

# CNC machine tool health assessment using Dynamic Bayesian Networks

D.A. Tobon-Mejia, K. Medjaher, N. Zerhouni

*FEMTO-ST Institute, AS2M department, UMR CNRS 6174 - UFC /  
ENSMM / UTBM, 25000 Besançon, France (e-mail:  
diego.tobon@femto-st.fr).*

---

**Abstract:** The failure of critical components in physical systems may have negative consequences on the availability, the productivity, their security and on the environment. Thus, the assessment of the critical component's health condition, which can be done in the diagnostic and prognostic framework, should be constantly ensured. In this paper, a contribution on the assessment of the health condition of the cutting tool from a Computer Numerical Control (CNC) machine tool and the prediction of its remaining useful life before its complete failure is addressed. The proposed method is based on the use of monitoring data and relies on two main phases: an off-line phase and an on-line phase. During the first phase, the raw data provided by the sensors are processed to extract reliable features. These latter are then fed as inputs to the learning algorithms in order to generate relevant models that best represent the behavior of the cutting tool. The second phase is an assessment one, which uses the constructed models to identify the current health state and to compute the remaining useful life and the associated confidence value. The method is applied on monitoring data gathered during several cuts of the CNC tool and simulation results are given and discussed.

*Keywords:* Diagnostic, Prognostic, Remaining Useful Life, Maintenance strategies, Hidden Markov Models.

---

## 1. INTRODUCTION

The maintenance of industrial systems becomes an important task as it insures the availability, the reliability and the cost reduction. The maintenance may be corrective or predictive. In the first case, the repairing tasks are performed only after a failure is noticed, whereas in the second case, maintenance is not done systematically, but depending on the estimation of the machine's condition. The assessment of the health condition may thus play a major role as it permits to detect the current state and to track this latter to predict it in the future. In the proposed framework, diagnostic and prognostic can be used, both by exploiting the monitoring data gathered during the machine's operation. Diagnostic aims at assessing the component's current condition and identifying the cause of its failure, whereas prognostic is used to predict its future health state in order to anticipate the failure. More details can be found in (Venkatasubramanian (2005); Vachtsevanos et al. (2006); Jardine et al. (2006)).

Basically, failure prognostic consists in estimating the time before failure or Remaining Useful Life (RUL) and his confidence value. It can be realized by using three main approaches, namely: model-based prognostic, data-driven prognostic and experience-based prognostic (Lebold and Thurston (2001); Tobon-Mejia et al. (2011)). In this contribution only the data-driven prognostic is used. The choice of this approach among the others dwells in the fact that in the assessment of the condition of the cutting tools (cutters) from Computer Numerical Control (CNC) machine tools, the main problem is that deriving a behavioral

model in an analytical form that best fits the dynamic of the tool's wear is not a trivial task. Furthermore, finding experience data for a long period of time is not easy in practice. This is why the utilization of the data provided by the monitoring system may be a trade-off between the model-based prognostic and the experience-based prognostic. This paper presents a method to transform the raw monitoring data into relevant models representing the behavior of the CNC cutting tool for its whole period of utilization.

The proposed method relies on two main phases: a learning phase and an assessment phase. During the first phase, the raw data are used to extract reliable features, which are then used to learn behavioral models representing the dynamic of the degradation of the cutter. The degradation is modeled using a Mixture of Gaussians Hidden Markov Model (MoG-HMM) represented by a Dynamic Bayesian Network (DBN). This probabilistic graphical model allows to use continuous observations and also to speed up the inference by using the algorithms proposed for DBNs Murphy (2002). In the second phase, the learned model is exploited online to assess the current health of the cutter and to estimate the value of the RUL and its associated confidence value. The section 2 presents some elements required to understand the framework of the method. Then, the section 3 is dedicated to the principle and the steps of the proposed method. In section 4, the method is applied using real data (taken from PHM Society (2010)) recorded during several tool cuts and simulation results are given and discussed. Finally, a conclusion and the ongoing work are given at the end of the paper.

## 2. DATA-DRIVEN DIAGNOSTIC AND PROGNOSTIC

This section is dedicated to the presentation of the method proposed to assess the health condition of the CNC cutting tool and to estimate its remaining useful life. Before introducing the steps of the method, a brief introduction of some necessary prerequisites would help the reader understanding it. These prerequisites concerns: data clustering, Mixture of Gaussians Hidden Markov Models (MoG-HMMs), DBNs and curve fitting.

### 2.1 Data clustering

Cluster analysis is a technique used to identify groups of individuals or objects that are similar to each other but different from those in other groups. Cluster analysis is principally used by researchers in social sciences and medicine like Hillhouse and Adler (1997) who used it to identify the factors that cause stress in nurses or in biology like Landgraf et al. (2001) who used hierarchical cluster analysis to identify interfaces and functional residues in proteins. In the field of clustering two main categories of methods exist, each one has its advantages and its drawbacks (Abonyi and Feil (2007)):

- Hierarchical clustering: this is the most straightforward method of clustering, it can be either agglomerative or divisive.
- Partitional clustering: in this category, the called k-means algorithm is used, where  $k$  is the number of clusters one defines, and each case is assigned to the cluster for which its distance to the cluster mean is the smallest. The K-means algorithm is a partitioning clustering technique. This powerful tool is suitable for analyzing and visualizing spatial patterns.

### 2.2 Mixture of Gaussians Hidden Markov Models

The MoG-HMM is primarily an Hidden Markov Model (HMM). This last is a statistical model used to represent stochastic processes and in which the states are not directly observed (see Rabiner (1989)). However the outputs (observations) dependent on the hidden state are visible. Each hidden state has a probability distribution over the possible values of the observations. An example of a three states left-right HMM is shown in Fig. 1.

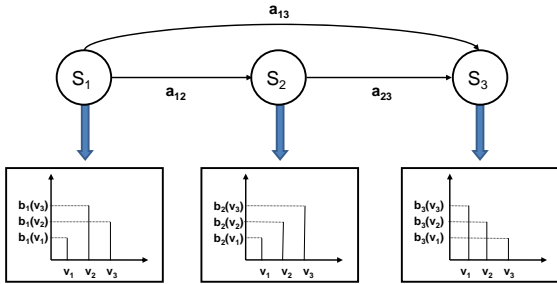


Fig. 1. A three state left-to-right HMM.

A discrete HMM is completely defined by the following parameters:

- N: number of states in the model. The individual states are  $1, 2, \dots, N$ , and the state at time  $t$  is defined as  $s_t$ .

- K: the number of distinct observations for each state. The individual observation symbols are denoted as  $V = v_1, v_2, \dots, v_K$ .
- A: the state transition probability distribution,  $A = a_{ij}$ , where  $a_{ij} = P[s_{t+1} = j | s_t = i]$ ,  $1 \leq i, j \leq N$ .
- B: the observation probability distribution of a state  $i$ ,  $B = b_i(k)$ , where  $b_i(k) = P[v_k | s_t = i]$ ,  $1 \leq i \leq N$ ,  $1 \leq k \leq K$ .
- $\pi$ : the initial state distribution  $\pi = \pi_i$ , where  $\pi_i = P[s_1 = i]$ ,  $1 \leq i \leq N$ .

For simplicity and clarity of presentation, a compact notation ( $\Theta = \pi, A, B$ ) is used for each HMM. In practice, HMMs are used to solve three typical problems (see Rabiner (1989) for more details): identification, decoding and learning.

The problem with the discrete HMMs is that in most real applications the observations are continuous signals. In this case, one can use a MoG-HMM, where each signal can be expressed as a combination of a finite number of mixtures, each one of the form:

$$b_i(O) = \sum_{m=1}^M C_{jm} \xi(O, \mu_{jm}, U_{jm}), 1 \leq j \leq N \quad (1)$$

where  $O$  is the observation vector being modeled,  $C_{jm}$  is the mixture coefficient for the  $m^{th}$  mixture in state  $i$  and  $\xi$  is a Gaussian with mean vector  $\mu_{jm}$  and covariance matrix  $U_{jm}$  for the  $m^{th}$  mixture component in state  $j$ . Similarly to an HMM, a MoG-HMM is completely defined by three parameters: the state transition matrix  $A$ , the observation matrix  $B$  and the initial probability distribution  $\pi$ . Moreover, for a MoG-HMM the observation matrix  $B$  is modeled by a Gaussian density with a mean  $\mu$ , a standard deviation  $\sigma$  and a mixture matrix  $M$ .

### 2.3 MoG-HMMs as Dynamic Bayesian Networks (DBNs)

In the last decade a new tool, namely the Dynamic Bayesian Networks (DBNs), derived from the artificial intelligence domain became popular thanks to its modeling, graphical representation and inference capabilities. Murphy (2002) presents the DBN as an extension of the traditional Bayesian network (BN) to model probability distributions over semi-infinite collections of random variables,  $Z_1, Z_2, Z_3 \dots$ . For the HMM case,  $Z$  is partitioned in  $Z = (S_t, V_t)$  to represent the hidden and the observed variables of a state space model. A DBN is defined to be a pair,  $(B1, B')$ , where  $B1$  is a BN that defines the prior probability of all variables known as  $P(Z_1)$ , and  $B'$  is a two-slice temporal Bayes net (2TBN), which defines  $P(Z_t | Z_{t-1})$  by means of a directed acyclic graph (DAG):

$$P(Z_t | Z_{t-1}) = \prod_{i=1}^N P(Z_t^i | Pa(Z_t^i)) \quad (2)$$

where  $Z_t^i$  is the  $i^{th}$  node at time  $t$ , which could be a component of  $S_t, V_t$  and  $Pa(Z_t^i)$  are the parents of  $Z_t^i$  in the graph. The nodes in the first slice of a 2TBN do not have any associated parameters, however the nodes in the second slice has associated conditional probability distributions (CPD), which defines  $P(Z_t^i | Pa(Z_t^i))$  for all  $t > 1$ .

A DBN is a powerful tool that allows modeling sequential

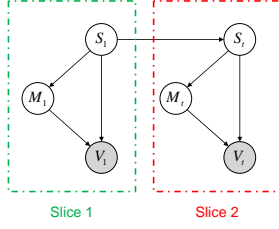


Fig. 2. A MoG-HMM represented by a DBN.

data. The main advantage in modeling HMMs by DBNs is that the inference can be done in  $O(T)$  time instead of  $O(T^3)$ , where  $T$  is the length of the data sequence, by using a junction tree algorithm (Murphy (2002)). An example of a MoG-HMM represented by a DBN is shown in Fig. 2. In this model,  $S$  represents the hidden states,  $M$  the mixture coefficients and  $V$  the observation distributions. For an illustration purpose, the corresponding CPDs for the first slice are presented below.

$$\begin{aligned} S_1 &= \pi \\ M_1 &= M_t = C_{jm} \\ V_1 &= V_t = \xi(O, \mu_{jm}, U_{jm}) \end{aligned}$$

Once the parameters of the model are completely defined, the learning algorithms for DBN can be used to correctly estimate the CPD values.

#### 2.4 Curve fitting

Curve fitting, or parameter estimation, is a numerical process that is used to represent a set of experimentally measured data points by some hypothetical analytical function Vohnout (2003). The points are constrained to a supposed polynomial function with a given order. Let us suppose  $y$  represents the output data and  $x$  the input data. By using numerical algorithms the input and output data will be fitted to the target function, which can be of  $p$  order, as shown in the following equation:

$$y(x, \beta) = \beta_0 + \beta_1 x + \dots + \beta_p x^p = \sum_{j=0}^p \beta_j x^j \quad (3)$$

where  $\beta_M$  stands for the parameters of the supposed polynomial function to be defined.

### 3. INTEGRATED DIAGNOSTIC AND PROGNOSTIC METHOD

An integrated diagnostic and prognostic method for estimating the current health state of a high-speed CNC machine tools and predicting its remaining useful life is proposed in the following of the paper. The method is a nondestructive one and relies on the utilization of the monitoring data provided by the sensors installed on the machine to track its tool's condition measured by the magnitude of the wear. In this method the amount of wear per cut is learned by using the features extracted from the monitoring data, instead of doing it by using a mathematical model of the wear for which material and environmental coefficients are not trivial to estimate. Moreover, the proposed method can be applied to assess the health state of a large range of physical components (drill bits, milling cutters, tool bits, rolling elements, etc.). The principle of the method relies on two main phases

(as shown in Fig. 3): a learning phase and an assessment phase. In the first phase, which is realized off-line, the tool's wear data provided by the monitoring sensors are first processed to extract reliable features. these latter are then fed as inputs to the k-means algorithm in order to partition them and to identify which information correspond to a specific health state of the cutting tool. The appropriate features are extracted from the raw monitoring signals by using a dedicated feature extraction procedure. In the feature matrix  $F$ , each column vector (of  $D$  features at a cut  $c$ ) corresponds to a snapshot on the raw signal, and each cell  $f_{dc}$  represents the feature  $d$  at a cut  $c$ . The matrix of features is finally used to estimate the parameters of different MoG-HMMs, where each MoG-HMM is represented by a DBN. The advantage of using several features instead of only one is that it can happen that a single feature may not capture all the information related to the behavior of the physical component.

The learned models are divided into two groups. The first group contains the general models per wear level, resumming all the data from all the learning histories and stored in a data base called "global model wear base". The second group feeds a model data base which contains one model per wear level and learning history named "local wear model base". In this way, if  $W$  are the wear stages and  $H$  the learning histories, the global model wear base contains  $W$  models (one per wear stage), whereas the local wear model base contains  $W \times H$  models (one model per wear stage and history). This latter permits to obtain the state sequence and to compute how much wear is acquired in each model state (hidden states) by using the Viterbi algorithm Viterbi (1967) to find the cuts belonging to a particular state (Fig. 4). Thus, by assuming that the wear in each state is governed by a Gaussian distribution, it is possible to estimate the mean and the standard deviation of a given wear  $(\mu(Wr_w^h(s_i)), \sigma(Wr_w^h(s_i)))$  and his variation  $(\mu(\Delta Wr_w^h(s_i)), \sigma(\Delta Wr_w^h(s_i)))$  for a particular hidden state of the modeled MoG-HMM as shown in the following equations.

$$\begin{aligned} \mu(Wr_w^h(s_i)) &= \phi = \frac{1}{T_c} \sum_{c=st}^{C_1} wr_w^h(c) \\ \mu(\Delta Wr_w^h(s_i)) &= \gamma = \frac{1}{T_c} \sum_{c=st+1}^{C_1} (wr_w^h(c) - wr_w^h(c-1)) \end{aligned} \quad (4)$$

$$\begin{aligned} \sigma(Wr_w^h(s_i)) &= \frac{1}{T_c} \sum_{c=st}^{C_1} [wr_w^h(c) - \phi]^2 \\ \sigma(\Delta Wr_w^h(s_i)) &= \frac{1}{C_1} \sum_{c=st+1}^{T_c} [(wr_w^h(c) - wr_w^h(c-1)) - \gamma]^2 \end{aligned} \quad (5)$$

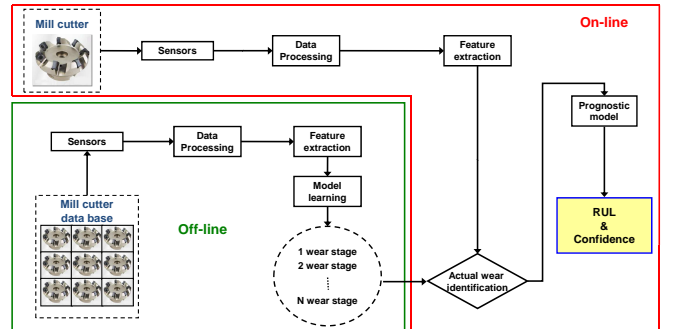


Fig. 3. The two phases of the proposed method.

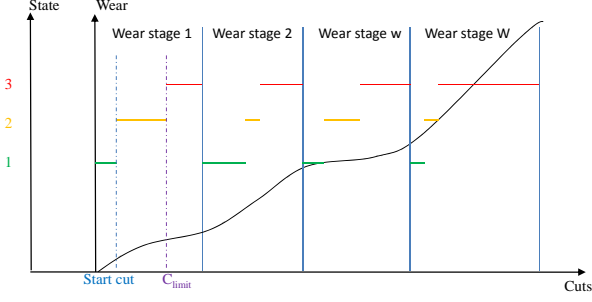


Fig. 4. A Viterbi decoding sequence.

In the previous equations  $w_r^h$  stands for the wear associated to the stage  $w = 1, \dots, W$  (from the clustering) and the history  $h = 1, \dots, H$ ,  $s_i$  is the  $i^{th}$  hidden state and  $c$  is the cut index constrained by the quantity of cuts from the cluster analysis ( $st = \text{start cut}$ ,  $C_l = C_{\text{limit}}$  and  $T_c = C_l - st + 1$ ). A compact representation of the learned model, which can then be used to perform health condition diagnostic and prognostic, is given by the following expression:

$$\lambda = \left( \begin{array}{l} DBN_w(\theta), DBN_w^h(\theta), \mu(Wr_w^h(s_i)), \\ \mu(\Delta Wr_w^h(s_i)), \sigma(Wr_w^h(s_i)), \sigma(\Delta Wr_w^h(s_i)) \end{array} \right) \quad (6)$$

In the above expression,  $\lambda$  represents the model,  $DBN_w(\theta)$  are the parameters of the DBN that capture the behavior of the wear stage  $w$  summarizing all the histories  $H$ ,  $DBN_w^h(\theta)$  are the parameters of the DBN that capture the behavior of the wear stage  $w$  from the history  $h$ ,  $\mu(Wr_w^h(s_i))$  and  $\mu(\Delta(Wr_w^h(s_i)))$  are respectively the mean wear and the mean wear variation of the state  $i$  in the stage wear  $w$  learned from the history  $h$ . In the same way,  $\sigma(Wr_w^h(s_i))$  and  $\sigma(\Delta(Wr_w^h(s_i)))$  stand for the standard deviation of the wear and the standard deviation of the wear variation in the state  $i$  for the wear stage  $w$  learned from the history  $h$ .

The on-line phase of the diagnostic and prognostic method consists in exploiting the learned models to identify the tool's current condition and to estimate the value of the corresponding RUL. The identification of the current condition is based on the evaluation of the likelihood  $P(O|\lambda)$  of the models (belonging to the global model wear base) over the observations. By knowing the current wear stage and by using the wear accumulation and the wear variation learned during the off-line phase, the tool's RUL and its associated confidence value can be calculated. The dedicated procedure to estimate the RUL and the associated confidence value during the on-line phase, based on the use of the model " $\lambda$ ", is explained through the following steps:

- The first step consists in detecting the appropriate  $DBN_w(\theta)$  that best fits and represents the on-line observed sequence of features. Indeed, the features are continuously fed to the set of learned models in the global model wear base and the likelihood is calculated in order to select the appropriate model. The selected model is then used to define the current wear stage (Fig. 4).
- The second step concerns the identification of the  $DBN_w^h(\theta)$  that best fits the observations knowing the actual wear stage  $P(O|\lambda, W)$ . The DBN inference

algorithm is thus applied on the local wear model base (belonging to the known wear stage) to compute the likelihood and to choose the best DBN. Then, the Viterbi algorithm is applied on the selected model in order to find the hidden state sequence, which permits to identify the current wear, the wear amount and the wear variation by choosing the most persistent state in the last observation sequence. This state number is stored in a global state sequence as  $G_w^h(s_i)$  in the cell  $c$ , which contains the current and the old states.

$$\text{state sequence} = (G_w^h(s_i)_1, G_w^h(s_i)_2, \dots, G_w^h(s_i)_c), \\ c = \text{current cut}$$

- The current global state sequence and the wear information  $\mu(Wr_w^h(s_i))$ ,  $\mu(\Delta(Wr_w^h(s_i)))$ ,  $\sigma(Wr_w^h(s_i))$  and  $\sigma(\Delta(Wr_w^h(s_i)))$  are used in the third step to estimate the wear vector. The idea is to estimate at each outcome the amount of mean wear and the bounds using the confidence value. The current state of each cell is compared with the precedent state (precedent cell). If the states are same the mean wear variation  $\mu(\Delta(Wr_w^h(s_i)))$  is added and the bounds are defined by using the confidence factor  $n$ , otherwise the wear and the bounds are updated. Three values of wear are estimated: the maximum ( $\widehat{Wr}_u(c)$ ), the mean ( $\widehat{Wr}_m(c)$ ) and the minimum ( $\widehat{Wr}_l(c)$ ) as shown in the following equations.

$$\widehat{Wr}_u(c) = \begin{cases} \mu(Wr_w^h(s_i)) + n \cdot \sigma(Wr_w^h(s_i)), \\ \text{If } c=1 \vee [G_w^h(s_i)]_c \neq [G_w^h(s_i)]_{c-1} \end{cases}$$

$$\widehat{Wr}_m(c) = \begin{cases} \widehat{Wr}_m(c-1) + [\mu(\Delta Wr_w^h(s_i)) + n \cdot \sigma(\Delta Wr_w^h(s_i))], \\ \text{If } [G_w^h(s_i)]_c = [G_w^h(s_i)]_{c-1} \end{cases}$$

$$\widehat{Wr}_l(c) = \begin{cases} \mu(Wr_w^h(s_i)), \\ \text{If } c=1 \vee [G_w^h(s_i)]_c \neq [G_w^h(s_i)]_{c-1} \end{cases}$$

$$\widehat{Wr}_l(c) = \begin{cases} \widehat{Wr}_l(c-1) + [\mu(\Delta Wr_w^h(s_i)) - n \cdot \sigma(\Delta Wr_w^h(s_i))], \\ \text{If } [G_w^h(s_i)]_c = [G_w^h(s_i)]_{c-1} \end{cases} \quad (7)$$

- Finally, in the fourth step the estimated wear is used to compute the RUL. This latter is obtained by using the information stored in the three vectors obtained in the previous phase  $\widehat{Wr} = (\widehat{Wr}_u, \widehat{Wr}_m, \widehat{Wr}_l)$ . Each vector is fitted to a polynomial function  $\widehat{P}(\widehat{Wr}, \beta)$  with the same order than the current wear stage ( $\beta = w$ ) to avoid linear variations and to smooth the data. The RUL is calculated by taking the difference between the wear limit ( $W_{\text{limit}}$ ) and the mean of the three smoothed wear estimations. As given in the following expression:

$$\widehat{P}(\widehat{Wr}, \beta) = \sum_{i=1}^w \beta_i \widehat{Wr}^i \\ RUL(c) = \left( W_{\text{limit}} - \frac{\widehat{P}(\widehat{Wr}_u(c), \beta) + \widehat{P}(\widehat{Wr}_m(c), \beta) + \widehat{P}(\widehat{Wr}_l(c), \beta)}{3} \right) \quad (8)$$

#### 4. APPLICATION AND SIMULATION RESULTS

The CNC machine tools are widely used in industry for achieving productivity performance goals. Statistically, 20% of the down-time of these machines is attributed to the cutting tool failure, resulting in productivity and economic losses (Kurada and Bradley (1997)). Thus, the prediction of the amount of wear before reaching the wear limit of the cutters may help improving the reliability, the availability and the safety of CNC machines, while insuring the surface roughness requirements and reducing maintenance costs.

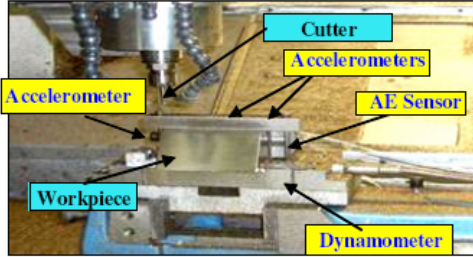


Fig. 5. The CNC milling testbed.

The wear diagnostic and prognostic method presented in the previous section is tested on the "prognostic data challenge 2010" data base taken from PHM Society (2010), which contains several histories of high-speed CNC milling machine cutters used until a significant wear stage. The authors of the experiments recorded the monitoring data from dynamometer, accelerometer and acoustic sensors during the cut process and measured the amount of wear after each cut. The experimental data correspond to three tests realized under constant conditions. The cutting parameters were: the spindle speed of the cutter 10400 *rpm*, the feed rate 1555 *mm/min*, the *Y* depth of cut (radial) 0.125 *mm* and the *Z* depth of cut (axial) 0.2 *mm*. The data were acquired at 50 *kHz/channel*. For simulation purposes three condition monitoring data are used (two for learning and one for testing), each cutter was considered as worn at the end of its associated history. The simulation results shown hereafter are obtained by choosing randomly the cutters 4 and 6 for learning and the cutter 1 for testing. During the learning phase, five wear stages are considered to define the size of the global model wear base being consequent by the analysis of the number of tool degradation regions made by Ertunc and Oysu (2004). To classify the different features in the different wear stages, the wear file record of the three flutes of each cutter are classified using the k-means algorithm. The figure 6 shows the result of the number of cuts in each cluster for the training sets. For example, for the cutter 6 the features from the first 57 cuts belong to the first wear stage, the next 137 cuts are classified in the wear stage 2, the 43 following cuts are in the wear stage 3, the 38 coming ones are classified in the wear stage 4 and the last 40 are in the wear stage 5.

For the learning and prognostic phases, different features are extracted from the raw signals provided by the monitoring sensors by following the recommendations stated in Abellan-Nebot and Romero Subiron (2010), where the relevant features used in modeling machining operations are reported. From the dynamometer signals the retained features are the root mean squares (*rms*), the peak, and

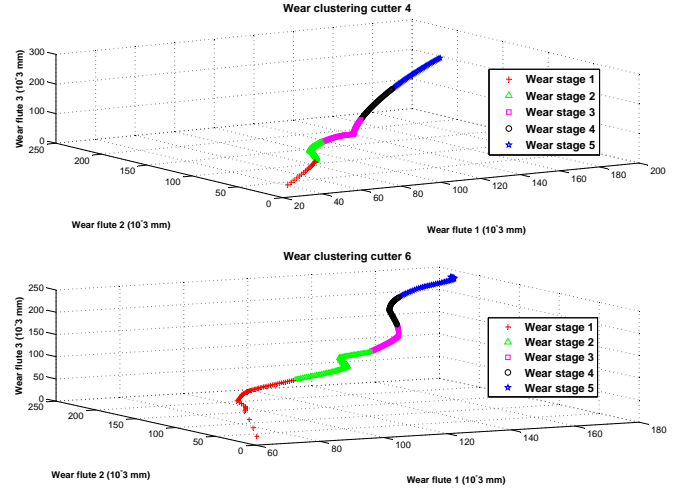


Fig. 6. The clustering results.

the standard deviation (*std*); from the accelerometer signals the *rms* and the kurtosis are computed and finally, for the acoustic emission sensor the mean and the *std* are extracted. A total of 17 features are extracted for each cut. The parameters of the DBNs in the global and in the local model wear base are first randomly initialized and constrained to a left-right MoG-HMM (to represent the degradation phenomena), and then the extracted features are fed to the learning algorithms in order to re-estimate the initialized parameters (CPDs). The number of mixtures in each MoG-HMM represented by a DBN is set to two. A total of fifteen DBNs (5 in the global model base and 10 in the local one) are obtained by using the learning algorithms of the DBN toolbox proposed by Murphy (2002). The re-estimated numerical values of the CPDs parameters related to  $\pi$ ,  $A$  and  $M$  of the MoG-HMM related to the global DBN for the first wear stage are given below:

$$\pi = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 0.85 & 0.15 & 0 \\ 0 & 0.94 & 0.06 \\ 0 & 0 & 1 \end{pmatrix}, M = \begin{pmatrix} 0.57 & 0.43 \\ 0.76 & 0.24 \\ 0.69 & 0.31 \end{pmatrix}$$

As an example the mean wear and the mean wear variation and their associated standard deviation belonging to the first wear stage learned from the history associated to the cutter 6 are presented in table 1.

Table 1. Estimated parameters in ( $10^{-3}$  *mm*)

Parameter	$s_1$	$s_2$	$s_3$
$\mu (Wr_1^6)$	26.6591	52.8704	48.7299
$\sigma (Wr_1^6)$	3.4549	18.8284	29.1340
$\mu (\Delta Wr_1^6)$	4.0446	1.5394	0.9644
$\sigma (\Delta Wr_1^6)$	0.7901	0.7869	0.4391

In order to simulate an on-line wear monitoring, the data corresponding to the cutter 1 is used as "test" history. A simulation result of the predicted mean wear after each cut is shown in Fig. 7, where the monitoring data used concerns the 315 cuts and the wear at each flute was measured. The figure 8 shows the RUL prediction, where the wear threshold is supposed known in advance. In this application a threshold limit of  $140 \cdot 10^{-3}$  *mm* is considered, which represents a small tolerance in the surface roughness of the finished part (good surface quality). In

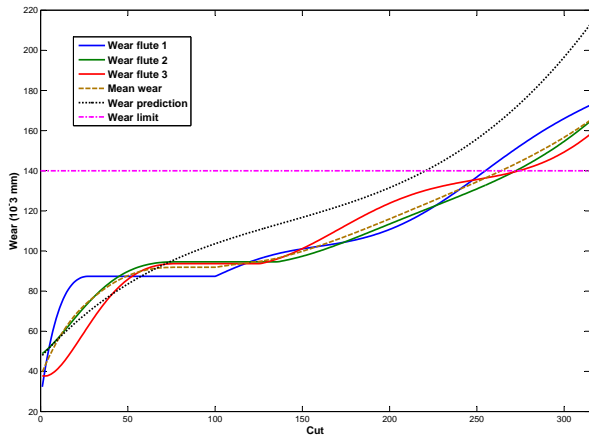


Fig. 7. Wear estimation.

Fig. 8 one can see that the precision of the estimated RUL is near to the real value. However, the prediction of the RUL value is slightly under the real value because the confidence bounds were fixed at 95%. This confidence value gives wide limits leading to an earlier prognostic, which can be useful in a sense that it suggests a tool replacement before reaching the wear limit in order to protect the work piece and the CNC machine from damage and breakdowns.

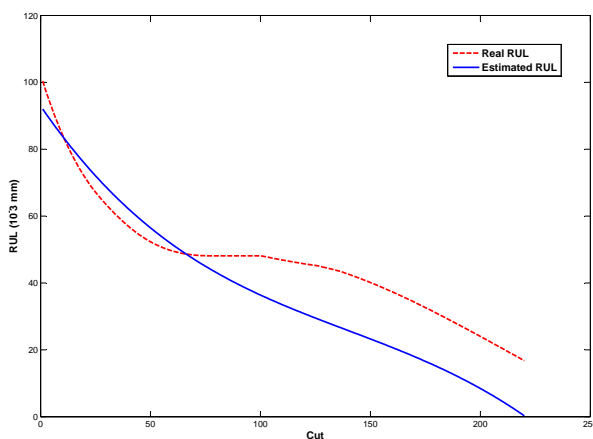


Fig. 8. The predicted RUL.

## 5. CONCLUSION

A condition monitoring, diagnostic and prognostic data-driven method has been proposed in this paper. The main idea of the work relies on the transformation of the monitoring data into relevant models that capture the degradation's behavior of the machining tools. The choice of a data-driven method instead of a physical model of the degradation dwells in the fact that obtaining this latter may not be a trivial task due to the complexity of the physical phenomenon of the wear, which is not easy to model analytically. Thus, the utilization of MoG-HMMs-DBNs allowed to model the wear in a stochastic way by taking the different stages of the degradation. The identification of the current condition of the tool combined with the models learned in the off-line phase allowed to calculate the value of the RUL.

The ongoing work concerns the identification of the optimum number of states in the MoG-HMMs and the evaluation of the method's performance using other real data sets to evaluate the suitability for real industrial applications. Also, we are working on the on-line implementation of the method on a real laboratory testbed related to bearings' degradation.

## REFERENCES

- Abellan-Nebot, J.V. and Romero Subiron, F. (2010). A review of machining monitoring systems based on artificial intelligence process models. *International Journal of advanced manufacturing technology*, 47, 237–257.
- Abonyi, J. and Feil, B. (2007). *Cluster Analysis for Data Mining and System Identification*. Birkhäuser Basel.
- Ertunc, H.M. and Oysu, C. (2004). Drill wear monitoring using cutting force signals. *Mechatronics*, 14(5), 533 – 548.
- Hillhouse, J.J. and Adler, C.M. (1997). Investigating stress effect patterns in hospital staff nurses: Results of a cluster analysis. *Social Science & Medicine*, 45(12), 1781 – 1788.
- Jardine, A.K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483 – 1510.
- Kurada, S. and Bradley, C. (1997). A review of machine vision sensors for tool condition monitoring. *Computers in Industry*, 34, 55–72.
- Landgraf, R., Xenarios, I., and Eisenberg, D. (2001). Three-dimensional cluster analysis identifies interfaces and functional residue clusters in proteins. *Journal of Molecular Biology*, 307(5), 1487 – 1502.
- Lebold, M. and Thurston, M. (2001). Open standards for condition-based maintenance and prognostic systems. In *Maintenance and Reliability Conference (MARCON)*.
- Murphy, K.P. (2002). *Dynamic Bayesian Networks: Representation, Inference and learning*. Ph.D. thesis, University of California.
- Rabiner, L.R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, volume 77 (2), 257–286.
- PHM Society (2010). PHM data challenge 2010. In <https://www.phmsociety.org/competition/phm/10>.
- Tobon-Mejia, D., Medjaher, K., Zerhouni, N., and Tripot, G. (2011). Estimation of the remaining useful life by using wavelet packet decomposition and hmms. In *IEEE Aerospace Conference*.
- Vachtsevanos, G., Lewis, F.L., Roemer, M., Hess, A., and Wu, B. (2006). *Intelligent fault diagnosis and prognosis for engineering systems*. Wiley.
- Venkatasubramanian, V. (2005). Prognostic and diagnostic monitoring of complex systems for product lifecycle management: Challenges and opportunities. *Computers & Chemical Engineering*, 29(6), 1253 – 1263.
- Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. *IEEE Transaction on Information Theory*, 13, 260 – 269.
- Vohnout, K.D. (2003). Curve fitting and evaluation. In *Mathematical Modeling for System Analysis in Agricultural Research*, 140 – 178. Elsevier Science, Amsterdam.