Parameter Inference Under Uncertainty in End-Milling γ'-Strengthened Difficult-to-Machine Alloy

Farbod Akhavan Niaki
International Center for Automotive Research,
Clemson University,
Greenville, SC 29607

Durul Ulutan
Mechanical Engineering,
Bucknell University,
Lewisburg, PA 17837

Laine Mears
International Center for Automotive Research,
Clemson University,
Greenville, SC 29607

Nickel-based alloys are those of materials that are maintaining their strength at high temperature. This feature makes these alloys a suitable candidate for power generation industry. However, high wear rate and tooling cost are known as the challenges in machining Ni-based alloys. The high wear rate causes a rapid failure of the tool, and therefore, fewer data will be available for model development. In addition, variations in material properties and hardness, residual stress, tool runout, and tolerances are some uncontrollable effects adding uncertainties to the currently developed models. To address these challenges, a probabilistic Bayesian approach using Markov Chain Monte Carlo (MCMC) method has been used in this work. The MCMC method is a powerful tool for parameter inference and quantification of embedded uncertainties of models. It is shown that by adding a prior probability to the observation probability, fewer experiments are required for inference. This is specifically useful in model development for difficult-to-machine alloys where high wear rate lowers the cardinality of the dataset. The combined Gibbs–Metropolis algorithm as a subset of MCMC method has been used in this work to quantify the uncertainty of the unknown parameters in a mechanistic tool wear model for end-milling of a difficult-to-machine Ni-based alloy. Maximum of 18% error and average error of 11% in the results show a good potential of this modeling in prediction of parameters in the presence of uncertainties when limited experiments are available.
[DOI: 10.1115/1.4033041]

Keywords: nickel-based alloys, uncertainty, Markov Chain Monte Carlo, milling, wear, Bayesian

1 Introduction

Understanding tool wear is highly important since it causes 20% of machining downtime [1]. A worn out tool deteriorates surface finish and dimensional accuracy and is eventually detrimental to machine health; this is an especially important consideration for abrasive, thermally insulative alloys with abnormally high-specific cutting energy, known as difficult-to-machine materials. Studying tool wear in different manufacturing processes goes back over a half century [2]. Since that time, several researchers have proposed various wear models to predict tool wear. However, the model parameters were generated mainly under certain experimental conditions, i.e., for 2D orthogonal cutting operation with specific tool geometry. Therefore, those models cannot be used for a wide range of materials, tool geometry, or cutting conditions. In addition to that, several uncontrollable factors in machining such as variation in material properties, existence of residual stress, tool runouts, and tolerances affect the dynamic interaction between tool and workpiece. Consequently, different dynamics for progressive tool wear can be observed. To address these unknown and uncontrollable factors, a probability-based approach can be utilized to quantify them as uncertainties in model parameters. In this way, parameters will be described with probability density functions and not deterministic values.

MCMC methods as one of the probabilistic approaches for parameter inference have gained more attraction recently in machining. Mehta et al. used this method for parameter identification in cutting force and spindle idle power model in turning [3,4]. In another effort, Karandikar et al. compared the performance of two MCMC methods, namely, mesh grid and Metropolis algorithms for inference on Taylor's tool life model parameters in milling [5,6]. Mesh grid algorithm is easy to implement as a brute force technique, but has a high computational cost in high dimensional space. Therefore, it suffers from a curse of dimensionality [7]. On the other hand, Metropolis algorithm is more capable to handle computation in high dimensional domains [7,8]. Bayesian inference was also investigated by Long et al. for studying the effect of flank wear and crater wear on 2D slip line force model in cutting CK45 steel [9]. Probabilistic modeling also influenced the tool wear condition monitoring area. The Kalman filter, particle filter, Bayesian networks, and hidden Markov models (HMMs) are the main monitoring and diagnosing probabilistic tools available. The Kalman and particle filters were specifically used for progressive tool wear estimation of γ'-strengthened alloys [10,11]. Akhavan Niaki et al. showed that the particle filter outperforms the Kalman filter due to its robustness for selection of non-Gaussian probability functions for tool wear state or model parameters [11]. Bayesian networks are another probabilistic tool, which can be used for diagnosing/classifying variation sources in the process. The main idea behind these networks is to discover the underlying conditional probabilities between the potential parameters affecting the outputs. Dey and Stori developed a Bayesian belief network to diagnose the root cause of process variation in sequential machining operation while considering tool wear as one of the sources of variation [12]. While Bayesian networks are proven practical in many decision-making applications, they are considered as static models. As an alternative, dynamic Bayesian networks known as HMMs were developed for classification of sequential patterns in speech signals [13] and then found their way into many different applications including machining. HMMs are successfully used for tool wear classification [14,15], but they are less informative than continuous tracking algorithms such as the Kalman or particle filters since the progress of wear cannot be estimated.

Manuscript received August 19, 2015; final manuscript received March 10, 2016; published online April 15, 2016. Assoc. Editor: Radu Pavel.

While probabilistic monitoring approaches have been widely studied in the literature for conventionally used materials in industry such as steel or aluminum, there are few works dedicated to the hard-to-machine alloys, specifically nickel-based alloys. Reviewing the recent research articles [16–18] demonstrates the lack of knowledge about proper model and strategies for monitoring wear condition for these materials. Nickel-based alloys are known to maintain their strength at elevated temperatures. Due to this unique characteristic and their corrosion and creep resistance, they are widely used in high-end industries. Nevertheless, because of high strength, poor thermal conductivity, and rapid work hardening of these materials, the tool wear rate of inserts during the machining process is relatively high and more unpredictable as compared to other conventional materials. This high wear rate makes establishing an accurate tool wear model a challenging task because only a limited number of experiments can be completed before the tool fails. This is where the role of quantifying uncertainties for the model parameters becomes important. While running tests with very mild cutting conditions can increase the number of available data for model establishment, this could be time-consuming and expensive in an industrial domain. Moreover, excessive wear causes surface tearing, high roughness, and induces significant tensile residual stress on the surface, which eventually leads to fatigue in service, reworking parts, or scrapping expensive materials (a cost comparison is given in Fig. 1) and lowers productivity rate. Unlike the conventionally used deterministic approaches (e.g., regression analysis), which only operate on the collected data, the power of probabilistic data analysis in the form of MCMC methods lies in the fact that due to incorporating initial knowledge to the current knowledge, fewer experiments are required for parameter inference and model establishment. Therefore, this method has the potential to be used for hard-to-machine alloys which exhibit high wear rate and cost.

The objective of this work is to investigate the performance of the Gibbs–Metropolis algorithm as one of the MCMC methods for parameter calibration in mechanistic tool wear model in milling a γ'-strengthened hard-to-machine alloys when a limited number of experiments is available. The organization of this work is as follows: The theoretical background of MCMC including Gibbs and Metropolis algorithms is described in Sec. 2. In Sec. 3, the selected mechanistic tool wear model and unknown parameters are identified. The experimental setup is described in Sec. 4. Implementation of MCMC methods is discussed in Sec. 5, and results and conclusions are provided in Secs. 6 and 7.

2 Theoretical Background

2.1 Bayes Rule. Bayesian data analysis is a powerful tool used for statistical inference. Thomas Bayes introduced the basic formulation, known as the Bayes rule, in the 18th century [19]. According to the Bayes rule, the probability of an event θ is derived by multiplying the initial belief (or previous knowledge of analyst, denoted as the *prior probability*) with the conditional likelihood function $p(y/\theta)$ as in Eqs. (1) and (2), where $p(\theta/y)$ is a *posterior probability* of event θ, $p(\theta)$ is the initial belief, and $p(y)$ is a marginal distribution

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \tag{1}$$

$$p(y) = \int p(y|\theta)p(\theta)d\theta \tag{2}$$

Assuming independent and identically distributed (i.i.d.) observations, the likelihood function, $p(y/\theta)$, is simplified as the product of observation probabilities as in Eq. (3). In many cases, finding a closed-form solution for marginal distribution is somewhat tedious or even impossible [20], so it is convenient to assume $p(y)$ as a normalizing constant which simplifies Eqs. (1)–(4). Although finding the closed-form solution of the posterior probability distribution $p(\theta/y)$ is possible, in many cases numerical approximations such as MCMC methods (e.g., Gibbs sampler or Metropolis algorithm) are proposed which simply generate samples from the posterior probability instead of calculating it

$$p(Y|\theta) = P(y_1, ..., y_2|\theta) = \prod_{i=1}^{n} p(y_i|\theta) \tag{3}$$

$$p(\theta|Y) \propto p(Y|\theta)p(\theta) \tag{4}$$

2.2 Gibbs Sampler. The Gibbs sampler is proposed as one of the MCMC algorithms to sample from a posterior distribution when a closed-from solution of posterior probability distribution is intractable [21]. To implement the Gibbs sampler, the closed-form solution for full conditional probability distribution of each parameter given all the remaining parameters is required. To illustrate how Gibbs sampler works, one needs to consider a linear model as in Eq. (5), where $Y\epsilon\{y_1,\ldots,y_n\}$ is the set of observations, β_i is an unknown coefficient, $X=[x_1,\ldots,x_i]^T$ is a set of known variables, and ε is a measurement error, which is assumed to be normally distributed with zero mean and unknown variance σ^2. In this case, there are two unknown values, β_i and σ^2, that should be identified in Bayesian framework

$$Y = \beta_i^T X + \varepsilon \tag{5}$$

The joint posterior probability density of the unknowns can be written as Eq. (6). The term $p(\beta_i/\sigma^2, x_i, y_1,\ldots,y_n)$ is called the full conditional of β_i, and the term $p(\sigma^2/x_i, y_1,\ldots,y_n)$ is called the marginal distribution of σ^2, where finding its closed-form solution is tedious except with some special assumptions. However, the Gibbs sampler states that if the full conditional of unknown parameters, β_i and σ^2, are known, samples taken from them belong to their joint posterior distribution. The full conditional of β_i can be written as Eq. (7), where $p(\beta_i)$ is an initial belief for unknown coefficients with mean β_0 and variance Σ_0

$$p(\beta_i, \sigma^2|X, y_1, \ldots y_n) \propto p(\beta_i|\sigma^2, X, y_1, \ldots y_n)p(\sigma^2|X, y_1, \ldots y_n) \tag{6}$$

$$\begin{aligned} p(\beta_i|\sigma^2, X, y_1, \ldots y_n) &\propto p(y_1, \ldots y_n|\beta_i, \sigma^2, X)p(\beta_i|\sigma^2, X) \\ &= p(y_1, \ldots y_n|\beta_i, \sigma^2, X)p(\beta_i) \end{aligned} \tag{7}$$

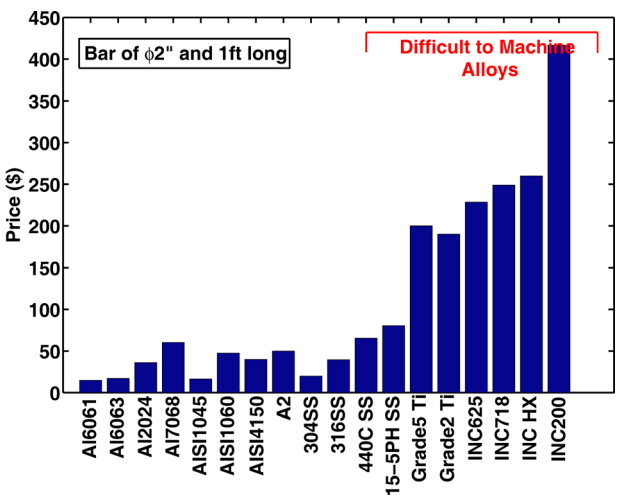

Fig. 1 Price comparison of different materials (source: McMaster-Carr)

Assuming i.i.d. observations and after some simplification, the full conditional of β_i is derived as a normal distribution with mean β_n and covariance Σ_n shown in the following equations:

$$p(\beta_i|\sigma^2, X, y_1, \ldots y_n) \propto N\left(\beta_n, \sum\nolimits_n\right) \tag{8}$$

$$\beta_n = \left(\sum\nolimits_0^{-1} + \frac{X^{\mathrm{T}}X}{\sigma^2}\right)^{-1}\left(\sum\nolimits_0^{-1}\beta_0 + \frac{X^{\mathrm{T}}Y}{\sigma^2}\right) \tag{9}$$

$$\sum\nolimits_n = \left(\sum\nolimits_0^{-1} + \frac{X^{\mathrm{T}}X}{\sigma^2}\right)^{-1} \tag{10}$$

The next is to calculate the full conditional of σ^2. This can be written as Eq. (11), where $p(\sigma^2)$ is the initial belief of measurement variance. It has been shown by Hoff [20] that this distribution can be considered as an inverse-gamma distribution with ν_0 and σ_0^2 as a sample size and sample variance of the prior, respectively (Eq. (12))

$$\begin{aligned} p(\sigma^2|\beta_i, X, y_1, \ldots y_n) &\propto p(y_1, \ldots y_n|\beta_i, \sigma^2, X)p(\sigma^2|\beta_i, X) \\ &= p(y_1, \ldots y_n|\beta_i, \sigma^2, X)p(\sigma^2) \end{aligned} \tag{11}$$

$$p(\sigma^2) \propto \mathrm{IG}\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right) \tag{12}$$

By substituting Eq. (12) into Eq. (11), the full conditional of measurement error variance is calculated as Eq. (13), where SSE is a sum of squared errors equivalent to $\sum_{i=1}^{n}\left(Y_i - \beta^{\mathrm{T}}X\right)^2$, and n is a number of observations. Therefore, the full conditional probability distribution of unknown parameters in linear systems is available and the Gibbs sampler can be easily used to draw samples to characterize the posterior distribution of parameters. This algorithm is described in Fig. 2

$$\begin{aligned} p(\sigma^2|\theta, y_1, \ldots y_n) &\propto \left((\sigma^2)^{\frac{n}{2}}\exp\left\{-\sigma^2\sum_{i=1}^{n}(y_i-\theta)^2/2\right\}\right) \times \left((\sigma^2)\frac{\nu_0}{2-1}\exp\left\{-\sigma^2\nu_0\sigma_0^2/2\right\}\right) \\ &= (\sigma^2)^{\frac{\nu_0+n}{2-1}} \times \exp\left\{-\sigma^2 \times \left[\nu_0\sigma_0^2 + \sum(y_i-\theta)^2\right]\Big/2\right\} = \mathrm{IG}\left(\frac{\nu_0+n}{2}, \frac{\nu_0\sigma_0^2 + \mathrm{SSE}}{2}\right) \end{aligned} \tag{13}$$

2.3 Metropolis Algorithm. The Gibbs sampler is an easy-to-implement and practical algorithm in the case of linear models, while for nonlinear models or in cases where closed-form solution for the full conditional distribution of unknown parameters is not available, it does not work properly. Another technique was proposed by Nicholas Metropolis in 1953, where a distribution called *proposal density* is used to approximate the posterior distribution of parameters [20]. Since the proposal density does not fully capture the features of the posterior distribution, an acceptance–rejection method is implemented to reject the samples that are generated from regions with lower probability of occurrence. Unlike the Gibbs sampler that accepts all the samples, the Metropolis algorithm just accepts a portion of them. The step-by-step guide for Metropolis algorithm is shown in Fig. 3.

In practical applications, a symmetric probability density (i.e., a normal distribution with zero mean and arbitrary variance δ^2) is used as the proposal density function. Even though Lynch reported that finding the variance δ^2 is more of an art than science and it depends on the experience level of the user [22], there exist several studies in the literature on techniques for finding an optimal proposal density variance [23–25]. The proper choice of a proposal density function plays a critical role in the acceptance rate of candidate points. If a very small variance is chosen for δ^2, it takes a longer time for the Markov chain to converge to the true values and if a very large variance is chosen the rejection rate of drawn samples increases and this affects the efficiency of the chain. It is reported that the 25–35% acceptance rate can be considered as appropriate for Markov chain [20]. In this work, a method of Solonen [26] based on calculating the Jacobian matrix was used to come up with the proposal density variance. Note that in addition to Metropolis algorithm, there are many other algorithms that are developed such as Metropolis–Hasting algorithm, delayed rejection algorithm, or adaptive Metropolis [20,25,27] to increase the performance of sampling procedure. However,

Gibbs Sampler Algorithm

- (0) Find the full condition of unknowns
- (1) start with k=1
  - Draw a sample β_i^k from full condition of β_i:
  $$\beta_i^k \sim N(\beta_n, \Sigma_n)$$
- (2) Use the drawn sample β_i^k
  - Calculate sum of squared error:
  $$SSE = \sum_{i=1}^{n}\left(Y_i - \beta_i^{k^T}X_i\right)^2$$
  - Draw a sample σ_k^2 from the full conditional of σ^2:
  $$\sigma_k^2 \sim IG\left(\frac{v_0+n}{2}, \frac{v_0\sigma_0^2+SSE}{2}\right)$$
- (3) Add (+1) to k and GO to line (1)

Fig. 2 Gibbs sampler algorithm

Metropolis Algorithm

- (0) Select the proposal density function
- (1) Select the starting point as β_i^1 , σ_1^2
- (2) FOR k=1:N
  - Select a candidate point q_k from proposal density:
  $$q_k \sim N(0, \delta^2)$$
  - Calculate a candidate point β_i^*:
  $$\beta_i^* = \beta_i^k + q_k$$
- (3) Calculate the r ratio
$$r = \frac{p(\beta_i^*|\sigma^2, X, y_1, \ldots, y_n)}{p(\beta_i^k|\sigma^2, X, y_1, \ldots, y_n)} =$$
$$\exp\left\{\frac{-1}{2}\left[\frac{SSE(\beta_i^*) - SSE(\beta_i^s)}{\sigma_s^2} + N(\beta_0, \Sigma_0)_{\beta_i^*} - N(\beta_0, \Sigma_0)_{\beta_i^k}\right]\right\}$$
- (4) Select a point U from the uniform distribution
$$U \sim uniform[0,1]$$
- (5) Find the acceptance ratio
$$\alpha = \min\{U, r\}$$
- (6) IF $U<\alpha$
  - Accept the candidate point $\beta_i^{k+1} = \beta_i^*$

  OTHERWISE
  - Reject the candidate point $\beta_i^{k+1} = \beta_i^k$
- (8) END IF
- (9) END FOR

Fig. 3 Metropolis algorithm

exploring these algorithms is out of the scope of this work. The last part of this section explains the Solonen algorithm to approximate the proposal density variance δ^2. Solonen showed that an optimal value of the covariance matrix δ^2 can be approximated as Eq. (14), where MSSE is the minimum SSE derived after plugging in the optimal value of β_i that minimizes squared error function and J is a Jacobian matrix of outputs with respect to the unknowns as in Eq. (15) [26]

$$\delta^2 = (J^{\mathrm{T}}J)^{-1}\mathrm{MSSE} \tag{14}$$

$$J = \left[\frac{\partial Y}{\partial \beta_1^{\mathrm{opt}}} \quad \cdots \quad \frac{\partial Y}{\partial \beta_i^{\mathrm{opt}}}\right] \tag{15}$$

3 Mechanistic Model of Tool Wear

Altintas and Yellowley showed that the tangential force in milling can be described as Eq. (16) [28], where K and c are the constants, $\bar{h}$ is the mean chip thickness, a_p is the depth of cut, f is the feedrate, and φ is the instantaneous angle of rotation (Fig. 4)

$$F_t = K\bar{h}^c a_p f \sin\varphi \tag{16}$$

It was shown that with an increase in tool wear, the magnitude of the cutting force increases as well [30]. Waldrof et al. showed that the change in magnitude of tangential force (F_t) is a function of material hardness (H), friction coefficient (μ), tool flank wear (Vb), and tool wear length (s) as in Eq. (17) [31]. Shao et al. derived average cutting power in milling [29] which is simplified in this work as Eq. (18), with the assumption of constant depth of cut. In Eq. (18), parameter P represents the cutting power, N is the cutting speed, and K_1, K_2, and K_3 are the unknown parameters that need to be identified

$$F_t^{\mathrm{wear}} = \mu VbHs \tag{17}$$

$$P = K_1 N f^{K_2} + K_3 NVb \tag{18}$$

4 Experimental Setup

The material used for this experimental study is a hard-to-machine nickel-based material known as Rene-108 (R-108) [26,27]. An OKUMA GENOS M460-VE three-axis computer numerical control machine was used to end mill (in the down-milling direction) rectangular blocks of size $60 \times 80 \times 25$ mm, using a 5% water-soluble coolant. A two-flute indexable tool holder with a diameter of 15.875 mm was used, and the width of cut was chosen to be 9.5 mm that corresponds to 60% tool engagement, as this was the maximum manufacturer recommendation for the particular tool holder. Full length of the blocks (60 mm as the cutting distance) was utilized for machining. At the chosen width of cut, eight tests were conducted on the block: four tests with two replications. Four additional tests were also conducted to cover the full range of cutting conditions for validating the results. The depth of cut for each pass is kept constant at 0.5 mm, and the cutting speed and feedrate were changed as excitation factors for parameter identification. A tool monitoring adaptive controller from Caron Engineering was used for measuring spindle current to monitor spindle power consumption in real time. However, due to the low sampling frequency of the commercial software (~50 Hz), a national instruments NI-cRIO9103 data-acquisition (DAQ) device was programed to capture the data at a higher rate. To measure spindle power in high sampling frequency, the output of the transducer (Fig. 5) was directed to the NI9215 analog input module mounted on NI-cRIO9103 chassis programed with LABVIEW. Data were collected in voltage at a sampling rate of 10.24 kHz.

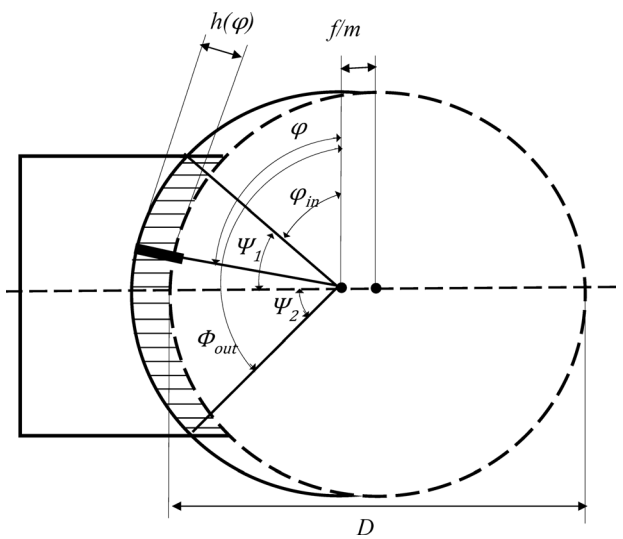

Fig. 4 Milling schematic [29]

Tests for R-108 were designed in a fashion that the effect of each parameter can be observed, therefore selecting a high and a low level of cutting speed and feedrate. Two levels of cutting speed at 25 and 50 m/min, two levels of feedrate at 0.1 and 0.2 mm/rev, and a constant depth of cut of 0.5 mm were used to identify the constants K_1, K_2, and K_3 in Eq. (18). These cutting parameters were selected at both their relatively mild values to show the behavior of the inserts under normal machining conditions, and at their relatively aggressive values to show the behavior of the inserts under high material removal rate conditions. Cutting parameters for validation tests were selected between the mild and aggressive values. Design of experiment (DoE) used in this work is shown in Table 1. Spindle power consumption was measured for each test. The change in spindle power consumption is shown in Fig. 6, the increase in the power illustrates the developing tool flank wear during the cutting process. The mean value of cutting power between 42 and 48 mm cutting distances (70–80% of total cutting distance) was selected as the average cutting power affected by tool flank wear at each test. One portion of this value is contributed to the power required to cut the workpiece (i.e., $K_1 N f^{K2}$), and another portion is due to the effect of tool wear on increasing cutting power magnitude (i.e., $K_3 NVb$).

Inserts used in this work were Sandvik Coromill (R390-11 T3 08 M-PM 1030). The 1030 grade is recommended by Sandvik for milling R-108 due to its resistance to material build-up on the cutting edge and plastic deformation [32]. Fresh unworn inserts were used for each test, and flank wear at the end of each pass on the bottom edge of each insert was measured using scanning electron microscope and average flank wear was calculated. Measured tool flank wear for tests 1–4 is shown in Fig. 7.

5 Inference on Model Parameters

The objective of this work is to identify unknown parameters K_1, K_2, and K_3 and measurement error variance σ^2 when a limited number of experiments exist. Due to the nonlinearity of Eq. (18), finding the full conditional distribution of the unknown parameters is impossible but the full conditional of measurement error variance (σ^2) is available as shown in Eq. (13). It is possible to use a combined Gibbs–Metropolis algorithm to characterize the posterior distribution of unknowns. In such case, the Metropolis algorithm is used to generate samples from unknown parameters K_1, K_2, and K_3, and Gibbs sampler is used to characterize the distribution of measurement error variance σ^2. In Fig. 8, the flow chart of the combined algorithms is shown. Since the full conditional of σ^2 is available, all the samples drawn by using the Gibbs algorithm were accepted automatically, but a rejection/acceptance method should be implemented for the samples generated by the Metropolis algorithm. The prior belief for unknown parameters is

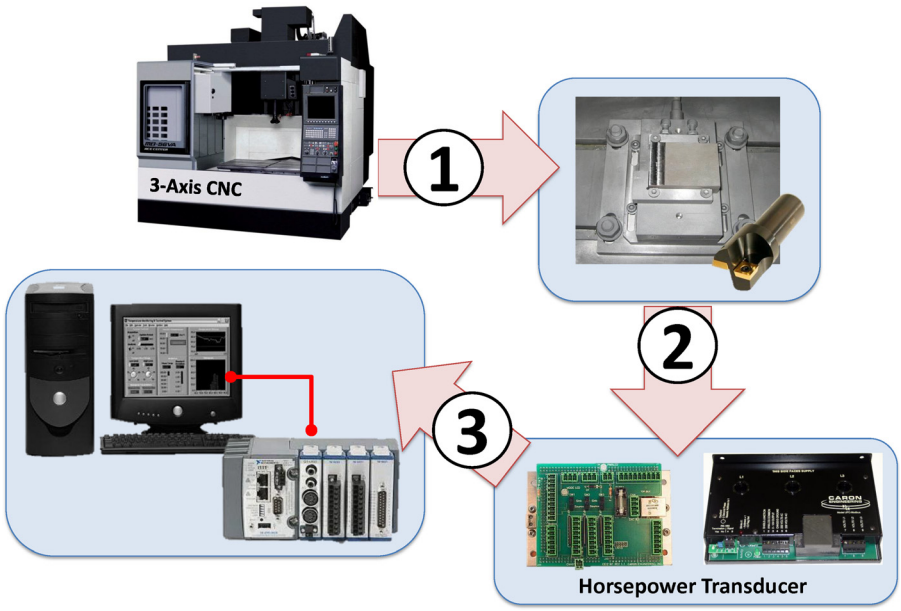

Fig. 5 Data acquisition with NI-cRIO9103

chosen as normal distribution with the mean of 1 and a large variance (Eq. (19)). The prior belief for measurement covariance is chosen as inverse-gamma function with $\nu_0 = 1$ and $\sigma_0^2 = 100$ (Eq. (20)). The choice of prior distribution is arbitrary but should be a rational guess since no previous knowledge was available about the parameters. A selection of prior distribution far from the true posterior can increase the convergence time of the Markov chain to find the true values

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \sim \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \tag{19}$$

$$\sigma^2 \sim \mathrm{IG}\left(\frac{1}{2}, \frac{2}{100^2}\right) \tag{20}$$

To avoid singularity of the covariance matrix, spindle power consumption is multiplied by 1000 so that K_1 and K_3 are in the same absolute range as K_2. The rest of this study is based on the normalized values for K_1 and K_3. The optimal value of unknown parameters was calculated from data in Table 1 using an unconstrained derivative-free optimization method. This method uses the Simplex search algorithm described in Ref. [33]. At each iteration, new points are generated around the simplex, and the points with the lowest output function are rejected. The process is repeated until the optimal points that minimize the output function are found

$$\begin{bmatrix} K_1 & K_2 & K_3 \end{bmatrix}^{\mathrm{T}}_{\mathrm{opt}} = \begin{bmatrix} 0.94 & 1.21 & 0.15 \end{bmatrix}^{\mathrm{T}} \tag{21}$$

Table 1 DoE for end-milling Rene-108

Test #	V_c (m/min)	f (mm/rev)	P ($\times 10^{-3}$ hp)	Vb (μm)
1	25	0.1	36	88
2	25	0.2	57	73
3	50	0.1	82	85
4	50	0.2	154	113
5	25	0.1	36	88
6	25	0.2	47	82
7	50	0.1	62	97
8	50	0.2	165	82

The Jacobian matrix was calculated using K^{opt} for each test and consequently the Solonen formula used for proposal density covariance matrix calculation. The optimal points K^{opt} were used for initializing the Markov chain as well. The total number of points in the chain was selected as $N = 2000$

$$J = \begin{bmatrix} N_1 f_1^{K_2^{\mathrm{opt}}} & K_1 N_1 f_1^{K_2^{\mathrm{opt}}} \log\left(K_2^{\mathrm{opt}}\right) & N_1 V B_1 \\ \dots & \dots & \dots \\ N_8 f_8^{K_2^{\mathrm{opt}}} & K_1 N_8 f_8^{K_2^{\mathrm{opt}}} \log\left(K_2^{\mathrm{opt}}\right) & N_8 V B_8 \end{bmatrix}_{8\times 3} \tag{22}$$

$$\delta^2 = \begin{bmatrix} 0.87 & 0.78 & 0.51 \\ 0.78 & 0.74 & 0.52 \\ 0.51 & 0.52 & 0.41 \end{bmatrix} \tag{23}$$

$$K_i^0 = K_i^{\mathrm{opt}} = \begin{bmatrix} 0.94 & 1.21 & 0.15 \end{bmatrix}^{\mathrm{T}} \; i \in \{1, 2, 3\} \tag{24}$$

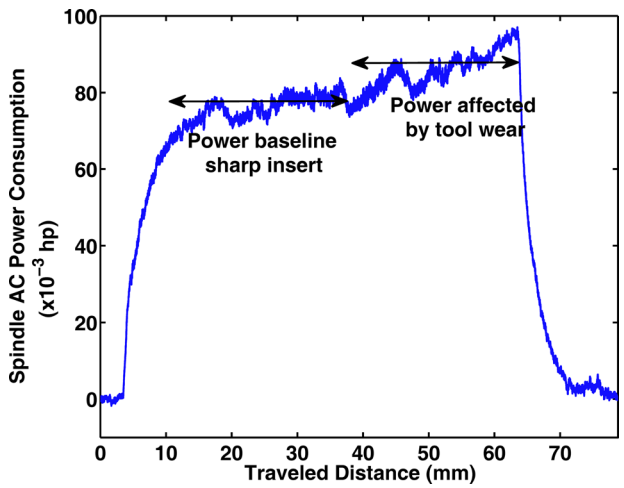

Fig. 6 Cutting power of test 3: $V_c = 50$ m/min and $f = 0.1$ mm/rev

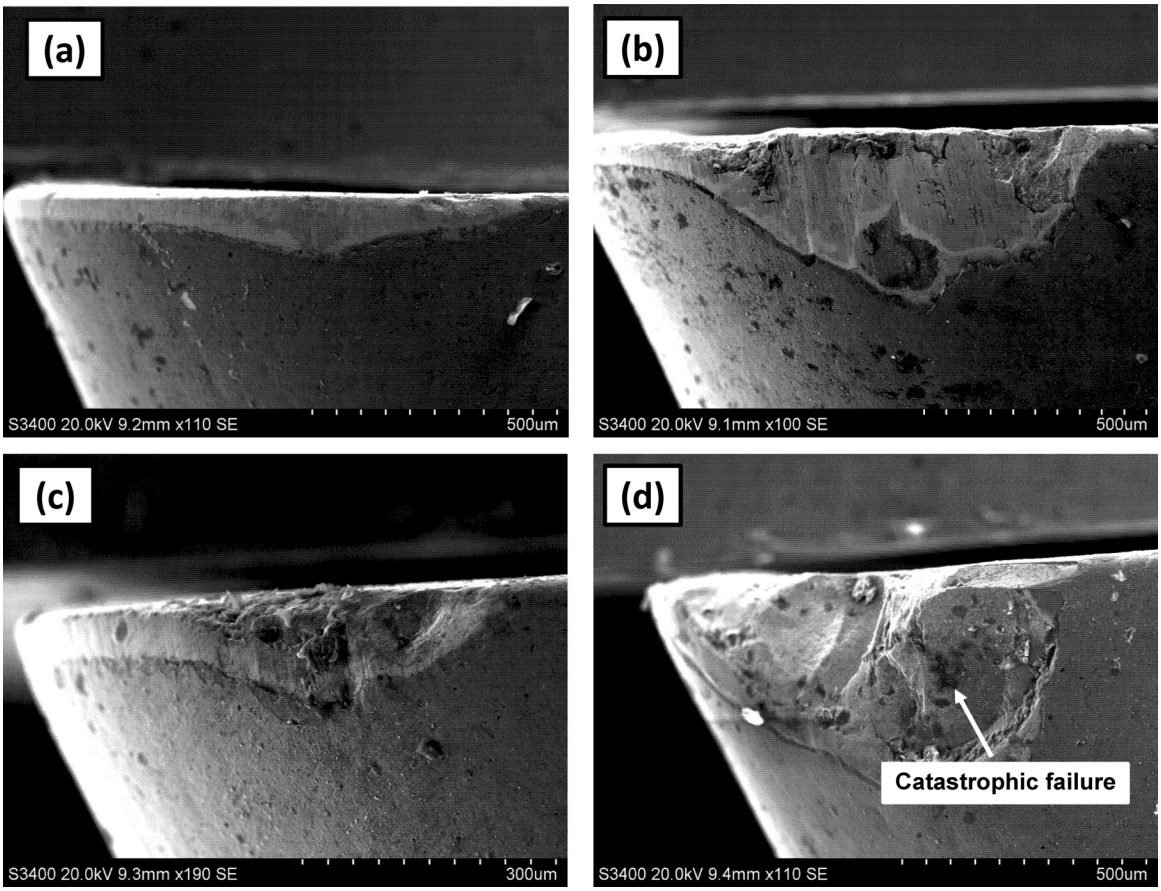

Fig. 7 Measured flank wear for tests 1–4: (*a*) test 1—$V_c = 25$ m/min and $f = 0.1$ mm/rev, (*b*) test 2—$V_c = 25$ m/min and $f = 0.2$ mm/rev, (*c*) test 3—$V_c = 50$ m/min and $f = 0.1$ mm/rev, and (*d*) test 4—$V_c = 50$ m/min and $f = 0.2$ mm/rev

6 Results

The MCMC trace plot for the parameters K_1, K_2, and K_3 following the procedure described in Sec. 5 with $N = 2000$ points is shown in Fig. 9(*a*). The first set of iterations is usually discarded as burn-in period to reduce the effect of initial errors at the start of the chain [20]. In this work, the first 20% of the iterations (400 points) were discarded as the burn-in period. After removing the first 400 points, the acceptance ratio was calculated as 19%. To assess the convergence of the Markov chain, the autocorrelation plot of samples shown in Fig. 9(*b*) was taken into consideration. In a converged chain, samples become uncorrelated to each other after some time lags. However, as it is shown in Fig. 9(*b*) samples are heavily correlated (more than 60% correlation with the lag of 20) which indicates that the Markov chain was not able to converge to the posterior density. Hoff stated that to improve the performance of Markov chain, the posterior variance of samples can be an efficient choice of proposal variance [20]. Therefore, to improve the current run (named as *pilot run*), the information

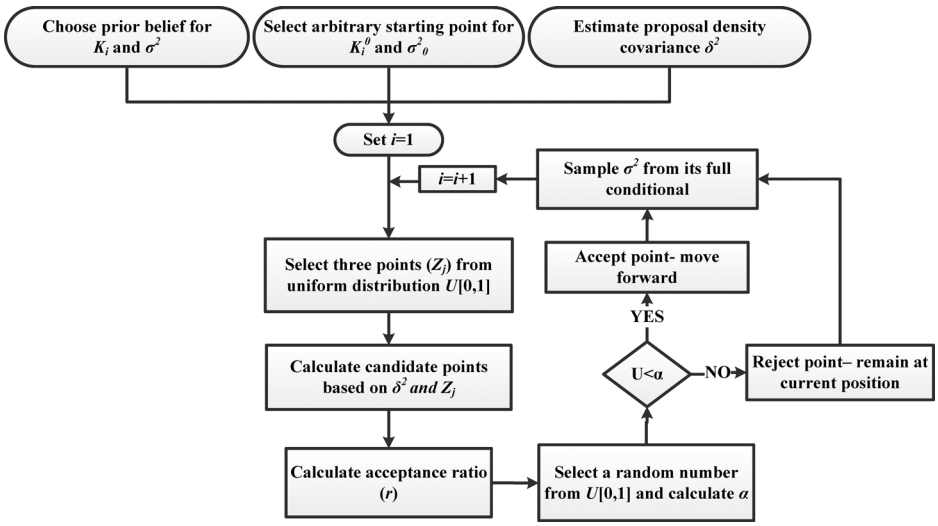

Fig. 8 Flowchart of combined Gibbs–Metropolis algorithm

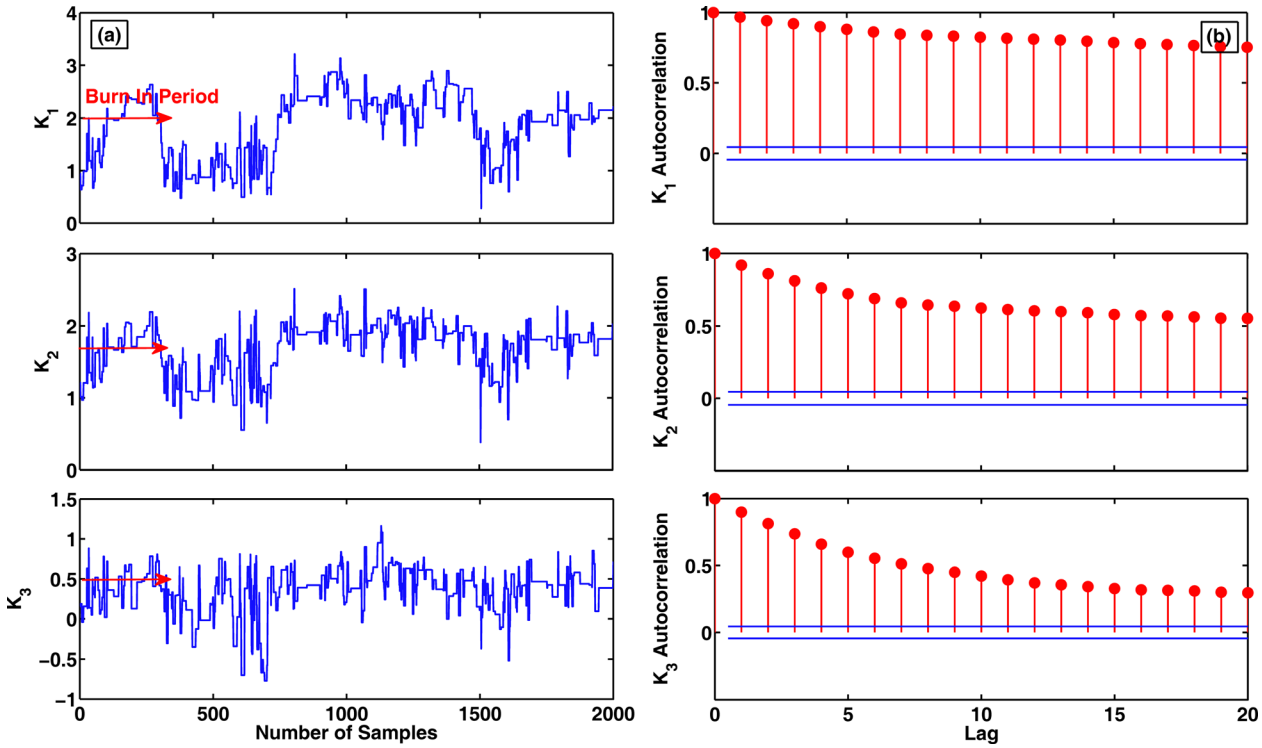

Fig. 9 Pilot run samples: (*a*) trace plot and (*b*) samples autocorrelation (diverged chain)

from the generated samples was used as the initial belief of parameters for a next MCMC run (named as *main run*). The mean and covariance matrices of the pilot run were extracted from the chain and implemented as a prior belief. The covariance matrix of parameters was also used as the proposal density covariance, and the final point of the chain was used as the initial point of the new chain. For the main run, the number of generated samples was selected as $N = 10{,}000$.

The trace plot and autocorrelation of samples are shown in Fig. 10, which reveals that after modifying the proposal density variance and initial prior the chain has converged and parameters are mixing well. After removing the first 20% of the samples

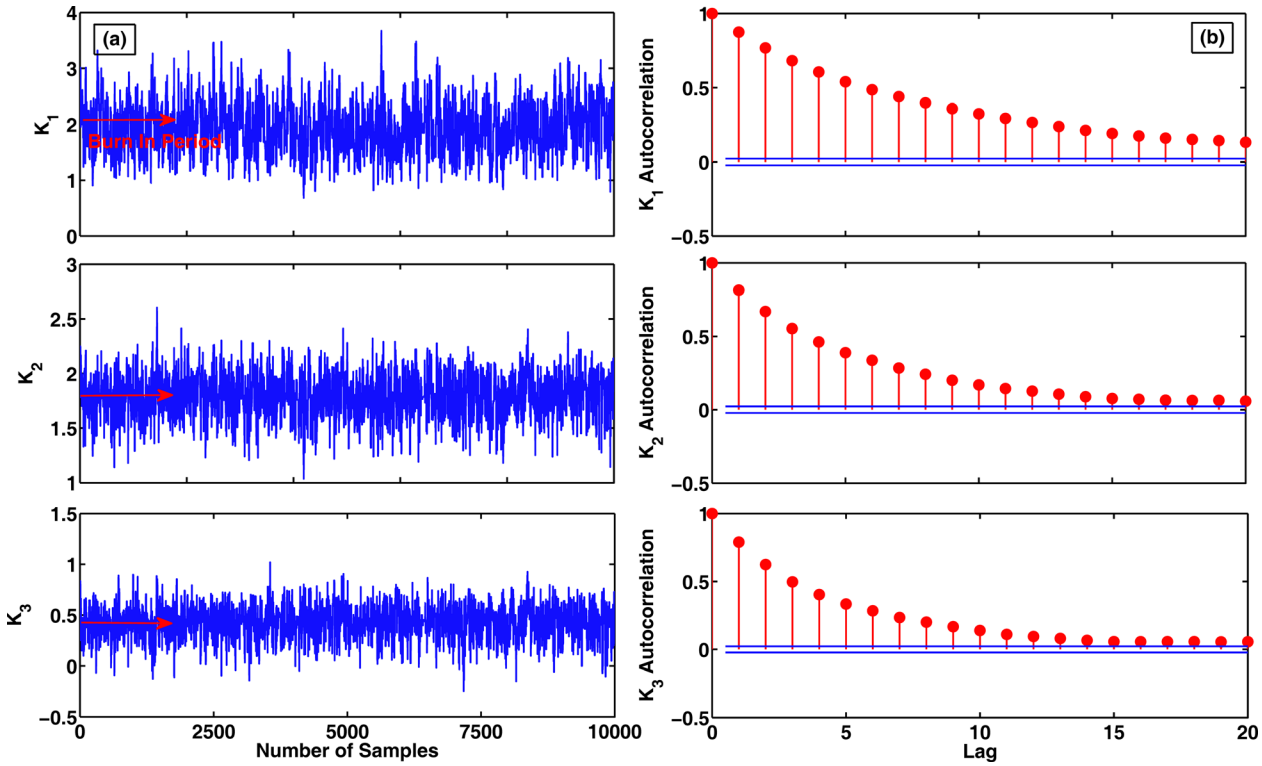

Fig. 10 Main run samples: (*a*) trace plot and (*b*) samples autocorrelation (converged chain)

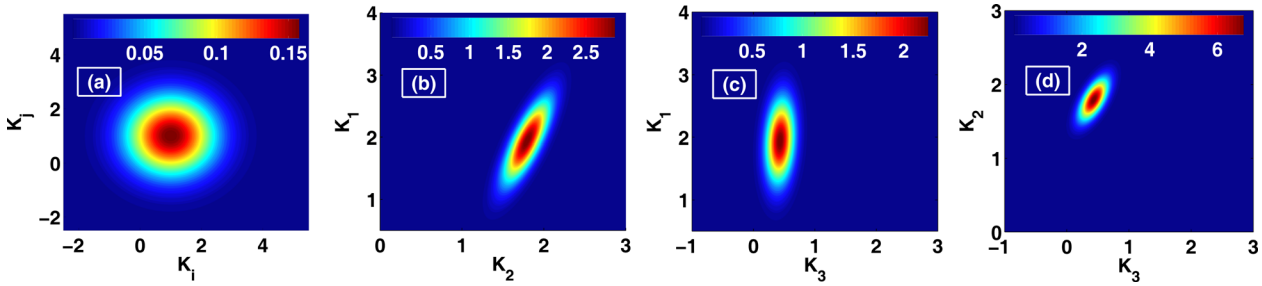

Fig. 11 Prior and posterior distributions after main run: (*a*) prior probability of K_i and K_j, (*b*) posterior probability of K_1 and K_2, (*c*) posterior probability of K_3 and K_1, and (*d*) posterior probability of K_3 and K_2

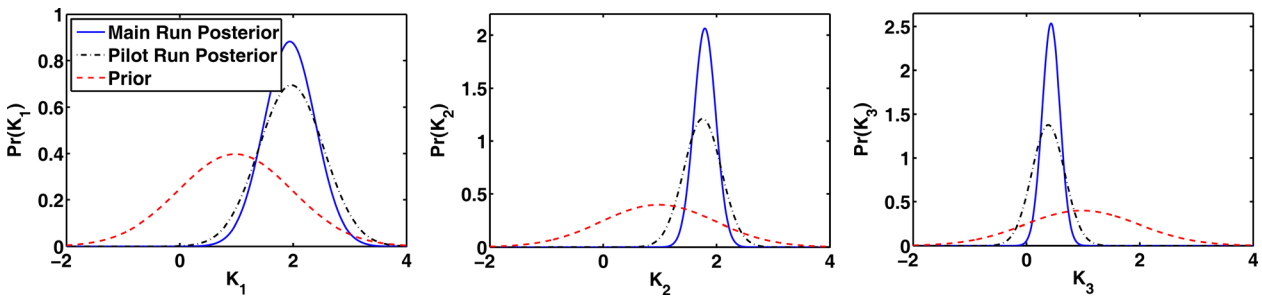

Fig. 12 Distribution of parameters for prior belief, pilot run, and main run (the *y*-axis is not normalized)

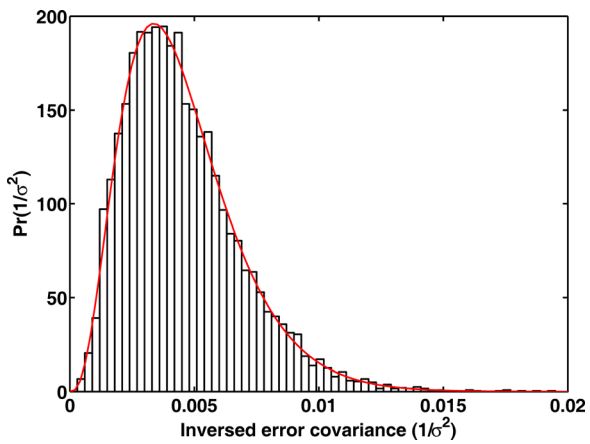

Fig. 13 Gamma distribution of the inverse of measurement error variance

(2000 points) as the burn-in period, a comparison of each pair of parameter's distribution is shown in Fig. 11, where contour (*a*) is the initial belief (prior distribution) from the pilot run, and contours (*b*)–(*d*) are the posterior belief of parameters. The multivariate posterior distribution of identified parameters is shown in Eq. (25). Figure 12 also demonstrates the improvement in the degree of uncertainties after each MCMC run. Analysis starts with an initial degree of uncertainties which is collected from previous available data or a rational guess. The initial belief as shown in Fig. 12 covers a wide range of possible values for unknown parameters (i.e., large variance, shown as dashed curve), however, by running the MCMC method and bringing new information, the range of possible values for unknown parameters shrinks and at the same time its probability distribution moves toward the true values of actual parameters (shown as solid and dotted dashed curves)

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \sim N\left(\begin{bmatrix} 1.94 \\ 1.79 \\ 0.43 \end{bmatrix}, \begin{bmatrix} 0.20 & 0.07 & 0.01 \\ 0.07 & 0.04 & 0.02 \\ 0.01 & 0.02 & 0.02 \end{bmatrix} \right) \tag{25}$$

Furthermore, the gamma distribution of the inverse of measurement error variance $(1/\sigma^2)$ generated from the Gibbs sampler is shown in Fig. 13, and the mode is chosen as the posterior value of measurement error variance $\sigma_N^2 = 51$. Now that the unknown parameters distribution is identified, the performance of the Bayesian inference can be evaluated using the validation tests. Table 2 shows the cutting parameters, spindle power consumption, and tool flank wear for each of the four new tests. It is fairly straightforward to find the posterior distribution of observations (posterior predictive distribution) as shown in Fig. 14.

The expectation is that the measured power be within the 95% confidence interval for each of the four tests. As shown in Fig. 15, the model is able to predict the measured power with good

Table 2 Validation tests cutting conditions, measured power, and tool flank wear

Test #	V_c (m/min)	f (mm/rev)	P ($\times 10^{-3}$ hp)	Vb (μm)
1	30	0.18	64	84
2	35	0.15	65	73
3	40	0.12	56	92
4	45	0.05	41	80

Posterior Predictive Distribution

- (1) FOR k=1:N
  - Sample K_i^k from posterior distribution
  - Sample σ_k^2 from its posterior distribution
  - Sample ε_k as measurement error

  $\varepsilon_k \sim N(0, \sigma_k^2)$
- (2) Calculate the power output

  $P_k = K_1^k N f^{K_2^k} + K_3^k NVB + \varepsilon_k$
- (3) END FOR

Fig. 14 Posterior predictive distribution algorithm

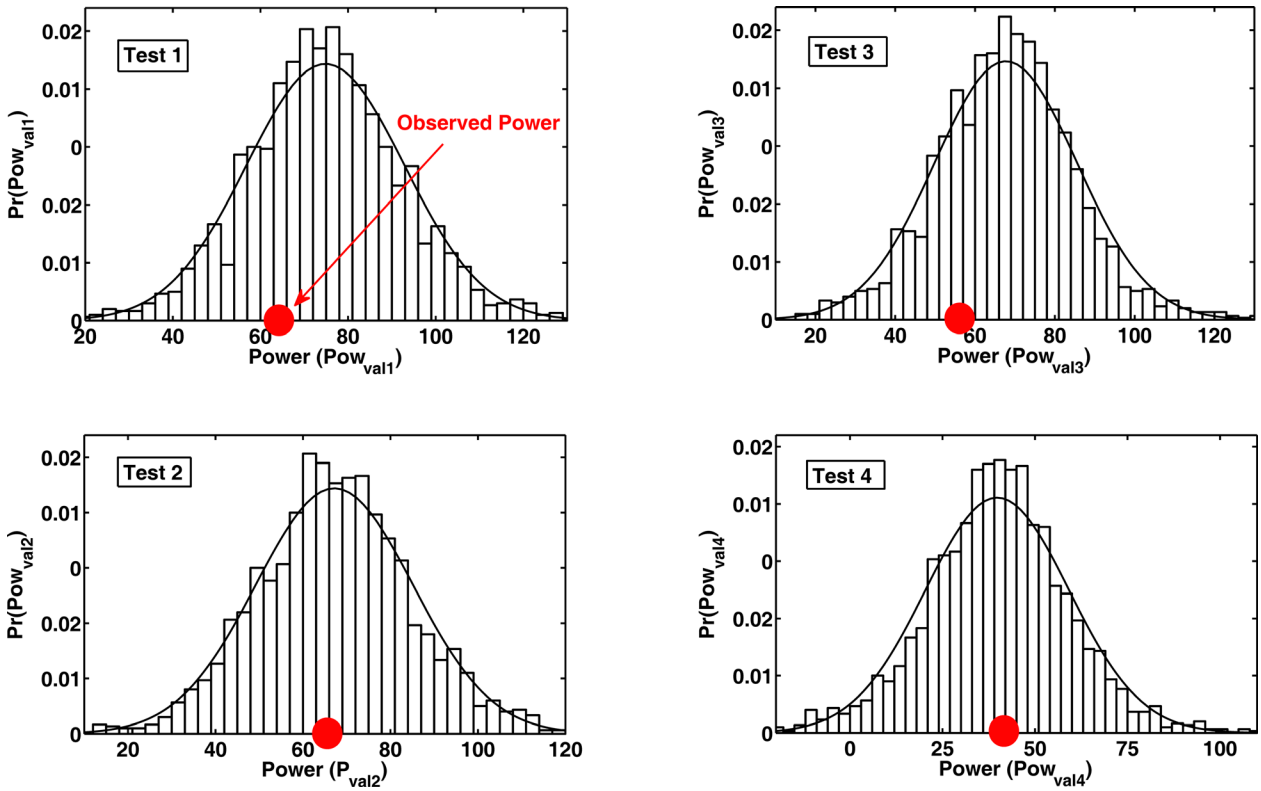

Fig. 15 Posterior predictive distribution, measured spindle power (-o- sign)—validation tests 1–4

Table 3 Percentage error of prediction and measurement

Test #	$P_{\text{measurement}}$ ($\times 10^{-3}$ hp)	$P_{\text{predicted}}$ ($\times 10^{-3}$ hp)	Error (%)
1	64	75	17
2	65	68	4
3	56	66	18
4	41	39	5

accuracy. Percent error between the measurement and the prediction mean for each test is compared in Table 3. Maximum 18% error indicates that the algorithm is capable to predict spindle power consumptions with good degree of accuracy, which implies validity of identified parameters. As demonstrated in Fig. 15, at test 4 some of samples of the posterior predictive distribution are generated in the negative area which is physically impossible. The nature of a large variance in power prediction in test 4 is due to the limited number of available experiments for establishing the model. Limited experiments cause a significantly large measurement error variance which produces a large variation in the output of the model. However, considering the mean of output prediction in comparison to the measured power, the error is in acceptable range (only 5% in test 4).

6.1 Elimination of Prior Belief's Effect. To eliminate and reduce the influence of the initial belief distribution, MCMC was run 50 times to investigate the convergence of unknown parameters mean and variance. On each run, the final points of the chain

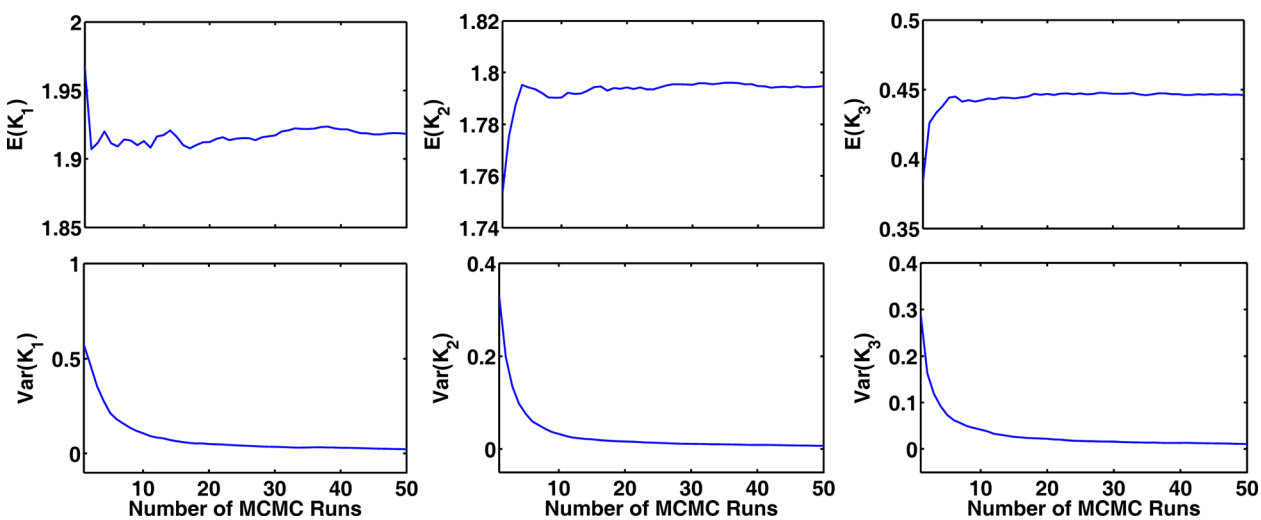

Fig. 16 Evolution of mean (E[K_i]) and variance (Var[K_i]) of parameters after 50 runs

were used as the initial points of a new chain and covariance of samples as a new proposal density covariance. As shown in Fig. 16, the evolution of mean and variance of each variable converged to a certain value. Variance (which represents uncertainties) never reaches zero which confirms the fact that uncertainties always exist but can be reduced. Note that although the variation in unknown parameters reduced significantly, this does not affect the measurement error variance since the same information from the experiments was used in the simulations. The reduction in measurement error variance was only 17%.

7 Conclusions

This work demonstrates the use of the Bayesian parameter identification in mechanistic model of the tool wear. The focus of this work is on hard-to-machine alloys which have been shown to have a poor machinability due to several reasons such as low thermal conductivity and high strength. High tool wear rate while machining these materials is a major challenge in industrial applications since it limits the productivity rate. In addition, excessive tool wear can damage surface quality and causes undesired residual stress beneath the machined surface. The aforementioned challenges also limit the available number of experiments for establishing accurate models since the cutting tool wears out quickly, and the materials are expensive. To have the accurate estimation of unknown model parameters, a Bayesian model establishment method, i.e., a combined technique of MCMC, was used in this work. This technique can be used for parameter identification when limited experiments are available, a feature that is beneficial and cost effective in studying hard-to-machine alloys, as well as a number of other physical phenomena. The main conclusions are summarized below:

- The combined Gibbs–Metropolis algorithm was formulated for prediction of unknown parameters in a nonlinear mechanistic cutting power model for milling of R-108. The Metropolis algorithm with symmetric proposal density was used for predicting the model parameters, while the Gibbs sampler was used for updating measurement error variance.
- A DoE study with mild and aggressive cutting conditions was used along with high-frequency DAQ to capture a wide range of tool wear and spindle power consumption. The performance of the algorithm improved significantly after using data from the first run of MCMC as prior belief for the second run. Predicted parameters were successful in estimating the spindle power consumption with a maximum 18% error and average error of 8.5%.

The results of this study can be used for monitoring tool wear while measuring cutting power by simply substituting the identified distribution of parameters into the mechanistic model as a measurement model in state-space format and then using the online Bayesian techniques such as Kalman filter or particle filter for estimating or predicting the tool wear during machining.

Acknowledgment

The authors wish to thank the National Science Foundation for support of this work under Grant No. 0954318.

References

[1] Vallejo, A. G., Jr., Nolazco-Flores, J. A., Morales-Menéndez, R., Sucar, L. E., and Rodríguez, C. A., 2005, "Tool-Wear Monitoring Based on Continuous Hidden Markov Models," *Progress in Pattern Recognition, Image Analysis and Applications*, Springer, Berlin, Heidelberg, pp. 880–890.
[2] Yen, Y., Söhner, J., Lilly, B., and Altan, T., 2004, "Estimation of Tool Wear in Orthogonal Cutting Using the Finite Element Analysis," J. Mater. Process. Technol., **146**(1), pp. 82–91.
[3] Mehta, P., Kuttolamadom, M., and Mears, L., 2012, "Machining Process Power Monitoring: Bayesian Update of Machining Power Model," ASME Paper No. MSEC2012-7277.
[4] Mehta, P., and Mears, L., 2013, "Model Learning in a Multistage Machining Process: Online Identification of Force Coefficients and Model Use in the Manufacturing Enterprise," ASME Paper No. MSEC2013-1144.
[5] Karandikar, J. M., Abbas, A. E., and Schmitz, T. L., 2014, "Tool Life Prediction Using Bayesian Updating—Part 1: Milling Tool Life Model Using a Discrete Grid Method," Precis. Eng., **38**(1), pp. 9–17.
[6] Karandikar, J. M., Abbas, A. E., and Schmitz, T. L., 2014, "Tool Life Prediction Using Bayesian Updating—Part 2: Turning Tool Life Using a Markov Chain Monte Carlo Approach," Precis. Eng., **38**(1), pp. 18–27.
[7] Daum, F., 2005, "Nonlinear Filters: Beyond the Kalman Filter," IEEE Aerosp. Electron. Syst. Mag., **20**(8), pp. 57–69.
[8] Doucet, A., De Freitas, N., and Gordon, N., 2001, *An Introduction to Sequential Monte Carlo Methods*, Springer, New York.
[9] Long, Y., Huang, Y., and Sun, X., 2010, "Combined Effects of Flank and Crater Wear on Cutting Force Modeling in Orthogonal Machining—Part II: Bayesian Approach-Based Model Validation," Mach. Sci. Technol., **14**(1), pp. 24–42.
[10] Henderson, A., 2012, "Updated Force Model for Milling Nickel-Based Superalloys," Ph.D. dissertation, Clemson University, Clemson, SC.
[11] Akhavan Niaki, F., Ulutan, D., and Mears, L., 2015, "Stochastic Tool Wear Assessment in Milling Difficult to Machine Alloys," Int. J. Mechatron. Manuf. Syst., **8**(3–4), pp. 134–159.
[12] Dey, S., and Stori, J., 2005, "A Bayesian Network Approach to Root Cause Diagnosis of Process Variations," Int. J. Mach. Tools Manuf., **45**(1), pp. 75–91.
[13] Rabiner, L., 1989, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. IEEE, **77**(2), pp. 257–286.
[14] Ertunc, H. M., Loparo, K. A., and Ocak, H., 2001, "Tool Wear Condition Monitoring in Drilling Operations Using Hidden Markov Models (HMMs)," Int. J. Mach. Tools Manuf., **41**(9), pp. 1363–1384.
[15] Zhu, K., Wong, Y. S., and Hong, G. S., 2009, "Multi-Category Micro-Milling Tool Wear Monitoring With Continuous Hidden Markov Models," Mech. Syst. Signal Process., **23**(2), pp. 547–560.
[16] Zhu, D., Zhang, X., and Ding, H., 2013, "Tool Wear Characteristics in Machining of Nickel-Based Superalloys," Int. J. Mach. Tools Manuf., **64**, pp. 60–77.
[17] Akhtar, W., Sun, J., Sun, P., Chen, W., and Saleem, Z., 2014, "Tool Wear Mechanisms in the Machining of Nickel Based Super-Alloys: A Review," Front. Mech. Eng., **9**(2), pp. 106–119.
[18] Thakur, A., and Gangopadhyay, S., 2016, "State-of-the-Art in Surface Integrity in Machining of Nickel-Based Super Alloys," Int. J. Mach. Tools Manuf., **100**, pp. 25–54.
[19] Gilks, W. R., Richardson, S., and Spiegelhalter, D., 1995, *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London.
[20] Hoff, P. D., 2009, *A First Course in Bayesian Statistical Methods*, Springer, New York.
[21] Andrieu, C., De Freitas, N., Doucet, A., and Jordan, M. I., 2003, "An Introduction to MCMC for Machine Learning," Mach. Learn., **50**(1–2), pp. 5–43.
[22] Lynch, S. M., 2007, *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*, Springer, New York.
[23] Haario, H., Saksman, E., and Tamminen, J., 1999, "Adaptive Proposal Distribution for Random Walk Metropolis Algorithm," Comput. Stat., **14**(3), pp. 375–396.
[24] Rosenthal, J. S., 2011, "Optimal Proposal Distributions and Adaptive MCMC," *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC, Boca Raton, FL, pp. 93–112.
[25] Haario, H., Saksman, E., and Tamminen, J., 2001, "An Adaptive Metropolis Algorithm," *Bernoulli*, International Statistical Institute, The Hague, The Netherlands, pp. 223–242.
[26] Solonen, A., 2006, "Monte Carlo Methods in Parameter Estimation of Nonlinear Model," Master's thesis, Lappeenranta University of Technology, Lappeenranta, Finland.
[27] Mira, A., 2001, "On Metropolis-Hastings Algorithms With Delayed Rejection," Metron, **59**(3–4), pp. 231–241.
[28] Altintas, Y., and Yellowley, I., 1989, "In-Process Detection of Tool Failure in Milling Using Cutting Force Models," ASME J. Manuf. Sci. Eng., **111**(2), pp. 149–157.
[29] Shao, H., Wang, H. L., and Zhao, X. M., 2004, "A Cutting Power Model for Tool Wear Monitoring in Milling," Int. J. Mach. Tools Manuf., **44**(14), pp. 1503–1509.
[30] Choudhury, S., and Rath, S., 2000, "In-Process Tool Wear Estimation in Milling Using Cutting Force Model," J. Mater. Process. Technol., **99**(1), pp. 113–119.
[31] Waldorf, D. J., Kapoor, S. G., and DeVor, R. E., 1992, "Automatic Recognition of Tool Wear on a Face Mill Using a Mechanistic Modeling Approach," Wear, **157**(2), pp. 305–323.
[32] Sandvik, C., 2006, "Steel Milling Stars, GC4240 and GC1030, C-1140, A Pair of Grades That Won't Crack Under Pressure," Sandvik Coromant, Sandviken, Sweden.
[33] Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E., 1998, "Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions," SIAM J. Optim., **9**(1), pp. 112–147.