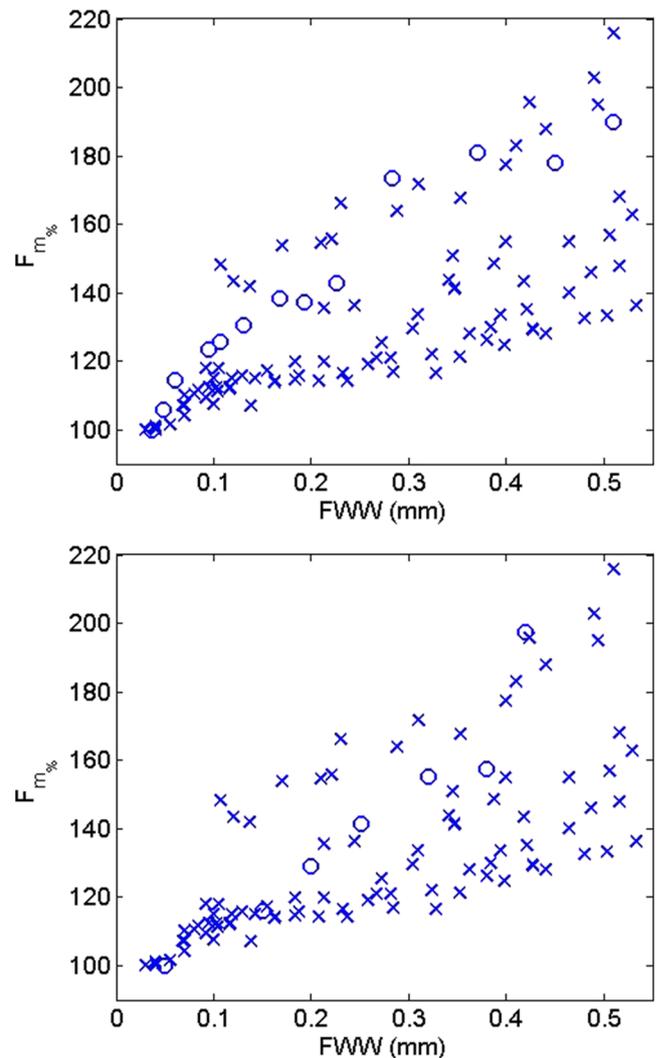


Fig. 20 Comparison of the training and prediction data set for $F_{m\%}$ at 3750 rpm (top) and 6250 rpm (bottom)

spindle speed can be easily incorporated by modifying Eq. 7 as

$$P(TW|F_1, F_2, \dots, F_n) \propto P(TW) \prod_{i=1}^n P(F_i|TW) \quad (15)$$

Equation 15 incorporates dependence on spindle speed by defining the terms in the equation conditional on the spindle speed. This implies that the likelihood (and the prior) is conditioned on the spindle speed and used for tool life prediction at the corresponding speed. The advantage of the classifier method is that it allows multiple sensor fusion for accurate tool wear predictions. In addition, the method is computationally insensitive to the number of sensor metrics. The user can decide a threshold probability of tool flank wear width exceeding the critical value (such as 0.3 mm) based on the application and user risk preference. For example, the threshold probability will be lower for machining an expensive forging as compared to a roughing operation.

8 Conclusions

A naïve Bayes classifier method for tool condition monitoring for both discrete and continuous case was described. The influence of tool wear on the mean force in the time domain and the sum of magnitudes for the tooth passing frequency and its harmonics in the frequency domain were used to train the classifier. In the discrete case, tool wear was divided into discrete states based on the flank wear width, and the probability of the tool wear being in any state is updated using force data. In the continuous case, the probability of the tool flank wear width is updated. The naïve Bayes classifier offers many advantages for tool condition monitoring. The method can be applied to multiple sensor signals, including power, acoustic emission, and vibration without loss of generality. The naïve assumption makes the method computationally inexpensive. To illustrate, using Intel i5 processor with a 4GB RAM on Windows 7, a single computation of posterior probability requires 0.5 ms. This is a result of the computation consisting of a multiplication and normalization. The number of training experiments can be optimized based on the likelihood values calculated. In addition, suitable sensor metrics can also be decided based on the likelihood uncertainty.

References

- Konstantinos S, Athanasios K (2014) Reliability assessment of cutting tool life based on surrogate approximation methods. *Int J Adv Manuf Technol* 71(5–8):1197–1208
- Yeo S, Khoo L, Neo S (2000) Tool condition monitoring using reflectance of chip surface and neural network. *J Intell Manuf* 11: 507–514
- Kurada S, Bradley C (1997) A review of machine vision sensors for tool condition monitoring. *Comp Ind* 34:55–72
- Rehorn AG, Jiang J, Orban PE (2005) State-of-the-art methods and results in tool condition monitoring: a review. *Int J Adv Manuf Technol* 26(7–8):693–710
- Prickett PW, Johns C (1999) An overview of approaches to end milling tool monitoring. *Int J Mach Tools Manuf* 39:105–122
- Dimla E Snr (2000) Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. *Int J Mach Tools Manuf* 40:10735–11098
- Burke LI, Rangwala S (1991) Tool condition monitoring in metal cutting: a neural network approach. *J Intell Manuf* 2(5):269–280
- Sick B (2002) On-line and indirect tool wear monitoring in turning with artificial neural networks—a review of more than a decade of research. *Mech Syst Signal Process* 16(4):487–546
- Dornfeld DA (1990) Neural network sensor fusion for tool condition monitoring. *Ann CIRP* 39:101–105
- Ghosh N, Ravi YB, Patra A, Mukhopadhyay S, Paul S, Mohanty AR, Chattopadhyay AB (2007) Estimation of tool wear during CNC milling using neural network-based sensor fusion. *Mech Syst Signal Process* 21:466–479
- Jemielniak K, Leszek K, Pawel W (1998) Diagnosis of tool wear based on cutting forces and acoustic emission measures as inputs to a neural network. *J Intell Manuf* 9(5):447–455
- Aliustaoglu C, Ertunc HM, Ocak H (2009) Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system. *Mech Syst Signal Process* 23(2):539–546
- Li X, Yuan Z (1998) Tool wear monitoring with wavelet packet transform—fuzzy clustering method. *Wear* 219(2):145–154
- Chen JC, Susanto V (2003) Fuzzy logic based in-process tool-wear monitoring system in face milling operations. *Int J Adv Manuf Technol* 21(3):186–192
- Kuo RJ, Cohen PH (1998) Intelligent tool wear estimation system through artificial neural networks and fuzzy modeling. *Artif Intell Eng* 12(3):229–242
- Li AU, Elbestawi MA (1996) Fuzzy clustering for automated tool condition monitoring in machining. *J Mech Syst Signal Proc* 10(5): 533–550
- Li X, Yao Y, Yuan Z (1997) On-line tool condition monitoring system with wavelet fuzzy neural network. *J Intell Manuf* 8(4):271–276
- Chen JC (1996) A fuzzy nets tool breakage detection system for end milling operations. *J Adv Manuf Technol* 12:153–164
- Dey S, Stori JA (2005) A Bayesian network approach to root cause diagnosis of process variations. *Int J Mach Tools Manuf* 45(1):75–91
- Atlas L, Ostendod M, Bernard GD (2000) Hidden Markov Models for monitoring tool wear. *IEEE Int Conf Acoust Speech Signal Proc* 6:3887–3890
- Ertunc HM, Loparo KA (2001) A decision fusion algorithm for tool wear condition monitoring in drilling. *Int J Mach Tools Manuf* 41(9): 1347–1362
- Ertunc HM, Loparo KA, Ocak H (2001) Tool wear condition monitoring in drilling operations using hidden Markov models (HMMs). *Int J Mach Tools Manuf* 41:1363–1384
- Vallejo AG, Nolasco-Flores JA, Morales-Menéndez R, Sucar LE, Rodríguez CA (2005) Tool-Wear Monitoring Based on Continuous Hidden Markov Models, *Progress in Pattern recognition. Image Anal Appl Lect Comp Sci* 3773:880–890
- Constantinides N, Benett S (1987) An investigation of methods for on-line estimation of tool wear. *Int J Mach Tools Manuf* 27(2):225–237
- Domingos P, Pazzani M (1997) Beyond independence: Conditions for the optimality of the simple Bayesian classifier. *Mach Learn* 29: 103–130

26. Duda RO, Hart PE, Stork DG (2012) Pattern classification. John Wiley & Sons
27. Zhang H (2004) The Optimality of Naive Bayes. Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference. AAAI Press, Miami Beach
28. Geiger D, Heckerman D (1997) A characterization of the Dirichlet distribution through global and local independence. *Ann Stat* 25: 1344–1369
29. Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis. Chapman and Hall CRC, Boca Raton