Proceedings of the ASME 2015 International Manufacturing Science and Engineering Conference
MSEC2015
June 8-12, 2015, Charlotte, North Carolina, USA

MSEC2015-9357

PARAMETER ESTIMATION USING MARKOV CHAIN MONTE CARLO METHOD IN MECHANISTIC MODELING OF TOOL WEAR DURING MILLING

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

Durul Ulutan
International Center for Automotive Research,
Clemson University
Greenville, SC 29607

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

ABSTRACT

Several models have been proposed to describe the relationship between cutting parameters and machining outputs such as cutting forces and tool wear. However, these models usually cannot be generalized, due to the inherent uncertainties that exist in the process. These uncertainties may originate from machining, workpiece material composition, and measurements. A stochastic approach can be utilized to compensate for the lack of certainty in machining, particularly for tool wear evolution. The Markov Chain Monte Carlo (MCMC) method is a powerful tool for addressing uncertainties in machining parameter estimation. The Hybrid Metropolis-Gibbs algorithm has been chosen in this work to estimate the unknown parameters in a mechanistic tool wear model for end milling of difficult-to-machine alloys. The results show a good potential of the Markov Chain Monte Carlo modeling in prediction of parameters in the presence of uncertainties.

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. Studying tool wear in different manufacturing processes goes back over a half century [2]. Since then, several researchers proposed various models to predict tool wear. There are many reasons that complicate tool wear studies such as its unknown dynamic evolution, dependence on material, temperature and cutting conditions, as well as existence of different tool wear mechanisms, and uncertainties in measuring tool wear [2]. The first step in studying tool wear is choosing an appropriate model. Models that are being used in tool wear studies can be divided into three different classes. Empirical models characterize tool wear through running multiple experiments in different cutting conditions and fitting an appropriate function to represent the relationship between the inputs and the outputs. Taylor's tool life model is the cornerstone to these types of models, and is used frequently [3-4]. Mechanistic models characterize tool wear as one of the inputs to the cutting force and power models. Fu et al. used cutting force model and Cuppini et al. used cutting power model to monitor tool wear [5-6]. Dynamic models characterize the dynamics of tool wear in the state space format. These models are very hard to derive and need extensive tests for validation. Danai et al. used dynamic models for online tool wear prediction during machining, however they showed that online prediction of crater wear is not accurate because of the incomplete model and inherent uncertainties of wear [7-8].

In all of the above models, one neglected factor is the stochastic behavior of tool wear. Due to this stochastic nature, it is essential to treat the model parameters as stochastic elements with certain probability distributions to be identified. Bayesian data analysis is a technique used in a wide variety of applications and it can be implemented for studying tool wear. Mehta et al. used this method for parameter identification in cutting force and spindle idle power model of turning [9-10]. Karandikar et al. used Bayesian parameter identification for estimating Taylor tool life model parameters in machining [11-12]. The power of Bayesian data analysis lies in the fact that because of incorporating initial knowledge to the current knowledge, fewer experiments are required for estimation. The Gibbs sampler and Metropolis sampling method are the most commonly used methods in Bayesian data analysis.

The objective of this work is to investigate the performance of the Metropolis-Gibbs algorithm as one of the Bayesian techniques for parameter identification in mechanistic tool wear model in milling hard-to-machine alloys when a limited number of experiments are available. The organization of this work is as follows: The theoretical background Markov Chain Monte Carlo including Gibbs and Metropolis sampler is described in section 2. In section 3, the selected mechanistic tool wear model and unknown parameters are identified. The experimental setup is described in section 4. Implementation of MCMC algorithm is discussed in Section 5, and results and conclusions are provided in sections 6 and 7.

THEORETICAL BACKGROUND

Bayes Rule

Bayesian data analysis is a powerful tool used for statistical inference. Thomas Bayes introduced the Bayesian inference and proposed the basic formulation, known as the Bayes rule, in the 18th century. According to the Bayes rule, the probability of an event θ (*i.e.* $p(\theta|y)$) can be derived by conditionally multiplying the initial belief or previous knowledge of analyst, θ, to the observed data set $p(y|\theta)$ and using the marginal distribution of observed data (*i.e.* $p(y)$) (Eq. 1). This marginal distribution can be derived as shown in Eq. (2).

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

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

In many cases of Bayesian inference, finding a closed form solution of the marginal distribution is somewhat tedious or even impossible, so it is more convenient to treat *p(y)* as a normalizing constant.

Gibbs Sampler

In general, non-conjugate priors or multi-parameter estimation can make sampling from the posterior distribution almost impossible. In this case, the Gibbs sampler is proposed [13] as one of the Markov Chain Monte Carlo methods to sample from a posterior distribution. To implement the Gibbs sampler, the closed form solution of full conditional of each parameter given all the remaining parameters is required. To illustrate how this approach works, one needs to consider a linear regression model (Eq. 3):

$$Y = \beta_i^T X + \varepsilon \tag{3}$$

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 the 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. To implement Gibbs sampler, closed from solution for the full conditional of unknown parameters is required. Assuming an initial belief for unknown coefficients with mean β_0 and variance Σ_0, full conditional of β_i can be written as:

$$p(\beta_i \mid \sigma^2, X, y_1, \ldots y_n) \propto N(\beta_n, \Sigma_n) \tag{4}$$

where β_n and Σ_n are formulated as shown in Eq. (5-6):

$$\beta_n = (\Sigma_0^{-1} + \frac{\mathrm{X}^T \mathrm{X}}{\sigma^2})^{-1} (\Sigma_0^{-1} \beta_0 + \frac{\mathrm{X}^T \mathrm{Y}}{\sigma^2}) \tag{5}$$

$$\Sigma_n = (\Sigma_0^{-1} + \frac{\mathrm{X}^T \mathrm{X}}{\sigma^2})^{-1} \tag{6}$$

The other step is calculating the full conditional of σ^2 given everything else. This can be written as below:

$$p(\sigma^2 \mid \beta_i, X, y_1, \ldots y_n) \propto IG(\frac{\nu_0 + n}{2}, \frac{\nu_0 \sigma_0^2 + SSE}{2}) \tag{7}$$

where v_0 and σ_0^2 are the sample size and sample variance of prior for σ^2, respectively, *SSE* is the sum of squared errors equal to $\sum_{i=1}^{n} (Y_i - \beta^T X)^2$ and *n* is the number of observations.

Metropolis Algorithm

The Gibbs sampler is an easy to implement and practical method in the case of linear models, while for nonlinear models or in cases where closed form solution of full conditional is not available, it does not work properly. Another technique was proposed by Nicholas Metropolis in 1953, where a new distribution called proposal density is used to approximate the actual target density, *i.e.* posterior distribution of parameters [14]. Because proposal density does not fully capture the features of target density, an acceptance-rejection method should be used to reject the samples that are generated from regions with lower probability of occurrence. Below is the step by step procedure for Metropolis algorithm [14].

(1) Selecting a candidate point from proposal density:

$$\beta_i^* = \beta_i^s + q_s \quad s \in \{1, 2, \ldots, N\}$$
$$q_s \sim N(0, \delta^2) \tag{8}$$

where β_i^* is the candidate point generated from perturbing the current position of chain, β_i^s is the current position of chain, and q_s is a sample from the proposal density function. In practical applications, a symmetric proposal density (*i.e.* normal distribution with zero mean and variance δ^2) is used as proposal density function. In this work, a method of Solonen [15] based on calculating the Jacobian matrix is used to find the proposal density variance.

(2) Calculating acceptance ratio:

$$r = \frac{p(\beta_i^* \mid \sigma^2, X, y_1, \ldots y_n)}{p(\beta_i^s \mid \sigma^2, X, y_1, \ldots y_n)} = \frac{p(y_1, \ldots y_n \mid \sigma^2, \beta_i^*, X)\, p(\beta_i^*)}{p(y_1, \ldots y_n \mid \sigma^2, \beta_i^s, X)\, p(\beta_i^s)} \quad (9)$$

$$r = \exp\{\frac{-1}{2}[\frac{SSE(\beta_i^*) - SSE(\beta_i^s)}{\sigma_s^2} + \ldots \quad (10)$$

$$\ldots N(\beta_i^*, \beta_0, \Sigma_0) - N(\beta_i^s, \beta_0, \Sigma_0)]\}$$

where $p(\beta_i^s)$ and $p(\beta_i^*)$ are calculated from the probability distribution of the initial belief, which is a normal distribution with mean β_0 and variance Σ_0.

(3) Selecting a point U from a uniform distribution [0,1]

(4) Finding $\alpha = min\{U, r\}$

(5) Accepting or rejecting the candidate point:

$$\begin{cases} \beta_i^{s+1} = \beta_i^* & \text{if } U < \alpha \\ \beta_i^{s+1} = \beta_i^s & \text{otherwise} \end{cases} \quad (11)$$

MECHANISTIC MODEL OF TOOL WEAR

Altintas et al. showed that the tangential force in milling can be described as below [16]:

$$F_t = K\bar{h}^c a_p f \sin\varphi \quad (12)$$

where K and c are 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 (Figure 1).

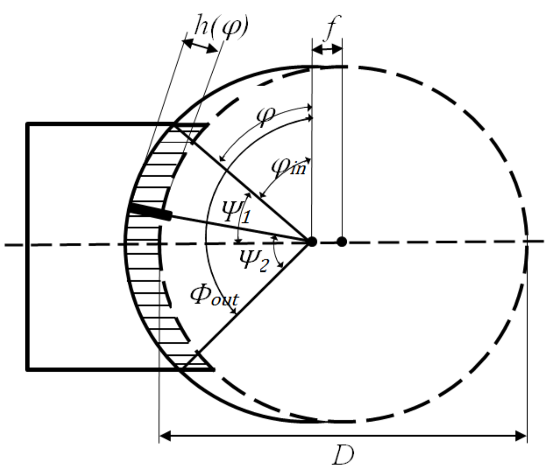

FIGURE 1: MILLING SCHEMATIC [17]

It was shown that with an increase in tool wear, the magnitude of the cutting force increases as well [18]. 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) (Eq. 13) [19]. Shao et al. derived average cutting power in milling [17] which is simplified in this work (Eq. 14), with the assumption of constant depth of cut, where P is the cutting power, N is cutting velocity and K_1, K_2 and K_3 are unknown constants in Eq. 14 that needs to be identified.

$$F_t^{wear} = \mu VBHs \quad (13)$$

$$P = K_1 N f^{K_2} + K_3 NVB \quad (14)$$

EXPERIMENTAL SETUP

Material used for this experimental study is Rene-108 (R-108). An OKUMA GENOS M460-VE 3-axis CNC machine was used to end mill (in down-milling direction) rectangular blocks of size 60 × 80 × 25 mm, using a water-soluble coolant. A 2-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) was utilized for machining. At the chosen width of cut, 8 tests were conducted on each block: 4 tests with 2 replications. 4 additional tests were also conducted to cover the full range of cutting conditions for validating the estimation. Depth of cut for each pass is kept constant at 0.5 mm, and cutting velocity and feedrate were changed as excitation factors for parameter identification. Tool Monitoring Adaptive Controller (TMAC) installed on the machine was used for measuring spindle current to monitor spindle power consumption in real time. However, due to the low sampling frequency of TMAC software (~50 Hz), an external data acquisition device (DAQ) was programmed to capture the data with high sampling rate. To measure spindle power in high sampling frequency, the output of the TMAC transducer (Figure 2) was fed into the NI9215 analog input module mounted on NI-cRIO9103 chassis programmed with LabVIEW2010. Data was collected in voltage at sampling frequency of 10.24 *KHz*.

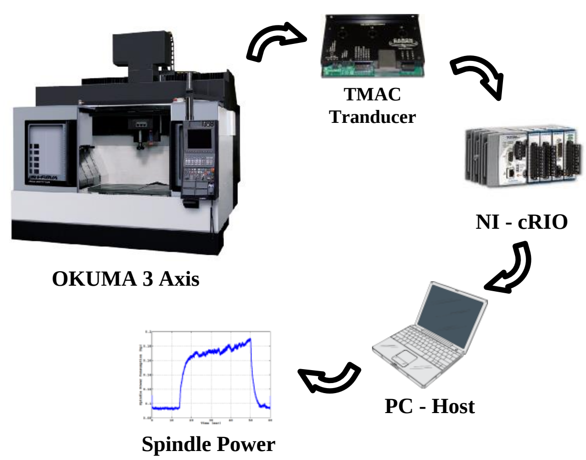

FIGURE 2: DATA ACQUISITION WITH NI-cRIO9103

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. 14. 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 pass. As shown in Figure 3 for cutting power at cutting velocity of 50 m/min and feedrate of 0.1 mm/rev, the increase in the power illustrates the developing tool flank wear during the cutting process. The mean value of cutting power between 42-48 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.* $K_1 N f^{K_2}$), 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 08M-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 [20]. Fresh unworn inserts were used for each test, and flank wear on the bottom edge of each insert was measured using an Olympus optical microscope and average flank wear was calculated. Measured tool flank wear for tests 1-4 is shown in FIGURE 4.

TABLE 1: DOE TABLE FOR END MILLING R-108

Test #	V_c (m/min)	f (mm/rev)	P (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

INFERENCE ON MODEL PARAMETERS

The target of this work is to identify unknown parameters K_1, K_2 and K_3 and measurement error variance σ^2 in the case of limited number of experiments. Due to the nonlinearity of Eq. 14, finding the full conditional of the unknown parameters is impossible but the full conditional of measurement error variance (σ^2) is available, so Hybrid Metropolis-Gibbs algorithm was used for inference. In FIGURE 5, the flow chart of the algorithm is shown.

Note that because full conditional of σ^2 is available, all the samples from Gibbs sampler are accepted automatically. The prior belief for unknown parameters is chosen as normal distribution with the equal mean (equal to 1) and a large covariance matrix (Eq. 15). The prior belief for measurement covariance is chosen as inverse gamma function with v_0=1 and $\sigma_0^2=100$ (Eq. 16).

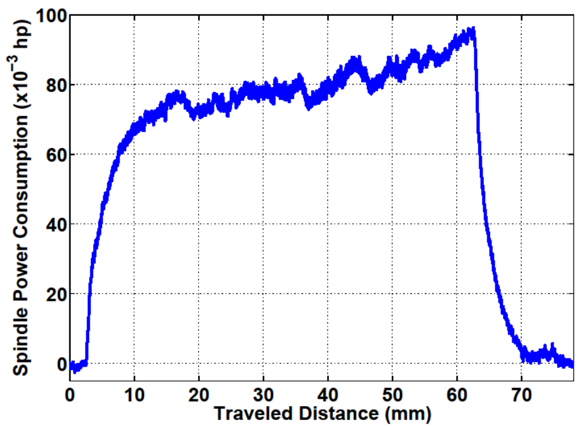

FIGURE 3: CUTTING POWER OF TEST 3, V_c=50 M/MIN, f=0.1 MM/REV

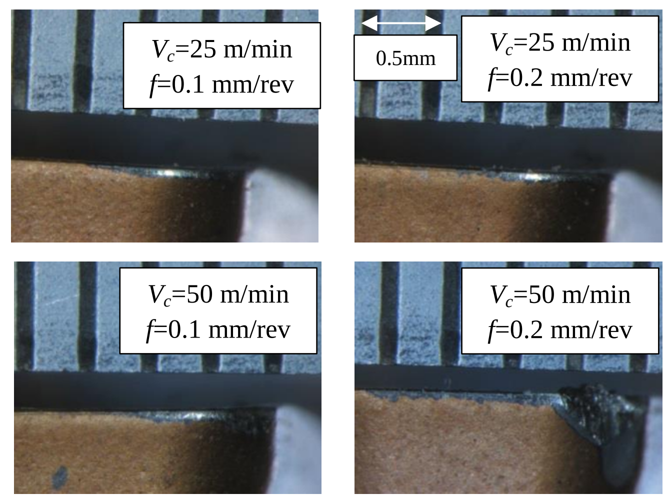

FIGURE 4: MEASURED FLANK WEAR FOR TESTS 1-4

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \sim \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \quad (15)$$

$$\sigma^2 \sim IG\left(\frac{1}{2}, \frac{2}{100^2}\right) \quad (16)$$

To avoid singularity of covariance matrix, spindle power consumption is divided by 1000 so that K_1 and K_3 are in the same range as K_2. The rest of this study is based on normalized value of K_1 and K_3. The optimal value of unknown parameters was calculated from data in Table 1 using unconstrained derivative free optimization method. This method uses simplex search algorithm described by [21]. 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 discovered. Note that this method only finds local maxima of the function.

$$\begin{bmatrix} K_1 & K_2 & K_3 \end{bmatrix}_{opt}^T = \begin{bmatrix} 0.94 & 1.21 & 0.15 \end{bmatrix}^T \quad (17)$$

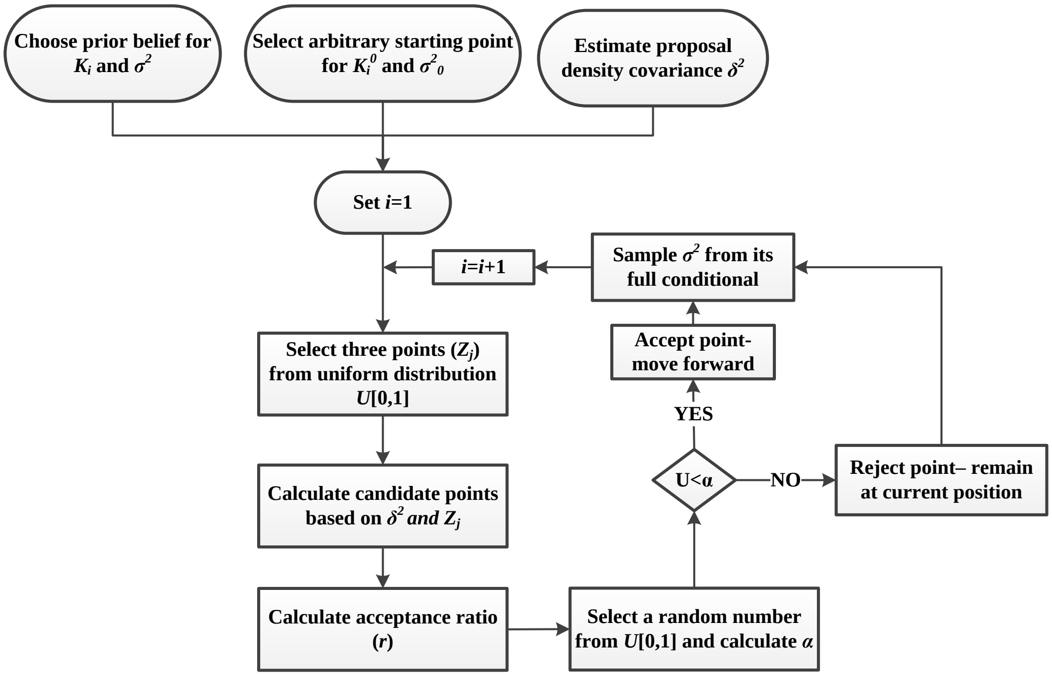

FIGURE 5: FLOWCHART OF Hybrid METROPOLIS-GIBBS ALGORITHM

The Jacobian matrix was calculated using K^{opt} for each test and proposal density covariance matrix was calculated using the Jacobian matrix of unknowns [15]. 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.

$$J = \begin{bmatrix} N_1 f_1^{K_2^{opt}} & K_1 N_1 f_1^{K_2^{opt}} \log(K_2^{opt}) & N_1 VB_1 \\ \ldots & \ldots & \ldots \\ N_8 f_8^{K_2^{opt}} & K_1 N_8 f_8^{K_2^{opt}} \log(K_2^{opt}) & N_8 VB_8 \end{bmatrix}_{8\times 3} \quad (18)$$

$$\delta^2 = \begin{bmatrix} 0.87 & 0.78 & 0.51 \\ 0.78 & 0.74 & 0.52 \\ 0.51 & 0.52 & 0.41 \end{bmatrix} \quad (19)$$

$$K_i^0 = K_i^{opt} = \begin{bmatrix} 0.94 & 1.21 & 0.15 \end{bmatrix}^{\mathrm{T}} \quad i \in \{1,2,3\} \quad (20)$$

RESULTS

MCMC trace plot for the parameters K_1, K_2 and K_3 following the procedure described in previous section with N=2000 points is shown in Figure 6. Usually the Markov chain needs some time to converge (burn-in period), so the first set of iterations are usually discarded to reduce the effect of initial errors at the start of the chain. In this work, the first 20% of the iterations (400 points) were discarded as burn-in period. After removing the first 400 points, the acceptance ratio was calculated as 19%. It is clear from low acceptance ratio and from Figure 6 that parameters were not able to converge to the posterior values. Current results can be used as the initial belief for the next MCMC run (Main Run). To do so, the mean and covariance matrices of the initial run (*i.e.* Pilot Run) were extracted from the chain and implemented as prior belief. The covariance matrix of parameters was used as the proposal density covariance as well, and the final point of the chain was used as the initial point of the new chain. For the Main Run, number of samples was N=10,000.

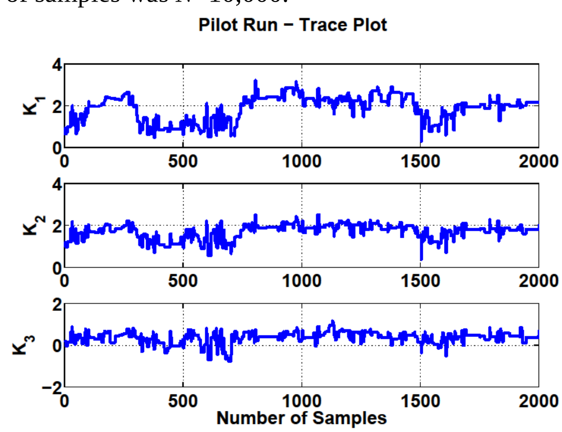

FIGURE 6: TRACE PLOT OF PARAMETERS IN PILOT RUN

A trace plot of parameters is shown in Figure 7. It was observed that the chain has converged and parameters mixed well. After removing the first 20% of the iterations (2,000 points) as burn-in period, a comparison of each pair of parameters distributions are shown in Figure 8 to 10, where contour (a) is the initial belief (prior distribution) from the Pilot Run, and contour (b) is the posterior belief (current knowledge) of parameters. The multivariate posterior distribution of identified parameters is shown in Eq. 21. Figure 11 also demonstrates the improvement in the degree of uncertainties after each MCMC run. In other words, 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 Figure 11 covers a wide range of possible values for unknown parameters (*i.e.* large variance, shown as red curve), however by running the MCMC algorithm 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 parameters (shown as blue and black curves).

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \sim N\left(\begin{bmatrix} 1.94 \\ 1.79 \\ 0.43 \end{bmatrix}, \begin{bmatrix} 0.20 & 0.07 & 0.01 \\ 0.07 & 0.04 & 0.02 \\ 0.01 & 0.02 & 0.02 \end{bmatrix} \right) \tag{21}$$

A gamma distribution of the inverse of measurement error variance ($1/\sigma^2$) which was generated from the Gibbs sampler is shown in Figure 12, and the mode is chosen as the posterior value of measurement error variance σ_N^2=44.

Now that the unknown parameters are 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 set of 4 new tests. It is fairly straightforward to find the posterior distribution of observation (posterior predictive distribution) with the following algorithm:

(1) Starting from i=1, $i \in \{1,2,...,N\}$

(2) Sample K_1^i, K_2^i and K_3^i from their posterior distribution

(3) Sample σ_i^2 from posterior distribution of measurement error variance

(4) Sample ε_i as measurement error, $\varepsilon_i \sim N(0,\sigma_i^2,)$

(5) Calculate the output for each set of 4 validation tests,

$$Y_i = K_1^i N f^{K_2^i} + K_3^i NVB + \varepsilon_i$$

The expectation is that the measured power be within the 95% confidence interval for each of the 4 tests. As shown in Figure 13 and 14, the model was able to predict the measured power with good 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 able to predict spindle power consumptions with good degree of accuracy which implies validity of identified parameters.

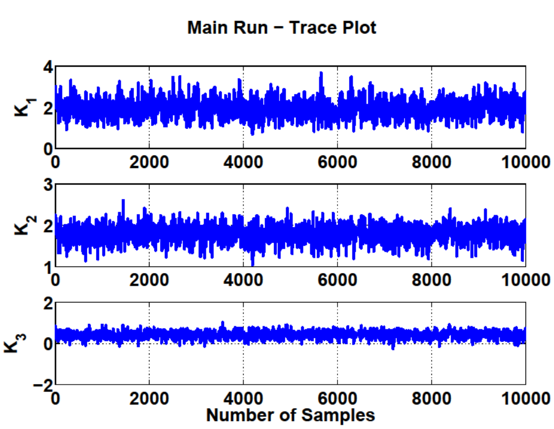

FIGURE 7: TRACE PLOT OF PARAMETERS IN MAIN RUN

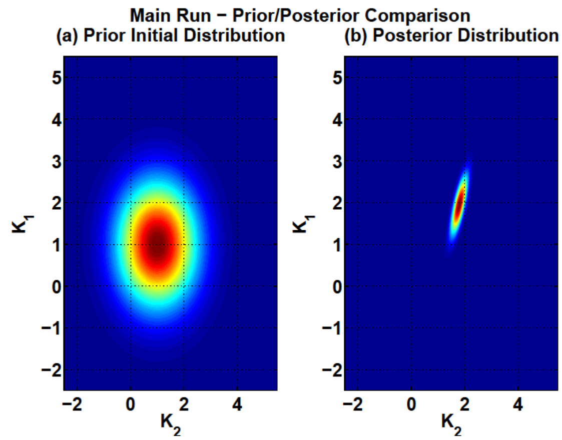

FIGURE 8: K_1 AND K_2 DISTRIBUTION: (a) PRIOR INITIAL BELIEF (b) POSTERIOR PROBABILITY OF PARAMETERS

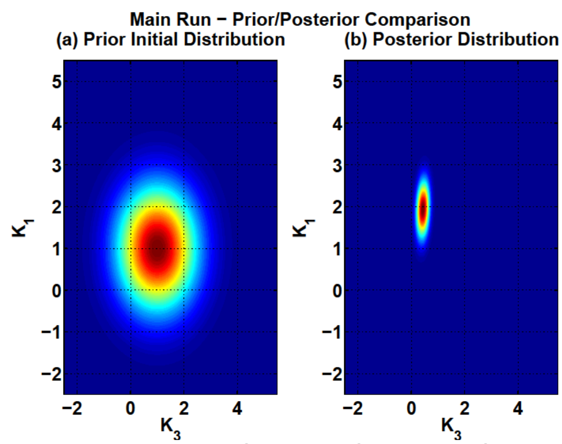

FIGURE 9: K_1 AND K_3 DISTRIBUTION: (a) PRIOR INITIAL BELIEF, (b) POSTERIOR PROBABILITY OF PARAMETERS

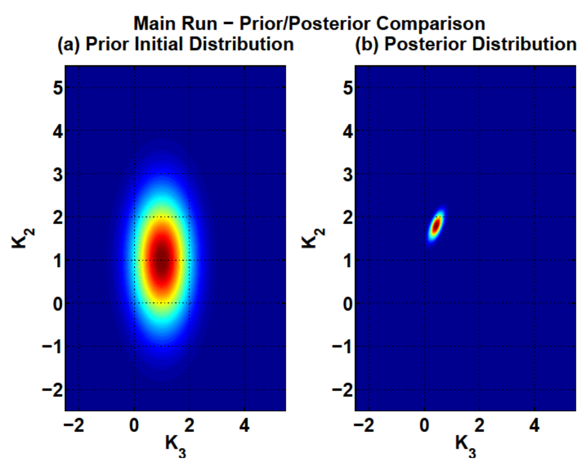

FIGURE 10: K_2 AND K_3 DISTRIBUTION: (a) PRIOR INITIAL BELIEF, (b) POSTERIOR PROBABILITY OF PARAMETERS

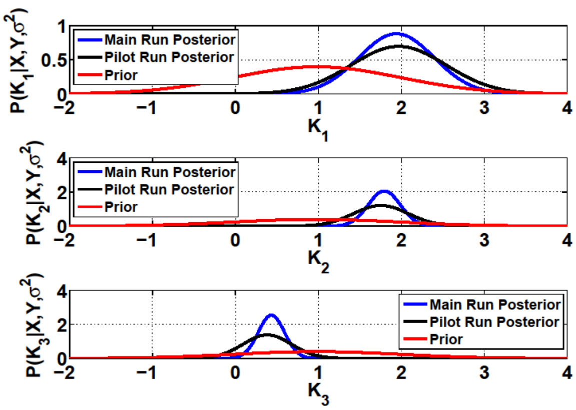

FIGURE 11: DISTRIBUTION OF PARAMETERS FOR PRIOR BELIEF, PILOT RUN AND MAIN RUN

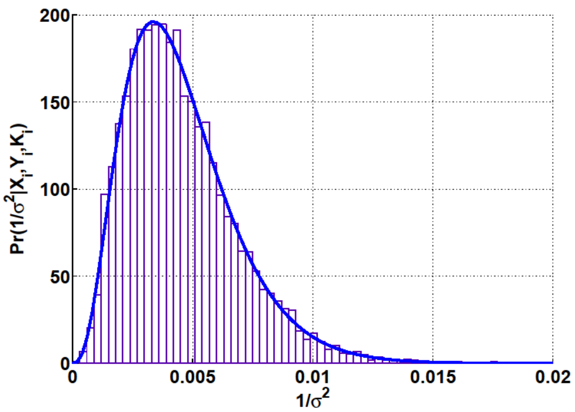

FIGURE 12: GAMMA DISTRIBUTION OF THE INVERSE OF MEASUREMENT ERROR VARIANCE

TABLE 2: VALIDATION TESTS CUTTING CONDITION, MEASURED POWER AND TOOL FLANK WEAR

Test #	V_c (m/min)	f (mm/rev)	P (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

TABLE 3: PERCENTAGE ERROR OF PREDICTION AND MEASUREMENT

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

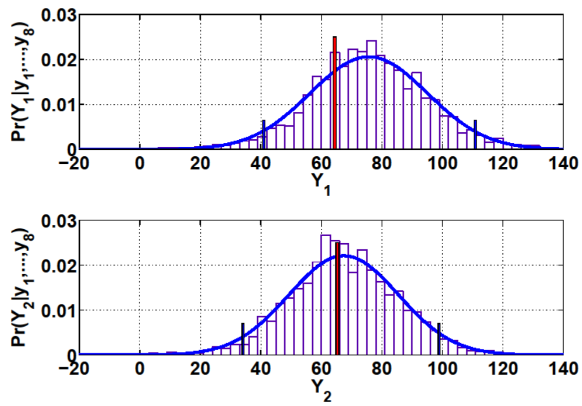

FIGURE 13: POSTERIOR PREDICTIVE DISTRIBUTION WITH 95% CONFIDENCE INTERVAL (BLACK BARS) AND SPINDLE POWER (RED BAR) – VALIDATION TESTS 1-2

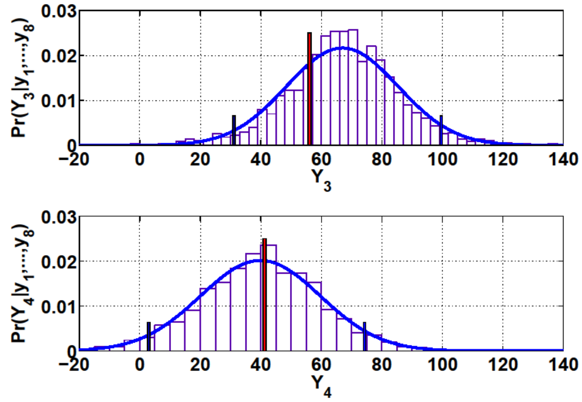

FIGURE 14: POSTERIOR PREDICTIVE DISTRIBUTION WITH 95% CONFIDENCE INTERVAL (BLACK BARS) AND SPINDLE POWER (RED BAR) – VALIDATION TESTS 3-4

CONCLUSIONS

This work deals with the Bayesian parameter identification of tool wear mechanistic model. This technique can be used when limited number of experiments are available, which is beneficial and cost effective in studying hard to machine alloys. The main conclusions are:

- The Hybrid Gibbs-Metropolis algorithm was formulated for prediction of unknown parameters in nonlinear mechanistic cutting power model in milling of R-108. Metropolis algorithm with symmetric proposal density was used for predicting the model parameters, while Gibbs sampler used for updating measurement error variance.
- A design of experiment with mild and aggressive cutting condition was used along with high frequency DAQ to capture wide range of tool wear and spindle power consumption. The performance of algorithm improved significantly after using a 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 when measuring cutting power by simply using the mean value of identified parameters and writing mechanistic tool wear equations in terms of tool flank wear as the output and spindle power consumption as the input. The identified parameters can also be used for online tracking of tool wear in industrial manufacturing plant with online Bayesian techniques such as Kalman filter or particle filter which is a subject of study for future work.

ACKNOWLEDGEMENTS

The authors wish to thank the National Science Foundation for support of this work under Grant No. 0954318. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Vallejo, A. G., Jr., Nolazco-Flores, J., Morales-Menendez, R., Sucar, L. E., & Rodriguez, C. A. (2005). Tool-wear monitoring based on continuous hidden markov models. In A. Sanfeliu, & M. Cortas, *Progress in pattern recognition, image analysis and applications* (pp. 880-890) Springer Berlin Heidelberg.

[2] Yen, Y., Söhner, J., Lilly, B., & Altan, T. (2004). Estimation of tool wear in orthogonal cutting using the finite element analysis. *Journal of Materials Processing Technology, 146*(1), 82-91.

[3] Tlusty, J. (2000). *Manufacturing processes and equipment* Prentice Hall.

[4] Xu, C., Xu, T., Zhu, Q., & Zhang, H. (2011). Study of adaptive model parameter estimation for milling tool wear. *Strojniški Vestnik-Journal of Mechanical Engineering, 57*(7-8), 568-578.

[5] Fu, H. J., DeVor, R. E., & Kapoor, S. G. (1984). A mechanistic model for the prediction of the force system in face milling operations. *Journal of Manufacturing Science and Engineering, 106*(1), 81-88. doi:10.1115/1.3185915

[6] Cuppini, D., D'errico, G., & Rutelli, G. (1990). Tool wear monitoring based on cutting power measurement. *Wear, 139*(2), 303-311. doi:http://dx.doi.org/10.1016/0043-1648(90)90052-C

[7] Danai, K., & Ulsoy, A. G. (1987). An adaptive observer for on-line tool wear estimation in turning, part I: Theory. *Mechanical Systems and Signal Processing, 1*(2), 211-225.

[8] Danai, K., & Ulsoy, A. G. (1987). An adaptive observer for on-line tool wear estimation in turning, part II: Results. *Mechanical Systems and Signal Processing, 1*(2), 227-240.

[9] Mehta, P., Kuttolamadom, M., & Mears, L. (2012). Machining process power monitoring: Bayesian update of machining power model. Paper presented at the *ASME 2012 International Manufacturing Science and Engineering Conference Collocated with the 40th North American Manufacturing Research Conference and in Participation with the International Conference on Tribology Materials and Processing,* 745-752.

[10] Mehta, P., & Mears, L. (2013). Model learning in a multistage machining process: Online identification of force coefficients and model use in the manufacturing enterprise. Paper presented at the *ASME 2013 International Manufacturing Science and Engineering Conference Collocated with the 41st North American Manufacturing Research Conference,* V002T02A026.

[11] Karandikar, J. M., Abbas, A. E., & Schmitz, T. L. (2014). Tool life prediction using bayesian updating. part 1: Milling tool life model using a discrete grid method. *Precision Engineering, 38*(1), 9-17.

[12] Karandikar, J. M., Abbas, A. E., & Schmitz, T. L. (2014). Tool life prediction using bayesian updating. part 2: Turning tool life using a markov chain monte carlo approach. *Precision Engineering, 38*(1), 18-27.

[13] Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain monte carlo in practice* Taylor & Francis.

[14] Hoff, P. D. (2009). *A first course in bayesian statistical methods* Springer.

[15] Solonen, A. (2006). *Monte Carlo Methods in Parameter Estimation of Nonlinear Model.*

[16] Altintas, Y., & Yellowley, I. (1989). In-process detection of tool failure in milling using cutting force models. *Journal of Manufacturing Science and Engineering, 111*(2), 149-157. doi:10.1115/1.3188744

[17] Shao, H., Wang, H. L., & Zhao, X. M. (2004). A cutting power model for tool wear monitoring in milling. *International Journal of Machine Tools and Manufacture, 44*(14), 1503-1509. doi:http://dx.doi.org/10.1016/j.ijmachtools.2004.05.003

[18] Choudhury, S., & Rath, S. (2000). In-process tool wear estimation in milling using cutting force model. *Journal of Materials Processing Technology, 99*(1), 113-119.

[19] Waldorf, D. J., Kapoor, S. G., & DeVor, R. E. (1992). Automatic recognition of tool wear on a face mill using a mechanistic modeling approach. *Wear, 157*(2), 305-323.

[20] Sandvik, C. (2006). *Steel milling stars, GC4240 and GC1030, C-1140, A pair of grades that won't crack under pressure*.

[21] Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the nelder--mead simplex method in low dimensions. *SIAM Journal on Optimization, 9*(1), 112-147.