TSK fuzzy modeling for tool wear condition in turning processes: An experimental study

Qun Ren a, *, Marek Balazinski a, Luc Baron a, Krzysztof Jemielniak b

a Mechanical Engineering Department, École Polytechnique de Montréal, C.P. 6079, succ. Centre-Ville, Montréal, Québec, Canada H3C 3A7
b Faculty of Production Engineering, Warsaw University of Technology, Narbutta 86, 02-524 Warsaw, Poland

1. Introduction

Tool condition has a strong influence on the resulting surface finish and dimensional integrity of the workpiece, as well as vibration levels of the machine tool. The information obtained from tool wear monitoring can be used for several purposes that include establishing tool change policy, economic optimization of machining operations, compensating for tool wear on-line and to some extent avoiding catastrophic tool failures (El-Gomayel and Bregger, 1998). Effective monitoring of a manufacturing process is essential for ensuring product quality and reducing production costs. Analysis, implementation and evaluation of machining processes present significant challenges to the manufacturing industry. The machining process varies considerably depending on the workpiece material, temperature, cutting fluids, chip formation, the tool material, chatter and vibration. The strain rate is extremely high compared to that of other fabrication processes as well (Moriwaki, 1983). Because the information obtained during machining process is/could be incomplete, imprecise, vague, contradictory, or deficient in some other way, there are very few established analytical models due to the difficulty in understanding the exact physics in most precision manufacturing processes. Researchers are required for quantitative evaluations of the transmission characteristics of cutting parameters and the calibration of the sensors (Lee, 1991).

Fuzzy modeling, known as qualitative modeling based on fuzzy logic (FL) (Zadeh, 1968, 1973), has the capability to model complex system behavior in such a qualitative way that the model is more effective and versatile in capturing the behavior of ill-defined systems with fuzziness or fully defined system with realistic approximation. Takagi–Sugeno–Kang (TSK) fuzzy modeling (Takagi and Sugeno, 1985; Sugeno and Kang, 1988), was proposed in an effort to develop a systematic approach to generating fuzzy rules from a given input–output data set. This model consists of rules with fuzzy antecedents and mathematical function in the consequent part. The antecedents divide the input space into a set of fuzzy regions, while consequents describe behavior of system in those regions. There are some works in modeling manufacturing process on using TSK-fuzzy model, for example, surface roughness prediction with genetic algorithm (GA) (Nandi, 2006), drilling performances (Nandi and Davim, 2009), processing times estimation based on genetic programming (Mucientes et al., 2009), etc.

The aim of this paper is to present TSK fuzzy modeling based on subtractive clustering method (Chiu, 1994) in order to accomplish the integration of multi-sensor information and tool wear information. It generates fuzzy rules directly from the input–output data acquired from sensors, and provides high accuracy and high reliability of the tool wear prediction over a wide range of cutting conditions. The experimental results show its effectiveness and satisfactory comparisons relative to other artificial intelligence methods.

© 2010 Elsevier Ltd. All rights reserved.
experimental results from the TSK fuzzy approach and compares it with other different artificial intelligence methods—neural network (NN), Mamdani FL and a neural network based fuzzy system (NF). Section 6 contains concluding remarks and future research recommendations.

2. Subtractive clustering based TSK fuzzy modeling

The structure of a fuzzy TSK model can be done manually based on knowledge about the target process or using data-driven techniques. Identification of the system using clustering involves formation of clusters in the data space and translation of these clusters into TSK rules such that the model obtained is close to the system to be identified.

The aim of Chiu’s (1994) subtractive clustering identification algorithm is to estimate both the number and initial location of cluster centers and extract the TSK fuzzy rules from input/output data. Subtractive clustering operates by finding the optimal data points to define a cluster center based on the density of surrounding data points. This method is a fast clustering method designed for high dimension problems with a moderate number of data points. This is because its computation grows linearly with the data dimension and as the square of the number of data points. A brief description of Chiu’s subtractive clustering method is as follows.

Consider a collection of data points \{x_{1}, x_{2}, ..., x_{n}\} specified by \( m \)-dimensional \( x_{i} \). Without loss of generality, assume the feature space is normalized, so that all data are bounded by a unit hypercube. Calculate potential for each point by using the following equation:

\[
p_i = \sum_{j=1}^{n} e^{-\alpha |x_i - x_j|^2}, \quad \alpha = 4/r_n^2
\]

(1)

where \(| | \cdot | |\) denotes the Euclidean distance. It is noteworthy that only the fuzzy neighborhood within the radius \( r_n \) is to the measure of potential.

After the potential of every data point has been computed, the data point with the highest potential is selected as the first cluster center. Assume \( x_1^* \) is the location of the first cluster center, and \( p_1^* \) is its potential value, then revise the potential of each data point \( x_i \) by the formula

\[
p_i \leftarrow p_i - p_1^* e^{-\beta |x_i - x_1^*|^2}
\]

(2)

where \( \beta = 4/r_n^2 \) and \( r_n = \eta r_p \).

When the potential of all data points have been reduce by (2), the data point with the highest remaining potential is selected as the second cluster center. Then further reduce the potential of each data points. Generally, after \( k \)th cluster center has been obtained, the potential of each data point is revised by formula

\[
p_i \leftarrow p_i - p_1^* e^{-\beta |x_i - x_k^*|^2}
\]

(3)

where \( x_k^* \) is the location of the \( k \)th cluster center and \( p_k^* \) is its potential value.

The process of acquiring new cluster center and revising potential repeats by using the following criteria:

- if \( p_k^* > \beta p_1^* \), accept \( x_k^* \) as a cluster center and continue.
- else if \( p_k^* < \beta p_1^* \)
  - reject \( x_k^* \) and end the clustering process.
- else
  - let \( d_{\text{min}} \) = shortest of the distances between \( x_k^* \) and all previously found cluster centers.
  - if \( \frac{d_{\text{min}}}{p_1^*} + \frac{p_k^*}{\beta} \geq 1 \)
    - accept \( x_k^* \) as a cluster center and continue.
  - otherwise
    - reject \( x_k^* \) and end the clustering process.

Using standard Gaussian membership degree can be expressed as

\[
\mu_k = e^{-\frac{1}{2} \left( \frac{x_i - x_k^*}{\sigma} \right)^2}
\]

Cluster center found in the training data are points in the feature space whose neighborhood maps into the given class. Each cluster center can be translated into a fuzzy rule for identifying the class.

A generalized type-1 TSK model can be described by fuzzy IF-THEN rules which represent input–output relations of a system. For a multi-input single-output (MISO) first-order type-1 TSK model, its \( k \)th rule can be expressed as

\[
\text{IF } x_1 \text{ is } Q_{1k} \text{ and } x_2 \text{ is } Q_{2k} \text{ and } \ldots \text{ and } x_n \text{ is } Q_{nk},
\]

\[
\text{THEN } Z = w_k^0 + p_k^1 x_1 + p_k^2 x_2 + \ldots + p_k^n x_n
\]

(5)

where \( x_1, x_2, ..., x_n \) and \( Z \) are linguistic variables \( Q_{1k}, Q_{2k}, ..., Q_{nk} \) and \( w_k^0 \) are the fuzzy sets on universe of discourses \( U, V, ..., W \), and \( w_k^0 \) are regression parameters.

With subtractive clustering method, \( x_j \) is the \( j \)th input feature \( f_{\text{input}} \) \( f_{\text{input}} \) \( (j \in [1, n]) \) and \( Q_{nk} \) is the MF in the \( k \)th rule associated with \( j \)th input feature. The MF \( Q_{nk} \) can be obtained as

\[
Q_{nk} = \exp \left[ -\frac{1}{2} \left( \frac{x_j - x_{nk}^*}{\sigma} \right)^2 \right]
\]

(6)
where $x^j_k$ is the $j$th input feature of $x^k$, the standard deviation of Gaussian MF given as

$$
\sigma = \sqrt{1/2\pi}
$$

(7)

The consequence parameter values of (5) can be obtained from a least squares estimation algorithms. Details are given in (Ren et al., 2008).

3. Experimental setup and data acquirements

3.1. Experimental setup

Fig. 1 shows the structure of tool condition monitoring system. In such system, real-time data are first acquired from sensors (e.g. dynamometers on Fig. 2) located as different locations of the workpiece, tool and machine-tool, then signal processing technique (amplifiers) are used to extract valid data, and decision making system (computers) is used to analyze the data and classify the results in order to make a more reliable estimation of the state of the tool and consequently of the machined parts themselves (Balazinski et al., 1994). Fig. 2 shows the experimental setup.

Fig. 3. Cutting parameters.

Fig. 4. Cutting force measurement.

Fig. 5. Subtractive clustering based fuzzy approach.
3.2. Data acquisiton

Cutting force measurement, currently the most reliable and accurate sensing method available in metal cutting, is one of the most commonly employed methods for on-line tool wear monitoring. It is frequently applied in turning processes because cutting force values are more sensitive to tool wear than other measurements such as vibration or acoustic emission (Du et al., 1995).

The experiments described in this paper were conducted on a conventional lathe TUD-50. A CSRPR 2525 tool holder equipped with a TiN—Al2O3—TiC coated sintered carbide insert SNUN 120408 was used in the test. To simulate factory floor conditions, six sets of cutting parameters were selected and applied in sequence as presented in Fig. 3. During machining, the feed force \( (F_f) \) and the cutting force \( (F_c) \) were recorded while the tool wear was manually measured after each test for a given feed rate \( (f) \).

For our purposes, tool wear \( (VB) \) was estimated from three input sources: \( f, F_f \) and \( F_c \). The choice of input variables was based on the following two observations: \( F_f \) is independent of \( f \) but rather depends on \( VB \) and the depth of cut, denoted \( q_p \). Moreover, \( F_c \) depends on \( q_p \) and \( f \), while being only weakly dependent on \( VB \). So, in this paper \( f \) and the measurement \( F_c \) are used to determine \( q_p \), and the measurement \( F_f \) is used to determine VB without requiring \( q_p \) as an input variable.

Cutting speeds were selected in such a way as to correspond to the same approximate tool life in each cut (shown in Fig. 3). VB was measured after carrying out each sequence. The value for \( F_f \) and \( F_c \) were measured by a single cut using a Kistler 9263 dynamometer (as shown in Fig. 2) during 5-s intervals while the cut was executed. Recent research has attempted to investigate the application of multiple sensors with complementing characteristics to provide a robust estimate of tool wear condition. Since the inserts used in the experiments had a soft, cobalt-enriched layer of the substrate under the coating, their tool life had a tendency to end suddenly after this coating wore through.

The experiments were carried out until a tool failure occurred. Two experiments were carried out until a tool failure occurred. In the first tests (designated W5) 10 cycles were performed until a sudden rise of the flank wear VB occurred, reaching approximately 0.5 mm. In the second test (designated W7) failure of the coating resulted in chipping of the cutting edge at the end of 9th cycle. W5 was devised for TSK fuzzy rule identification, while W7 was used to verify the performance of the different monitoring system. Fig. 4

| Table 1 |

Nine cluster centers obtained by using subtractive clustering with the four parameters initialized as \( r_0 = 0.25, \ z = 0.6, \ F = 1 \).

<table>
<thead>
<tr>
<th>Cluster</th>
<th>( f ) (mm)</th>
<th>( F_f ) (N)</th>
<th>( F_c ) (N)</th>
<th>VB (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.47</td>
<td>1397</td>
<td>389</td>
<td>0.145</td>
</tr>
<tr>
<td>2</td>
<td>0.47</td>
<td>1411</td>
<td>341</td>
<td>0.156</td>
</tr>
<tr>
<td>3</td>
<td>0.47</td>
<td>1044</td>
<td>405</td>
<td>0.158</td>
</tr>
<tr>
<td>4</td>
<td>0.47</td>
<td>1395</td>
<td>372</td>
<td>0.130</td>
</tr>
<tr>
<td>5</td>
<td>0.33</td>
<td>1070</td>
<td>365</td>
<td>0.148</td>
</tr>
<tr>
<td>6</td>
<td>0.33</td>
<td>1112</td>
<td>423</td>
<td>0.158</td>
</tr>
<tr>
<td>7</td>
<td>0.33</td>
<td>1332</td>
<td>347</td>
<td>0.102</td>
</tr>
<tr>
<td>8</td>
<td>0.33</td>
<td>1032</td>
<td>334</td>
<td>0.113</td>
</tr>
<tr>
<td>9</td>
<td>0.47</td>
<td>1455</td>
<td>508</td>
<td>0.200</td>
</tr>
</tbody>
</table>

| Table 2 |

TSK fuzzy model.

<table>
<thead>
<tr>
<th>rule</th>
<th>( \text{If then } y = p_1 x_1 + p_2 x_2 + p_3 x_3 + p_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.47}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1397}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 389}{0.25} \right) ). then ( VB = 2.93 \times 10^3 f - 0.13 F_f + 3.45 F_c - 1.44 \times 10^{12} ).</td>
</tr>
<tr>
<td>2</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.47}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1413}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 431}{0.25} \right) ). then ( VB = -6.14 \times 10^3 f - 7.05 \times 10^{-1} F_f - 0.34 F_c + 3.21 \times 10^{11} ).</td>
</tr>
<tr>
<td>3</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.33}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1044}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 405}{0.25} \right) ). then ( VB = 8.32 f - 0.64 \times 10^4 F_f - 4.51 \times 10^{-1} F_c + 5.85 ).</td>
</tr>
<tr>
<td>4</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.47}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1395}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 372}{0.25} \right) ). then ( VB = -5.90 \times 10^3 f - 0.34 F_f + 4.57 F_c + 2.80 \times 10^{12} ).</td>
</tr>
<tr>
<td>5</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.33}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1070}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 365}{0.25} \right) ). then ( VB = -10.67 f + 3.12 \times 10^2 F_f - 1.05 \times 10^{-2} F_c + 2.07 ).</td>
</tr>
<tr>
<td>6</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.33}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1112}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 423}{0.25} \right) ). then ( VB = 0.62 f - 1.34 \times 10^2 F_f + 2.61 \times 10^{-1} F_c + 4.90 ).</td>
</tr>
<tr>
<td>7</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.47}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1332}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 347}{0.25} \right) ). then ( VB = 1.14 \times 10^3 f - 1.43 \times 10^{-2} F_f - 0.18 F_c - 5.38 \times 10^{12} ).</td>
</tr>
<tr>
<td>8</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.33}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1032}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 334}{0.25} \right) ). then ( VB = 2.41 f - 1.14 \times 10^2 F_f + 1.30 \times 10^{-2} F_c + 1.20 ).</td>
</tr>
<tr>
<td>9</td>
<td>If ( AF_f = \exp \left( \frac{4 f - 0.47}{0.25} \right) ), ( AF_c = \exp \left( \frac{4 F_f - 1455}{0.25} \right) ), and ( AF_{q_p} = \exp \left( \frac{4 F_c - 508}{0.25} \right) ). then ( VB = 7.19 f - 9.09 \times 10^{-4} F_f - 3.28 \times 10^{-2} F_c + 2.50 \times 10^{12} ).</td>
</tr>
</tbody>
</table>
presents the cutting force components $F_f$ and $F_c$ versus VB obtained in the experiments. Apparently, $F_c$ is weakly dependent on VB. It is a function of the cutting parameters only. On the other hand, $F_f$ is independent of $F_c$, being affected only by VB. This provides an interesting opportunity to estimate the tool wear without requiring information about $a_p$.

4. TSK fuzzy modeling

For this experimental, TSK fuzzy modeling, in which subtractive clustering method is combined with a least-square estimation algorithm, was used to accomplish the integration of multi-sensor information and tool wear information. Detailed description for this method can be found in (Ren et al., 2008). The diagram of the algorithm is shown in Fig. 5.

By using TSK fuzzy approach, a nine rule TSK fuzzy model can be obtained to describe VB with $f$, $F_f$ and $F_c$ as input variables.

4.1. Cluster centers

Identification of the system using clustering involves formation of clusters in the data space and translation of these clusters into TSK rules such that the model obtained is close to the system to be identified. Extended subtractive clustering identification algorithm (Demirli et al., 2003) is used to estimate both the number and initial location of cluster centers and extract the TSK fuzzy rules from input/output data. The cluster radius is confined to the range $[0.15; 1.0]$ with a step size of $0.15$. The accept ratio and the reject ratio are both considered in the range $[0; 1.0]$ with a step size of $0.1$. The squash factor is considered in the range $[0.05; 2]$ with a step size of $0.05$. Running time of the proposed method for this experimental is about 5 min on a standard personal computer.

Table 1 lists the nine cluster centers obtained by subtractive clustering method from the first experiment W5.

4.2. Fuzzy rules

In this paper, the consequent functions are linear. Least-square estimation [8] is used to identify the consequent parameters of the TSK model, where the premise structure, premise parameters, consequent structure and consequent parameters were identified and adjusted recursively. Table 2 is the obtained fuzzy model for tool wear condition in this experiment.

5. Experimental results

Fig. 6 summarizes results of tool wear conditioning from W5 (learning) and W7 (testing) and compares them with several different artificial intelligence methods—NN, Mamdani FL and NF. Detailed parameters setting of these methods were described in Balazinski et al. (2002), which were applied to the same experimental arrangements. Fig. 6 shows that the results from the proposed method fit the experimental data better than other AI methods.

The quality of the tool wear estimation is evaluated by root-mean-square-error (rmse) (Eq. (8)) and maximum error (Eq. (9)) for each AI methods.

$$\text{rmse} = \sqrt{\frac{1}{N} \sum (\text{VB}_m - \text{VB}_e)^2} / N$$

$$\text{max} = \max (\text{VB}_m - \text{VB}_e)$$

where VB$_m$ and VB$_e$ are measured and estimated flank wear, respectively, and N is the number of patterns in the set (N=71 for the experiment W5 and N=66 for the experiment W7).

In Table 3, the TSK fuzzy approach has the lowest root-mean-square-error and the smallest maximum error.

The TSK fuzzy modeling program used for tool condition monitoring in this paper was developed by Geno-flou development group in Lab CAE at École Polytechnique de Montréal.

6. Conclusion

A TSK fuzzy approach using subtractive clustering is described in detail in this article. The experimental results show its effectiveness and a satisfactory comparison with several other artificial intelligence methods. But for tool condition monitoring, none of existing methods can overcome the defect that the models are difficult to estimate the errors of approximation. Our future

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Summary of root-mean-square-error (rmse) and maximum error (max) from the experimental results with different AI methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI method</td>
<td>Learning (W5)</td>
</tr>
<tr>
<td></td>
<td>rmse (mm)</td>
</tr>
<tr>
<td>Neural network</td>
<td>0.015</td>
</tr>
<tr>
<td>FDSS (Mamdani FL)</td>
<td>0.024</td>
</tr>
<tr>
<td>NF</td>
<td>0.014</td>
</tr>
<tr>
<td>TSK FL</td>
<td>0.006</td>
</tr>
</tbody>
</table>
research aims at developing a new approach to capture the uncertainty during tool wear monitoring.

References


