Precision Engineering 38 (2014) 18–27

Contents lists available at ScienceDirect

Precision Engineering

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

Tool life prediction using Bayesian updating. Part 2: Turning tool life using a Markov Chain Monte Carlo approach

Jaydeep M. Karandikar[a], Ali E. Abbas[b], Tony L. Schmitz[c,*]

[a] Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA
[b] Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[c] Department of Mechanical Engineering and Engineering Science, University of North Carolina at Charlotte, Charlotte, NC, USA

ARTICLE INFO

Article history:
Received 7 November 2011
Received in revised form 25 June 2013
Accepted 27 June 2013
Available online 22 July 2013

Keywords:
Tool wear
Taylor tool life
Bayesian updating
Discrete grid
Markov Chain Monte Carlo
Uncertainty

ABSTRACT

According to the Taylor tool life equation, tool life reduces with increasing cutting speed following a power law. Additional factors can also be added, such as the feed rate, in Taylor-type models. Although these models are posed as deterministic equations, there is inherent uncertainty in the empirical constants and tool life is generally considered a stochastic process. In this work, Bayesian inference is applied to estimate model constants for both milling and turning operations while considering uncertainty.

In Part 1 of the paper, a Taylor tool life model for milling that uses an exponent, *n*, and a constant, *C*, is developed. Bayesian inference is applied to estimate the two model constants using a discrete grid method. Tool wear tests are performed using an uncoated carbide tool and 1018 steel work material. Test results are used to update initial beliefs about the constants and the updated beliefs are then used to predict tool life using a probability density function. In Part 2, an extended form of the Taylor tool life equation is implemented that includes the dependence on both cutting speed and feed for a turning operation. The dependence on cutting speed is quantified by an exponent, *p*, and the dependence on feed by an exponent, *q*; the model also includes a constant, *C*. Bayesian inference is applied to estimate these constants using the Metropolis–Hastings algorithm of the Markov Chain Monte Carlo (MCMC) approach. Turning tests are performed using a carbide tool and MS309 steel work material. The test results are again used to update initial beliefs about the Taylor tool life constants and the updated beliefs are used to predict tool life via a probability density function.

1. Introduction

Tool wear can impose a significant limitation on machining processes, particularly for hard-to-machine materials such as titanium and nickel-based superalloys. Taylor first defined an empirical relationship between tool life and cutting speed using a power law [1]:

$$VT^n = C \tag{1}$$

where V is the cutting speed in m/min, T is the tool life in minutes, and n and C are constants which depend on the tool–workpiece combination. The constant C is defined as the cutting speed required to obtain a tool life of 1 min. Tool life is typically defined as the time required to reach a predetermined flank wear width (FWW), although other wear features (such as crater depth) may also be applied depending on the nature of the tool wear. The Taylor tool life equation can be extended to include other effects, such as feed rate [2]:

$$V^p f_r^q T = C \tag{2}$$

where f_r is the feed in mm/rev in turning and C, p, and q are constants which depend on the tool–workpiece combination. Note that in the extended Taylor tool life equation shown in Eq. (2), the constant C is dimensionless. The Taylor-type tool life model shown in Eq. (2) is deterministic in nature, but uncertainty exists due to: (1) factors that are unknown or not included in the model; and (2) tool-to-tool performance variation. For these reasons, tool wear is often considered to be a stochastic and complex process and, therefore, difficult to predict.

Previous efforts to model tool wear as a stochastic process are available in the literature [3–5]. Vagnorius et al. calculated the optimal tool replacement time by determining the probability of the tool failing before the selected time using a tool reliability function [3]. Liu and Makis derived a recursive formula to determine the cutting tool reliability. The maximum likelihood method was used to determine the unknown parameters in the reliability function

* Corresponding author. Tel.: +1 117046875086.
E-mail address: tony.schmitz@uncc.edu (T.L. Schmitz).

0141-6359/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.precisioneng.2013.06.007

[4]. Wiklund applied the Bayesian approach to monitor tool wear using in-process information [5]. The method presented in this paper uses Bayesian inference to predict tool life at the process planning stage. The distribution of the Taylor tool life constants, p, q, and C, are updated using experimental tool life results. The updated distributions of p, q, and C can then be used to predict tool life. The objective of the paper is to demonstrate the application of Bayesian updating to tool life prediction. The Taylor tool life model is used in this study, despite its potential limitations, because it is well-known and generally understood in the manufacturing community. Without loss of generality, the Bayesian updating method demonstrated in this paper can be applied to other available models [6].

2. Bayesian inference

Bayesian inference, which forms a normative and rational method for belief updating is applied in this work [7]. Let the prior distribution about an uncertain event, A, at a state of information, &, be $\{A|\&\}$, the likelihood of obtaining an experimental result B given that event A occurred be $\{B|A,\&\}$, and the probability of receiving experimental result B (without knowing A has occurred) be $\{B|\&\}$. Bayes' rule is used to determine the posterior belief about event A after observing the experiment results, $\{A|B,\&\}$ as shown in Eq. (3). Using Bayes' rule, information gained through experimentation can be combined with the prior prediction about the event to obtain a posterior distribution.

$$\{A|B, \&\} = \frac{\{A|\&\}\{B|A, \&\}}{\{B|\&\}} \tag{3}$$

As seen in Eq. (2), the Taylor-type tool life model assigns a deterministic value to tool life for the selected cutting speed and feed rate values. In contrast, Bayesian inference assigns a probability distribution to the tool life value at a particular cutting speed/feed rate combination. From a Bayesian standpoint, a variable which is uncertain is treated as a variable which is random and characterized by a probability distribution. The prior, or initial belief of the user, can be based on theoretical considerations, expert opinions, past experience, or data reported in the literature; the prior should be chosen to be as informative as possible. The prior is represented as a probability distribution and, using Bayes' theorem, the probability distribution is updated when new information becomes available (from experiments, for example). As a result, Bayesian inference enables a model to incorporate uncertainty in terms of a probability distribution and beliefs about this distribution to be updated based on experimental results.

In the Taylor-type tool life model provided in Eq. (2), there is uncertainty in the exponents, p and q, and in the constant, C. As a result, there is uncertainty in the tool life, T. This uncertainty can be represented as a joint probability distribution for C, p, and q and, therefore, for the tool life, T. Bayes' rule (Eq. (3)) can be used to update the prior joint distribution of C, p, and q using new information. The new distribution can then be used to update the distribution of tool life, T. In this case, the prior distribution $\{A|\&\}$ is the initial belief about constants C, p, and q. The updating of the constants can be completed using experimental data of tool life. For this case, Bayes' rule is:

$$\{p, q, C|T, \&\} \propto \{p, q, C|\&\}\{T|p, q, C, \&\} \tag{4}$$

where $\{p, q, C|\&\}$ is the prior joint distribution of p, q, and C, $\{T|p, q, C,\&\}$ is the likelihood of observing experimental result of tool life, T, given C, p, and q, and $\{p, q, C|T,\&\}$ is the posterior joint distribution of C, p, and q given an experimental result of tool life, T. Note that the denominator in Eq. (3), $\{B|\&\}$, acts as a normalizing constant. It is not included in Eq. (4).

According to Bayes' rule, the posterior distribution is proportional to the product of the prior and the likelihood. The prior is a three-dimensional joint distribution of the constants C, p, and q. The likelihood and, subsequently, the posterior are also three-dimensional joint distributions of C, p, and q. The grid-based method (see Part 1 of this paper) is computationally expensive for updating a joint distribution with three or more dimensions since it is dependent on the size of the grid. For example, a joint probability density function (pdf) of three variables with a grid size equal to 300 would require at least 2.7×10^6 computations for each update in the grid-based method. As an alternative, the Markov Chain Monte Carlo (MCMC) technique can be used to sample from multivariate posterior distributions for Bayesian inference [7]. Using the MCMC technique, samples can be drawn from the posterior multivariate distribution which can then be used to characterize the distribution. The single-component Metropolis–Hastings (MH) algorithm facilitates sampling from multivariate distributions without sensitivity to the number of variables [9,10]. The algorithm proceeds by considering a single variable at a time and sampling from a univariate proposal distribution. In this study, the single-component MH algorithm of the MCMC technique is used to sample from the joint posterior distribution of the constants C, p, and q. The remainder of the paper is organized as follows. Section 3 describes the use of the MH algorithm to sample from a univariate bimodal pdf and the application to Bayesian inference. Section 4 describes sampling from the joint posterior distribution using the single component MH algorithm. Tool life prediction using the posterior or the updated distributions of tool life is shown in Section 5. Section 6 compares the Bayesian approach to classical regression. Finally, the influence of prior and likelihood uncertainty is discussed in Section 7.

3. Markov chain Monte Carlo (MCMC) method

The Markov Chain Monte Carlo (MCMC) method is a sampling technique used to draw samples from a pdf. Samples are generated from the state space of the variable of interest using a Markov chain mechanism [8]. The most popular method for MCMC is the MH algorithm [9,10]. Let x be the variable of interest. The pdf of variable x is referred to as the target distribution and is denoted by $p(x)$. The MH algorithm uses a proposal distribution (pdf) denoted as $q(x)$. A candidate sample, x^*, drawn from the proposal distribution is either accepted or rejected depending on an acceptance ratio, A. In each iteration, the Markov chain moves to x^* if the sample is accepted. Otherwise, the chain remains at the current value of x. The algorithm proceeds for $N-1$ iterations to obtain N samples from the target distribution using the following steps.

1. Initialize the starting point x^0.
2. For $N-1$ iterations, complete the following four steps:
   a. draw a sample, x^*, from the proposal distribution; the pdf value is $q(x^*|x^i)$, where i denotes the current iteration and the distribution mean is x^i with a selected standard deviation
   b. sample u from a uniform distribution with a lower limit of zero and an upper limit of 1, $U(0, 1)$
   c. compute the acceptance ratio, $A = \min(1, (p(x^*)q(x^i|x^*)/p(x^i)q(x^*|x^i)))$, where $q(x^i|x^*)$ is the pdf value of the proposal distribution at x^i given a mean of x^* with the selected standard deviation, $p(x^*)$ is the pdf value of the target distribution at x^*, and $p(x^i)$ is the pdf value of the target distribution at x^i
   d. if $u < A$, then set the new value of x equal to the new sample, $x^{i+1} = x^*$; otherwise, the value of x remains unchanged, $x^{i+1} = x^i$.

3.1. Algorithm demonstration

To illustrate the algorithm, consider a bimodal pdf as the target distribution; see Eq. (5) [8]. The target distribution needs to be known only up to the normalization constant.

$$p(x) \propto 0.3e^{(-0.2x^2)} + 0.7e^{(-0.2(x-10)^2)} \quad (5)$$

For this example, a normal proposal distribution, $q(x)$, was chosen because a normal distribution facilitates convenient sampling of candidate values for x. For a normal distribution, the proposal distribution is conditioned on the current value of the chain, $q(x^*|x^i)$. The mean of the proposal distribution is x^i and a standard deviation of 10 was selected, i.e. $q(x) = N(x^i, 10)$. The starting point of the chain, x^0, was selected to be zero. At each iteration, i, the following steps were completed. First, a candidate sample, x^*, was randomly drawn from the proposal distribution, $N(x^i, 10)$. The candidate sample was drawn given the current value of the chain, $q(x^*|x^i)$. To illustrate, consider the first iteration. The chain starting point is $x^0 = 0$. Therefore, x^* is a random sample drawn from $N(0, 10)$. Assume the randomly selected value is $x^* = 2$ and it is accepted as $x^{(1)}$. In the second iteration, the random sample is drawn from $N(2, 10)$. If the sample is 12 and it is rejected, then the current value of $x^{(2)}$ remains at 2. In the third iteration the random sample will again be drawn from $N(2, 10)$.

Second, $p(x^*)$ and $p(x^i)$ were calculated using Eq. 5 for the target distribution. Third, $q(x^*|x^i)$ and $q(x^i|x^*)$ were calculated. Since a normal distribution is symmetric, the $q(x^*|x^i)$ and $q(x^i|x^*)$ are equal. Fourth, the acceptance ratio, A, was calculated. For the normal proposal distribution, the acceptance ratio simplifies to:

$$A = \min\left(1, \frac{p(x^*)}{p(x^i)}\right) \quad (6)$$

Fifth, A was compared to a random sample, u, drawn from a uniform distribution with a range from 0 to 1. Finally, if u was less than A, then the candidate sample was accepted and x^{i+1} was updated using $x^{i+1} = x^*$. Otherwise, the sample was rejected and $x^{i+1} = x^i$. These steps were repeated for $N-1$ iterations to obtain N samples of x from the target pdf described by Eq. (5).

The MH algorithm was carried out for 10,000 iterations. Fig. 1 shows the histogram of the 10,000 samples and target distribution from Eq. (5) (left) and x values for each iteration (right). It is seen that the samples approximate the target pdf quite well. Note that the histogram and target distribution were normalized to obtain a unit area.

Although the MH algorithm is effective for sampling from any target distribution, there are a number of considerations in its application. The success of the algorithm depends on the choice of proposal distribution. In theory, the chain should converge to the stationary target distribution for any proposal distribution [11,12]. However, the proposal distribution may affect the convergence and mixing of the chain. In general, the proposal distribution may be selected so that the sampling is convenient. If the proposal distribution was chosen to be uniform, then it is not dependent on the current value of x since x^* is drawn from a preselected range of x values. A uniform proposal distribution may therefore be less efficient because the random samples are independent of the current state of the chain. Using a normal proposal distribution where x^* is dependent on x^i is referred to as random walk Metropolis sampling, while the uniform proposal approach where x^* is independent of x^i is called independent Metropolis–Hastings sampling.

For a normal proposal distribution, the choice of the standard deviation can also affect the results. A larger standard deviation causes greater jumps in the domain or the state space of the variable. Thus, the candidate sample has a higher probability of being rejected, which yields $x^{i+1} = x^i$. On the other hand, while a smaller variance will tend to accept a higher number of random samples, this can result in slower convergence of the chain.

The number of iterations should be large enough to ensure convergence to the statistical moments of the target distribution. The convergence to the true statistical moments can be observed by repeating the algorithm using different starting values and varying the number of iterations. The starting value of the chain has no effect for a large number of iterations [13]. The initial iterations are typically discarded (called the burn-in time) and the chain subsequently settles to a stationary distribution. A practical way to evaluate convergence to the chain's stationary distribution is by observing the traces and histograms of the variables (e.g. see Fig. 1). Despite these potential limitations, the MH algorithm for MCMC works well and can effectively be used to draw samples from multivariate distributions.

3.2. Application to Bayesian inference

This section describes the application of MCMC to Bayesian inference. As stated in Section 2, Bayesian inference provides a formal way to update beliefs about the posterior distribution (the normalized product of the prior and the likelihood functions) using experimental results. The prior for this analysis is a joint pdf of the Taylor tool life constants, C, p, and q. As a result, the posterior is also a joint pdf of the constants. The MCMC technique is applied here because it can be used to sample from multivariate posterior distributions. The joint posterior pdf is the target pdf for the MCMC approach. Note that the normalizing constant for the posterior pdf is not required for sampling.

The MH algorithm was detailed for a single variable in Section 3. To sample from a joint pdf, the algorithm samples one variable at a time and then proceeds sequentially to sample the remaining variables. To illustrate, consider a joint target pdf of n variables: $x_1, x_2, x_3, \ldots, x_n$. First, the starting value for all the variables is initialized, $[x_1^0, x_2^0, x_3^0, \ldots x_n^0]$. The sequence of variable sampling does not influence the convergence of the algorithm, so let the algorithm proceed in the order, $x_1 \to x_2 \to x_3 \to \ldots x_n$. The sampling for each variable is carried out using a univariate proposal distribution for that variable. The target and the proposal pdf for each variable are conditioned on the current values of the other variables. For example, consider a candidate sample, x_1^*, drawn from the univariate proposal distribution for x_1. The candidate sample from the joint pdf is then $[x_1^*, x_2^0, x_3^0, \ldots x_n^0]$. The candidate sample, x_1^*, is either accepted or rejected given the current values of $x_2, x_3, \ldots, x_n$. Thus, the target pdf values of x_1^* and x_1^0 are conditional on the current values of the other variables, $x_2^0, x_3^0, \ldots x_n^0$ and are denoted as $p(x_1^*|x_1^0, x_2^0, \ldots x_n^0)$ and $p(x_1^0|x_1^0, x_2^0, \ldots x_n^0)$. The proposal univariate pdfs are also conditional on the current values of the chain and are denoted as $q(x_1^*|x_1, x_1, x_3 \ldots x_n)$ and $q(x_1|x_1^*, x_1, x_3 \ldots x_n)$ for x_1^* and x_1, respectively. The chain either stays at the current point, $[x_1^0, x_2^0, x_3^0, \ldots x_n^0]$ or moves to a neighboring point, $[x_1^*, x_2^0, x_3^0, \ldots x_n^0]$, which differs only in one component of the current state (x_1 in this case). The procedure is repeated for all variables in each iteration. The acceptance ratio is:

$$A = \min\left(1, \frac{p(x_1^*|x_2, x_3, \ldots, x_n) q(x_1^i|[x_1^*, x_2, x_3, \ldots, x_n])}{p(x_1^i|x_2, x_3, \ldots, x_n) q(x_1^*|[x_1^i, x_2, x_3, \ldots, x_n])}\right) \quad (7)$$

where the value of each of the four joint pdfs must each be calculated. The value of A is compared to a random sample, u, from a uniform distribution with a range from 0 to 1 and x_1^* is either accepted or rejected to obtain x_1^{i+1}. The algorithm is repeated using the updated values of each variable continually for the next variable. Thus, x_2^{i+1} is determined using $x_1^{i+1}, x_3^i, \ldots, x_n^i$, x_3^{i+1} is determined using $x_1^{i+1}, x_2^{i+1}, \ldots, x_n^i$, and so on for n variables. A single iteration updates all the variables. The algorithm is then

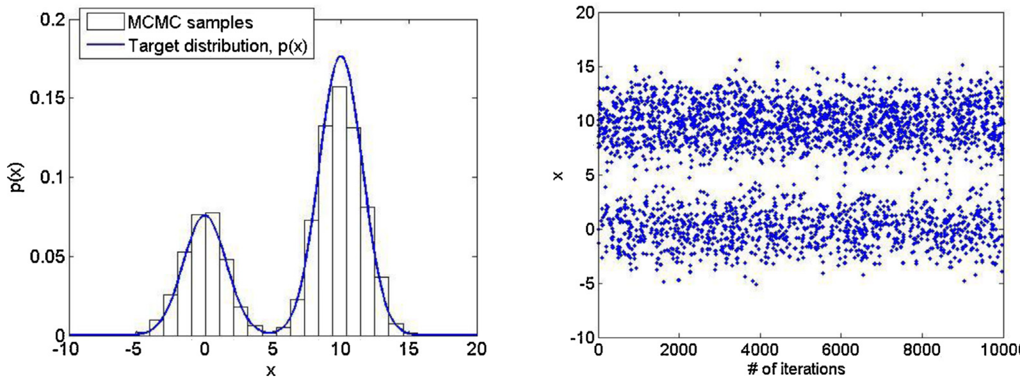

Fig. 1. Histogram of MCMC samples and target distribution (left) and *x* values for each iteration (right).

carried out for $N-1$ iterations to obtain samples from the joint target pdf. An alternative method is to sample from a joint proposal pdf and accept or reject it using the MH algorithm. However, it is much simpler to sample from univariate proposal distributions for each variable and is computationally less expensive.

4. Bayesian updating using the Markov Chain Monte Carlo method

In this section, the Markov Chain Monte Carlo (MCMC) method for Bayesian updating of Taylor tool life constants using experimental tool life data is described. According to Bayes' rule, the posterior distribution is proportional to the product of the prior and the likelihood. This is the process of learning, or updating beliefs, when experimental results are available.

4.1. Establishing the prior

The cutting tool used for wear testing was a coated carbide insert (Kennametal KC9110) and the workpiece material was forged AISI 4137 chrome alloy steel. The turning experiments were performed on an Okuma LC-40 CNC lathe. The first step in applying Bayesian inference was to determine the prior distribution. In this case, the prior was a joint probability distribution for the Taylor tool life constants, *C*, *p*, *and q*. The prior, or initial belief of the user, can be based upon theoretical considerations, expert opinions, past experiences, or data reported in the literature. For this case, the following information was applied. It was believed that:

1. in general, the value exponent *p* is greater than exponent *q*, due to a stronger influence of cutting speed on tool wear
2. the value of *p* is between 2 and 6 and *q* is between 1.5 and 3 [2]
3. the value of *C* is sensitive to the values of *p* and *q* due to the nature of the tool life equation and is in the range of 1×10^6 to 1×10^8 m/min.

In this case, information was available to supply only a general range of the constants *C*, *p*, and *q*. Therefore, the prior was assumed to be joint uniform distribution, i.e., it was equally likely to obtain any value within the specified range. The constants were assumed to be independent for the prior. In cases where experimental data using the same tool–material combination is available, a more informative prior (such as a normal distribution) can be selected. For this study, the marginal prior pdfs of the constants were specified as: $p=U(2, 6)$, $q=U(1, 5)$, and $C=U(10^6, 10^8)$, where *U* represents a uniform distribution and the parenthetical terms indicated the lower and upper values of the range.

4.2. Experimental setup and results

The prior *C*, *p*, and *q* joint distribution was updated using experimental results to obtain the posterior joint distribution. The initial outer diameter of the steel workpiece was 174.62 mm. The depth of cut was 4.06 mm and the length of cut for a single pass was 139.7 mm with a chamfer of 63.4 degrees at the end of each cut. The spindle speed was varied to maintain constant cutting speed with reducing workpiece diameter as additional cuts were completed. A digital microscope (60× magnification) was used to image the flank and rake surfaces within the lathe enclosure to avoid removing the insert from the tool holder during the wear testing. The wear status of the tool was recorded after each pass and the calibrated digital images were used to identify the flank wear width (FWW). The first test was completed using a cutting speed of 153.6 m/min and a feed per revolution of 0.51 mm/rev. Tool life was defined as the time required for the FWW to reach 0.4 mm. The time to reach a FWW of 0.4 mm was 22.47 min. Fig. 2 shows the images of the relief face at selected cutting times.

Two additional tests were performed at $\{V=192.01$ m/min, $f_r=0.61$ mm/rev$\}$ and $\{V=230.42$ m/min, $f_r=0.51$ mm/rev$\}$. Fig. 3 shows the growth in FWW for all three test conditions. The 'o' symbols denote the intervals at which the FWW was recorded. The tool life was linearly interpolated between adjacent intervals if it exceeded 0.4 mm at the final measurement interval to determine the tool life. The results of the three tests are summarized in Table 1. As expected, tool life reduces with increased cutting speed and feed.

4.3. Bayesian updating

As described in Section 3.2, a single-component MH algorithm was used to sample from the joint posterior pdf of the Taylor tool life constants: *p*, *q*, and *C*. The posterior joint pdf was the target pdf for the MH algorithm. The prior distribution of the constants was assumed to be a joint uniform distribution. As noted in Section 4.1, this distribution represents a less informative prior than a normal distribution with a mean and standard deviation.

The single-component MH algorithm proceeded as follows. First, the starting point for the Markov chain, $x^0=[C^0 p^0 q^0]$, was selected to be the midpoints of the uniform *C*, *p*, and *q* distributions, $x^0=[5\times10^7\ 4.0\ 3.0]$. The sampling was completed one coefficient

Table 1
Experimental tool life results used for updating.

Test #	Cutting speed (m/min)	Feed (mm/rev)	Tool life (min)
1	153.6	0.51	22.47
2	192.01	0.61	6.52
3	230.42	0.51	5.58

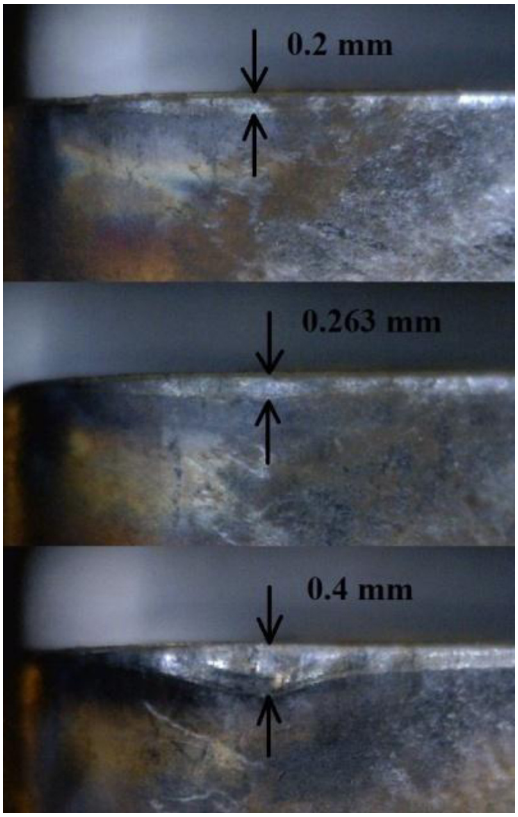

Fig. 2. Images of FWW at 60× magnification. The cutting times from top to bottom are {6.8, 15.5, and 22.4} min.

at a time in the order $C \rightarrow p \rightarrow q$. The proposal distribution for each constant was selected to be normal. The standard deviations of the proposal distribution of constants, C, p, and q were 1×10^7, 0.2, and 0.2, respectively. To begin, a candidate sample, C^*, was drawn from the proposal distribution of C. The posterior, or target, pdf values, of each constant were conditional on the values of the other coefficients. The posterior pdf for C, $p(C^0|p^0\,q^0)$, was the product of the prior and likelihood functions. The tool life value was calculated using the current state of the chain, $[C^0\,p^0\,q^0]$, as input to Eq. (2). Because tool wear is stochastic, there is uncertainty in tool life

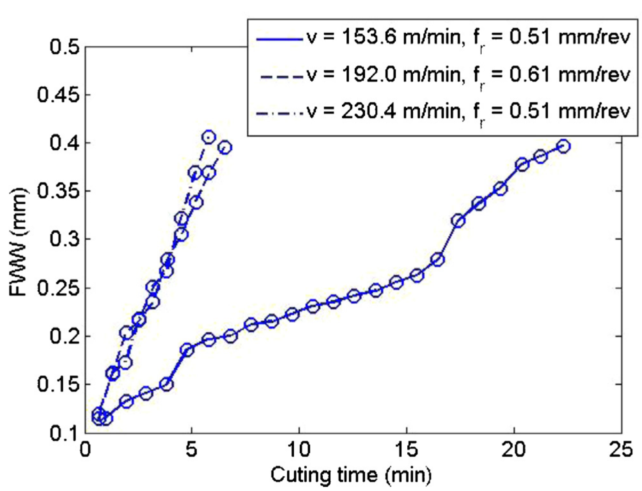

Fig. 3. Variation of FWW with cutting time at various test conditions.

predictions. Therefore, the tool life calculated using the current state of chain as input to Eq. (2) was assumed to be normally distributed with a standard deviation equal to 10% of the experimental tool life; the 10% level was based on the user's belief regarding experimental uncertainty in measured tool life. This gave a pdf for tool life calculated using the current state of the chain. The likelihood is the value of the pdf for the experimental tool life (listed in Table 1). Therefore, the likelihood described how likely it was to obtain the experimental tool life given the current state of the chain. For multiple measurement results, the total likelihood pdf was the product of the likelihood pdfs for all measurements. The same procedure was followed to determine the posterior pdf value for C^*, $p(C^*|p^i\,q^i)$. Since the proposal distribution was normal, the acceptance ratio was calculated using Eq. (8).

$$A = \min\left(1, \frac{p(C^*|p, q)}{p(C^i|p, q)}\right) \tag{8}$$

The acceptance ratio was compared to a random sample, u, from a uniform distribution (with a range from 0 to 1) to assign the value of $C^{(1)}$ to be either C^* or C^0. To update the constants, C, p, and q, the algorithm considered one coefficient at a time and then proceeded to sequentially update the remaining coefficients. For the order $C \rightarrow p \rightarrow q$, $C^{(1)}$ was used to update p^0. Next, $C^{(1)}$ and $p^{(1)}$ were used for q^0. A single iteration provided samples for all the three constants. This sequence was repeated for $N-1$ iterations giving N samples from the joint posterior pdf of the Taylor tool life constants. Note that the standard deviations of the proposal distributions affect the convergence of the chain. As a rule of thumb, the standard deviation should be large enough to draw adequate samples to explore the domain. However, a very large standard deviation leads to a higher probability of candidate samples being rejected.

The MH algorithm was exercised for 100,000 iterations. Fig. 4 shows the sample traces of the constants C (top left), p (top right), and q (bottom left) for all iterations. The initial burn-in time was selected to be 1000 iterations. Fig. 5 shows a comparison between the prior marginal pdfs and posterior sample histograms of the constants. The histograms represent the marginal posterior pdfs of the constants and were normalized to obtain a unit area. Note that the value of constant C is extremely sensitive to the exponents p and q due to the power law nature of the Taylor-type tool life equation. The distribution of the constants is due to the uncertainty in the tool life values. MCMC provides samples from the joint posterior pdf of the Taylor tool life constants, C, p, and q. The samples can then be used in a Monte Carlo simulation to determine the posterior tool life predictions at any speed and feed.

5. Tool life predictions

The samples from the joint posterior pdf of the Taylor tool life constants, C, p, and q, were used in a Monte Carlo simulation to determine the posterior tool life predictions. The posterior or the updated distribution of tool life can be used to predict tool life at cutting conditions other than the ones at which the tool wear experiments were performed. The posterior distribution was used to predict tool life for two new test conditions: {V = 192.01 m/min, f_r = 0.51 mm/rev} and {V = 230.42 m/min, f_r = 0.61 mm/rev}. Two tests were performed for each parameter combination. Other conditions were maintained constant and the same procedure was followed to measure tool life. As before, tool life was selected to be the time for the tool to reach a FWW of 0.4 mm. The tool life values obtained from these tests were compared to the predicted posterior distributions of tool life at the corresponding test parameters, V and f_r (see Figs. 6 and 7).

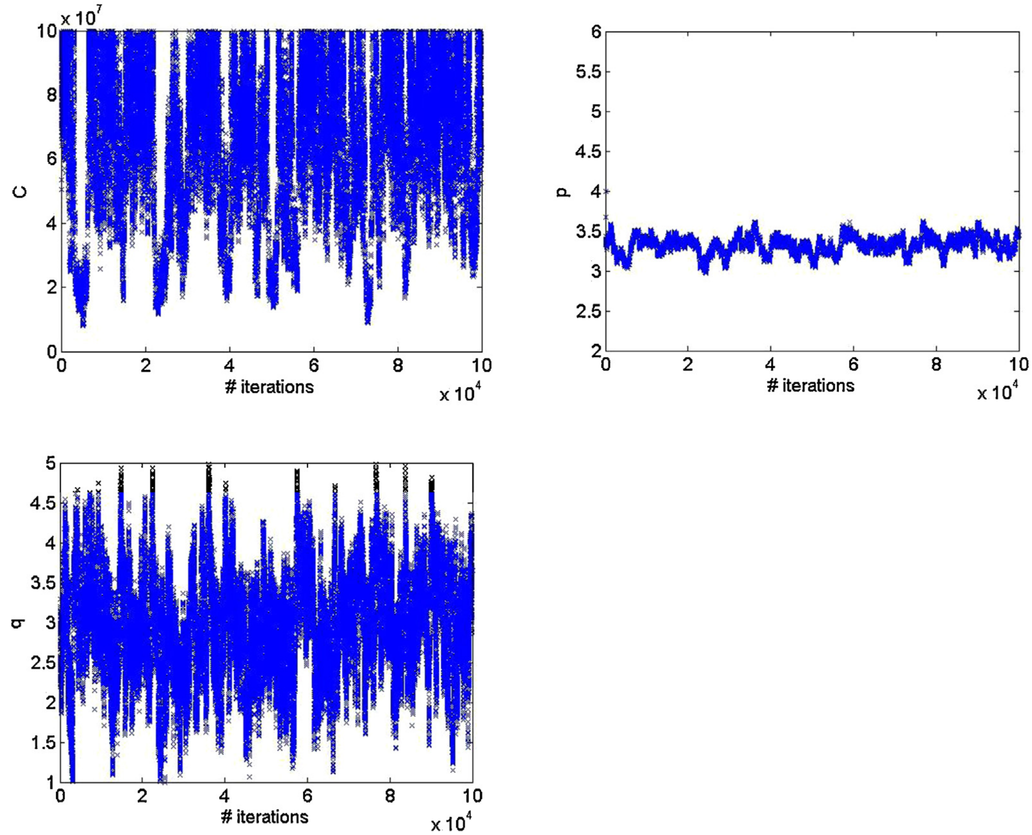

Fig. 4. Traces for C (top left), p (top right), and q (bottom left).

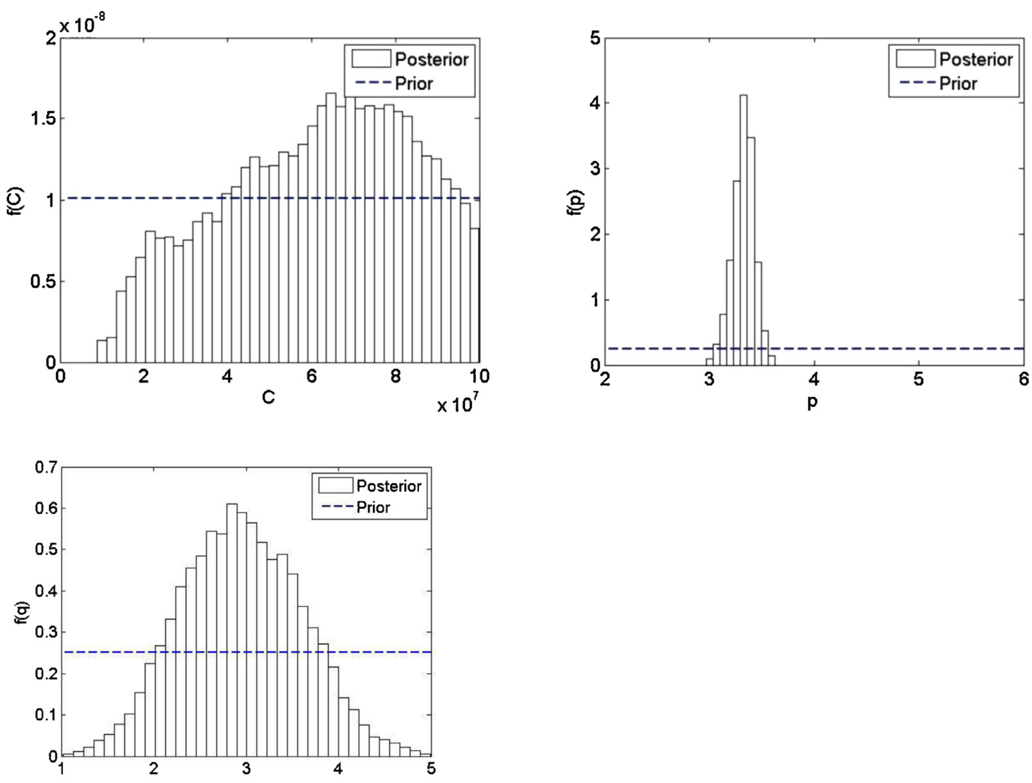

Fig. 5. Posterior and prior distributions of constants C (top left), p (top right), and q (bottom left).

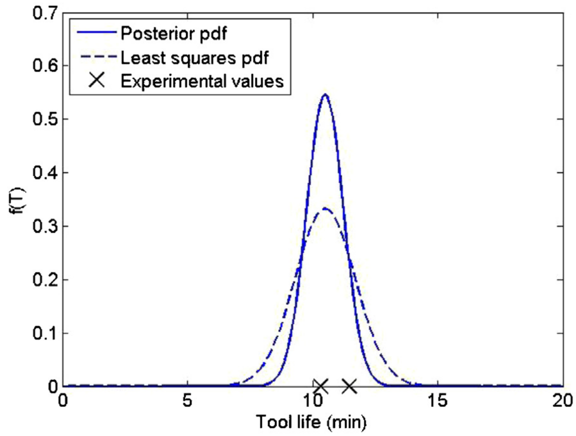

Fig. 6. Comparison of the posterior tool life and the deterministic prediction of tool life at V = 192.01 m/min and f_r = 0.51 mm/rev. The 'x' symbols denote experimental values and the 'o' symbol represents the curve fit prediction.

6. Comparison with the deterministic method

Bayesian inference assigns a probability distribution over a range of the variable(s) of interest, while deterministic methods (such as curve fitting) predict a single value with an associated uncertainty. For example, the Taylor tool life constants were calculated using the experimental tool life data listed in Table 1. The C, p, and q values were 9.83×10^7, 3.39, and 2.63, respectively. Note that the Bayesian estimate does not center on the least squares solution due to the non-linearity of the model. Fig. 5 shows that the prior restricts the possible values of C, causing a shift in the prediction. Thus, the Taylor-type tool life form for the experimental data was:

$$V^{3.39} f_r^{2.63} T = 9.83 \times 10^7 \quad (9)$$

To obtain a probabilistic prediction, a Monte Carlo simulation was performed to simulate the experimental values using the values obtained and the experimental uncertainty. The uncertainty (standard deviation) in the experimental result was again taken to be 10% of the measured value. 1×10^4 random samples were drawn from the normal distribution with the mean equal to the experimental value and standard deviation equal to 10% of the experimental value and the Taylor tool life constants were calculated for each sample combination of tool life values. The mean and standard deviation were 3.39 and 0.35 for p, 2.63 and 0.68 for q, and 7.26×10^7 and 6.37×10^7 m/min for C, respectively. Note that the mean and standard deviation of C are higher than the deterministic values shown in Eq. (9). This is due to the large variations in the value of C from the Monte Carlo simulation. The values of p, q, and C from the simulation were used to predict the distribution of tool life. Table 2 presents a comparison between the tool life distribution from the Monte Carlo simulation, the posterior tool life distribution after Bayesian updating, and the experimentally obtained values. Note that the mean values from the Monte Carlo simulation are similar to the mean Bayesian posterior predictions. However, the uncertainty is greater due to the small size of the experimental dataset. Bayesian inference takes into account prior beliefs and experimental evidence and, therefore, gives good results even for a small number of data points. Fig. 6 shows the experimental values at V = 192.02 m/min and f_r = 0.51 mm/rev (denoted by 'x'), the posterior distribution after updating, and the deterministic tool life mean predictions (denoted by 'o'). Fig. 7 shows the results for V = 230.42 m/min and f_r = 0.61 mm/rev. Using the Bayesian approach, a probability of tool failure can be determined and machining parameters can be selected accordingly depending on the nature of the operation and the risk preferences of the user. Additionally, while a statistical curve fit requires a large amount of data to achieve confidence in the fit parameters, Bayesian updating of an informed prior using only a few tests can lead to an accurate prediction with the inherent characterization of prediction uncertainty.

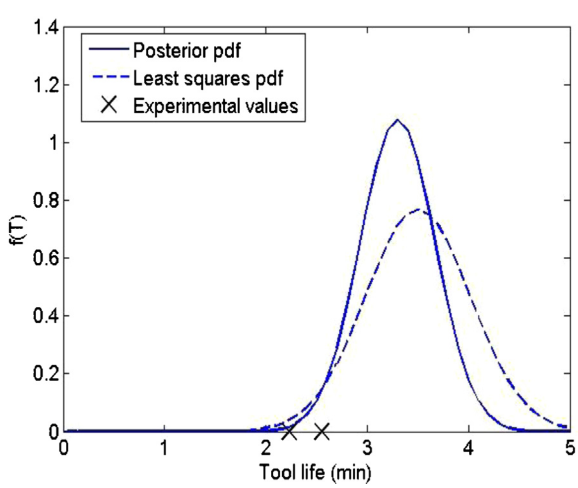

Fig. 7. Comparison of the posterior tool life and the deterministic prediction at V = 230.42 m/min and f_r = 0.61 mm/rev. The 'x' symbols denote experimental values and the 'o' symbol represents the curve fit prediction.

7. Effects of prior selection and likelihood uncertainty on tool life predictions

In Bayesian inference, the posterior distribution is the product of the prior and the likelihood distributions. Clearly, the posterior tool life predictions depend on the selection of the prior and the likelihood distributions. In this section, the influence of the prior distribution and likelihood uncertainty on the posterior is evaluated. First, the influence of the prior distribution on the posterior distribution is examined. As stated in Section 4.1, a uniform prior was selected for the tool wear study. A uniform prior represents a non-informative case, where the parameter value is equally likely to take any value within with the specified range. To evaluate the influence of the prior distribution on the posterior pdf, the algorithm was repeated using normal marginal pdfs as the prior for the constants. The marginal prior pdfs were selected as:

- $C = N(5 \times 10^7, 1.5 \times 10^7)$
- $p = N(4, 0.5)$
- $q = N(3, 0.5)$

where N denotes a normal distribution and the values in the parentheses identify the mean and standard deviation, respectively. Fig. 8 displays a comparison of the prior marginal pdfs and posterior marginal pdfs of the constants using the selected normal prior distributions. The posterior distributions were obtained after updating using the experimental results in Table 1. Fig. 9 compares the posterior tool life pdfs obtained using uniform and normal prior distributions at V = 230.42 m/min and f_r = 0.61 mm/rev. The experimental value is denoted as 'x' and the deterministic prediction is denoted by 'o'.

For a uniform prior, the posterior is the same as the likelihood. Therefore, in the case of non-informative priors, the posterior is only dependent on the experimental tool life data. However, a normal prior represents a more informative case where knowledge of the possible values of the variable of interest is available. An informative prior, such as the normal distribution, thus reflects the most

Table 2
Comparison between the predicted values obtained from curve fit, Bayesian updating and experimental values.

Test #	Cutting speed (m/min)	Feed rate (mm/rev)	Tool life (min)	Least square prediction (min)	Bayes prediction (min)
1	192.01	0.51	11.46	(10.5, 1.2)	(10.5, 0.73)
2	192.01	0.51	10.33		
4	230.42	0.61	2.23	(3.5, 0.52)	(3.3, 0.37)
5	230.42	0.61	2.55		

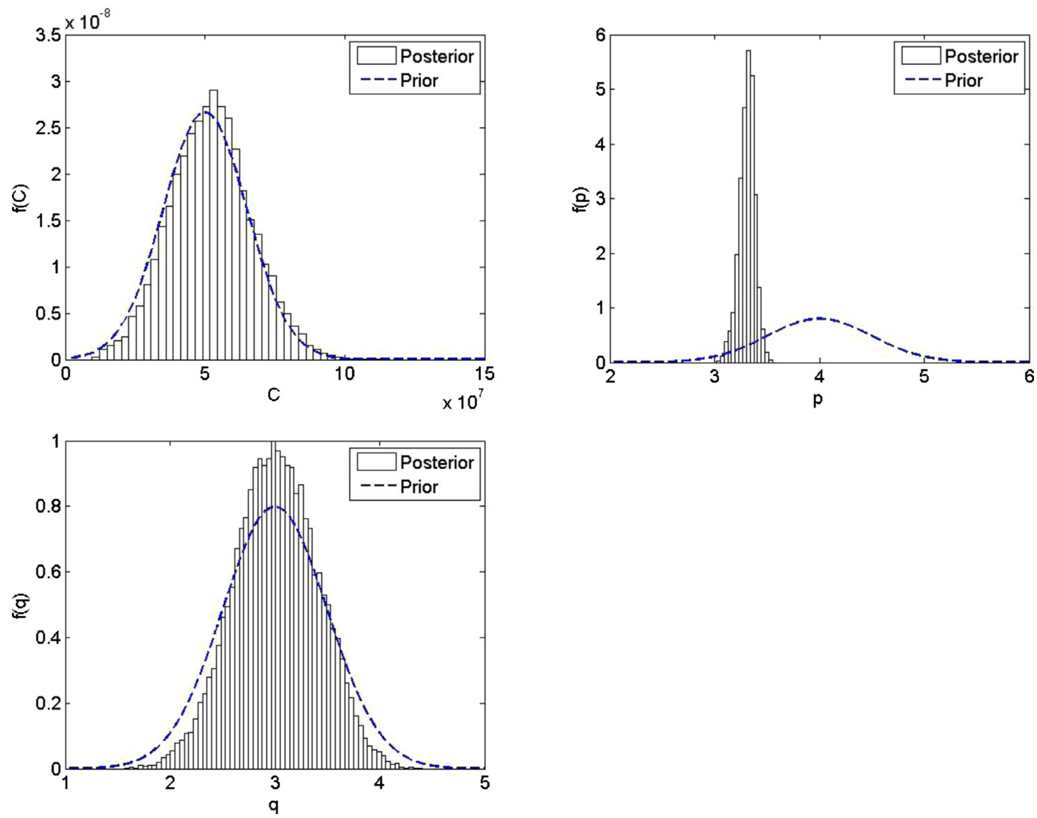

Fig. 8. Posterior and prior distributions of constants *C* (top left), *p* (top right), and *q* (bottom left).

likely values of the variable. The prior represents the initial degree of belief about the constants; if the initial belief is far from the true value, this affects the final results. In general, the prior should be chosen to be as informative as possible considering all the available information. If enough data or prior knowledge is not available, a uniform prior may be selected.

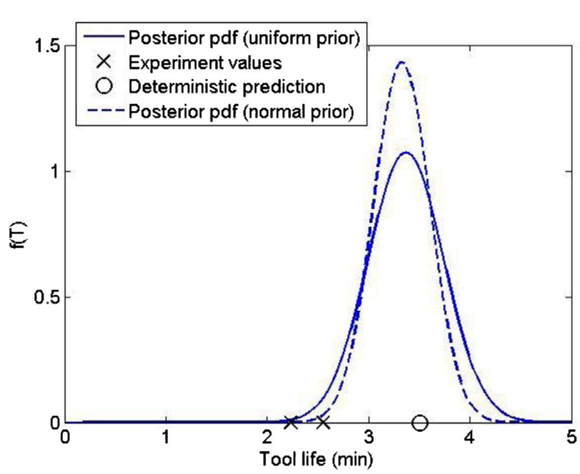

Fig. 9. Comparison of tool life predictions at $V = 230.42$/min and $f_r = 0.61$ mm/rev using uniform and normal priors.

Next, the influence of the likelihood uncertainty on the posterior tool life pdf was evaluated. For previous posterior, or target, pdf calculations in this study, an uncertainty of 10% (one standard deviation) of the experimental tool life was assumed. The assumed likelihood uncertainty was varied (1%, 5%, 10%, and 20%), and the algorithm was repeated for each case. Figs. 10 and 11 display the posterior pdfs for *C*, *p*, and *q* at 20% and 5% uncertainty, respectively. It is seen that the likelihood uncertainty affects the posterior distribution. Fig. 12 shows the posterior tool life pdf at $V = 230.42$/min and $f_r = 0.61$ mm/rev for different likelihood uncertainties. The experimental values are denoted as 'x' and the deterministic prediction as 'o'. Note that the posterior pdf at 1% likelihood uncertainty is not normalized for demonstration purposes. As shown in Figs. 10 and 11, the assumed standard deviation affects the spread of the posterior pdfs of the constants. The posterior tool life pdf obtained using a 20% likelihood uncertainty has a higher standard deviation than the posterior pdf obtained using a 10% likelihood uncertainty. Therefore, the prediction using a 20% uncertainty is more conservative. On the other hand, for low values of the standard deviation (less than 5%), the likelihood function approaches the deterministic prediction. This will result, in general, in a less conservative tool life prediction. As the likelihood uncertainty approaches zero, the posterior pdf standard deviation also approaches zero and the mean approaches the deterministic prediction in this case. In general, the value should be based on the uncertainty expected in the experimental result. Multiple

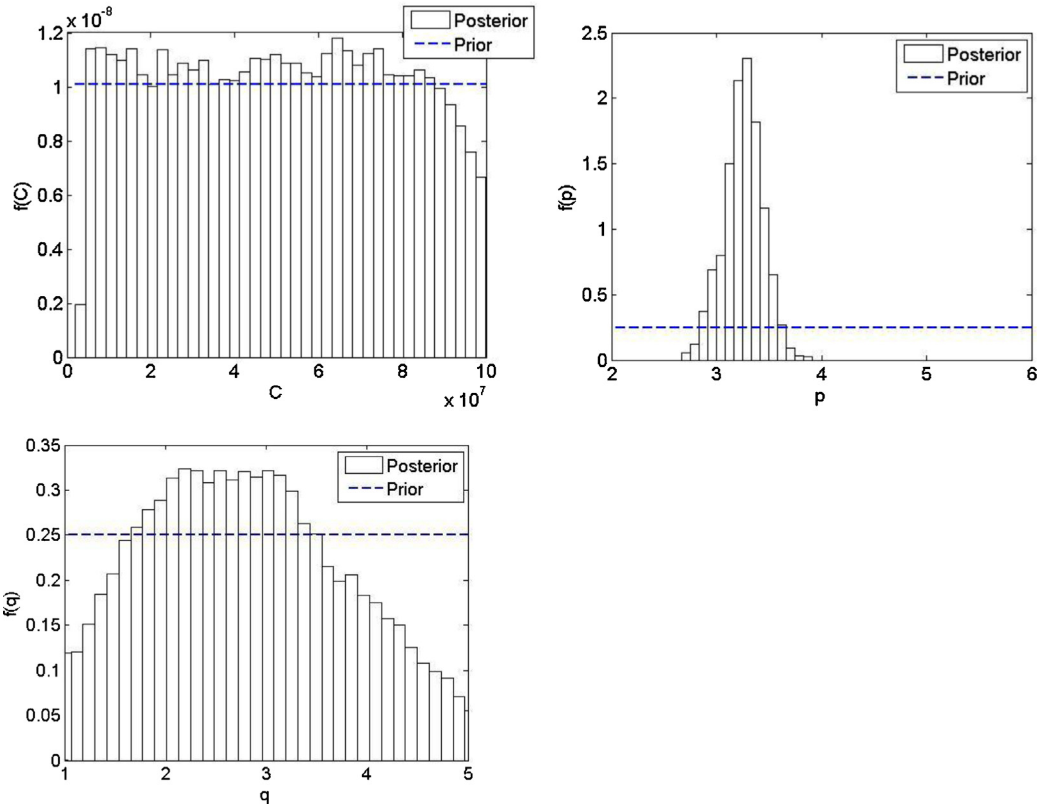

Fig. 10. Posterior and prior distributions of constants *C* (top left), *p* (top right), and *q* (bottom left) for a 20% likelihood uncertainty.

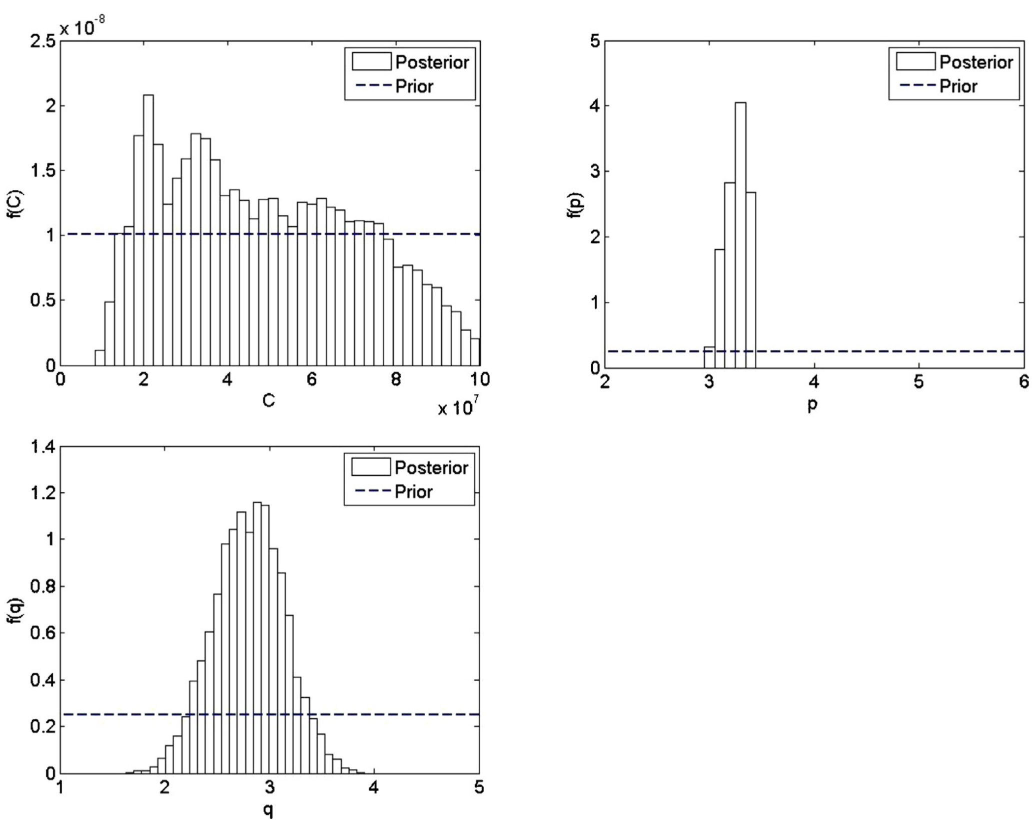

Fig. 11. Posterior and prior distributions of constants *C* (top left), *p* (top right), and *q* (bottom left) for a 5% likelihood uncertainty.

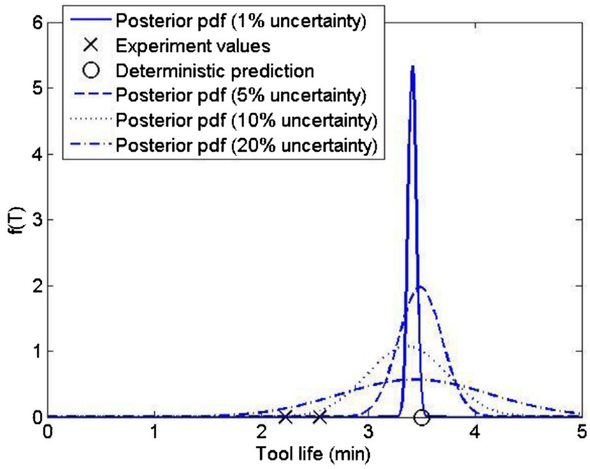

Fig. 12. Tool life predictions at V = 230.42/min and f_r = 0.61 mm/rev for different likelihood uncertainty levels. Note that the posterior pdf at 1% likelihood uncertainty is not normalized for demonstration purposes.

experiments could be performed at a single spindle speed to determine the distribution, for example. From the values of tool life obtained in the prediction set, a 10% standard deviation is reasonable.

8. Conclusions

A Bayesian inference approach to tool life prediction was demonstrated. The Taylor-type tool life equation constants, C, p, and q, were updated using experimental results for turning tests using a coated carbide insert and alloy steel workpiece. The single-component Metropolis–Hastings (MH) algorithm for the Markov Chain Monte Carlo (MCMC) approach was used to sample from the joint posterior pdf of the three constants. The samples were then used to determine the posterior distribution of tool life, which was subsequently used to predict tool life at different cutting conditions.

Bayesian inference assigns a probability distribution over a range of the variable(s) of interest. The probability distributions of the predictions can be updated when new information is available (in the form of experimental results, for example). When this new information is obtained, uncertainty in the prior distributions can be reduced. Bayesian inference provides a way to combine prior data with experimental values to update beliefs about an uncertain variable. By combining prior knowledge and experimental results, Bayesian inference reduces the number of experiments required for uncertainty quantification. When combined with rational decision making theories, an optimal sequence of experiments and value gained from experimental results can also be determined. Finally, the Metropolis–Hastings algorithm is a powerful tool for sampling from multivariate distributions. The single-component MH algorithm for MCMC facilitates updating of joint distributions without significant computational expense.

Acknowledgements

The authors gratefully acknowledge financial support from the National Science Foundation (CMMI-0927051 and CMMI-0926667). They would also like to thank E. Deane and M. Hernandez, University of Central Florida, and H.S. Kim, University of Florida, for their help with the tool wear experiments.

References

[1] Taylor FW. On the art of cutting metals. Transactions of the ASME 1906;28:31–248.
[2] Tlusty J. Manufacturing process and equipment. Upper Saddle River, NJ: Prentice Hall; 2000. p. 463.
[3] Vagnorius Z, Rausand M, Sorby K. Determining optimal replacement time for metal cutting tools. European Journal of Operational Research 2010;206:407–16.
[4] Liu H, Makis V. Cutting-tool reliability assessment in variable machining conditions. IEEE Transactions on Reliability 1996;45(4):573–81.
[5] Wiklund H. Bayesian and regression approaches to on-line prediction of residual tool life. Quality and Reliability Engineering International 1998;14:303–9.
[6] Astakhov VP. Tribology of metal cutting. Boston, MA: Elsevier; 2006.
[7] Gelman A, Carlin J, Stern H, Rubin D. Bayesian data analysis. Boca Raton, FL: Chapman and Hall/CRC Press; 2009.
[8] Andrieu C, Freitas N, de Doucet A, Jordan M. An introduction to MCMC for machine learning. Machine Learning 2003;50:5–43.
[9] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. Journal of Chemical Physics 1953;21(6):1087–92.
[10] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57:97–109.
[11] Roberts GO, Gelman A, Gilks WR. Weak convergence and optimal scaling of random walk metropolis algorithms. Annals of Applied Probability 1997;7(1):110–2.
[12] Chib S, Greenberg E. Understanding the Metropolis Hastings algorithm. The American Statistician 1995;49(4):327–35.
[13] Gelfand A, Smith A. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 1990;85(410):398–409.