Robust, reliable and applicable tool wear monitoring and prognostic : approach based on an Improved-Extreme Learning Machine.

Kamran Javed, Rafael Gouriveau, Noureddine Zerhouni, Ryad Zemouri, Xiang Li

HAL Id: hal-00719545

https://hal.science/hal-00719545

Submitted on 20 Jul 2012

Robust, reliable and applicable tool wear monitoring and prognostic: approach based on an Improved-Extreme Learning Machine

Kamran Javed, Rafael Gouriveau
Noureddine Zerhouni
FEMTO - ST Institute
24 rue Alain Savary
25000 Besançon, France
Email: rafael.gouriveau@femto-st.fr

Ryad Zemouri
Laboratoire d'Automatique du CNAM
75003 Paris, France
Email: ryad.zemouri@cnam.fr

Xiang Li
Singapore Institute of
Manufacturing Technology
71 Nayang Drive, Singapore 638075
Email: xli@SIMTech.a-star.edu.sg

Abstract—Although efforts in this field are significant around the world, real prognostics systems are still scarce in industry. Indeed, it is hard to provide efficient approaches that are able to handle with the inherent uncertainty of prognostics and nobody is able to *a priori* ensure that an accurate prognostic model can be built. As for an example of remaining problems, consider data-driven prognostics approaches: how to ensure that a model will be able to face with inputs variation with respect to those ones that have been learned, how to ensure that a learned-model will face with unknown data, how to ensure convergence of algorithms, etc. In other words, robustness, reliability and applicability of a prognostic approach are still open areas. Following that, the aim of this paper is to address these challenges by proposing a new neural network (structure and algorithm) that enhances reliability of RUL estimates while improving applicability of the approach. Robustness, reliability and applicability aspects are first discussed and defined according to literature. On this basis, a new connexionist system is proposed for prognostics: the Improved-Extreme Learning machine (Imp-ELM). This neural network, based on complex activation functions, enables to reduce the influence of human choices and initial parameterization, while improving accuracy of estimates and speeding the learning phase. The whole proposition is illustrated by performing tests on a real industrial case of cutting tools from a Computer Numerical Control (CNC) machine. This is achieved by predicting tool condition (wear) in terms of remaining cuts successfully made. Thorough comparisons with adaptive neuro fuzzy inference system (ANFIS) and existing ELM algorithm are also given. Results show improved robustness, reliability and applicability performances.

I. INTRODUCTION

With recent advances in modern technology, industrials and researchers are progressing toward enhanced maintenance support systems that aim at improving reliability and availability of critical engineering assets while reducing overall costs. More precisely, it is nowadays assumed that adequate actions (either maintenance tasks, either load profile changes) must be performed in a timely manner so that, critical failures that would lead to major breakdowns or huge wastes, can be avoided. Within this framework, to fulfill such time critical needs of modern industry, condition based maintenance (CBM) appears to be a promising activity that utilizes real monitoring data to facilitate relevant indicators and trends that depict the health of the system. The uses of CBM technology enables to perform maintenance actions at the right time when needed [1]–[6]. However, the question arises that, "how to utilize the health monitoring data of an asset to determine or predict its life time"? This enhanced application of CBM can be achieved through prognostics approaches. According to ISO 13381-1, prognostic is defined as "the estimation of time to failure and risk for one or more existing and future failure modes" [7]. It is also called the "prediction of a system's lifetime" as it is a process whose objective is to predict the remaining useful life (RUL) before a failure occurs given the current machine condition and its past operation profile [3]. Prognostics should thereby enable avoiding failure risks and inopportune spending [8].

According to literature, various approaches for prognostics exist [3], [4], [6]. However, real prognostic systems to meet industrial challenges are still scarce. This can be due to highly complex and nonlinear operational environment of real industrial systems, which makes it hard to establish effective prognostics approaches, *robust* enough to tolerate uncertainty, and *reliable* enough to show acceptable performance under diverse conditions. Following that, a great attention is paid to data-driven prognostics approaches (mainly techniques from artificial intelligence - AI). Such methods aim at transforming raw in situ data into appropriate information by performing a non-linear modeling of real systems. But, the *applicability* of data-driven approaches is also limited to industrial requirements such as the complexity of implementation, human choices like initial parametrization [9], [10]. Finally, huge problems can be pointed out: *robustness*, *reliability* and *applicability* of prognostics approaches are still open areas that need to be addressed. This is the global aim of this work.

More precisely, this paper emphasizes on the proposition of an AI based approach for prognostics that, firstly is able to face with challenges of robustness and reliability, and secondly is applicable for real industrial cases. An improved-extreme learning machine (Imp-ELM) algorithm is proposed, that is

basically an extension of the existing ELM algorithm for single hidden layer feed forward neural networks [11]. In addition, due to potential benefits of ELM over its close counterpart form AI domain, performances of the proposed connexionist system are discussed by performing tests on a real industrial case of cutting tools from a Computer Numerical Control (CNC) machine for which various AI methods have already been proposed before [10], [12], [13].

The paper is organized as follows. Section II addresses prognostic modeling challenges. This enables to point out the problems of robustness, reliability and applicability of prognostics approaches, and to define them. Section III presents the proposed Imp-ELM algorithm in context to potential benefits towards enhanced prognostics. Lastly, in section IV the whole proposition is illustrated by applying it to a real industrial case of CNC machine for predicting life span (wear) of a 6 mm ball-nose 3-flutes tungsten carbide milling cutters in a high speed milling operation. Results of an Imp-ELM approach are benchmarked with ELM and adaptive neuro fuzzy inference system (ANFIS) in order to discuss performances of the approach.

II. Data-driven prognostic modeling - some challenges

A. Background and limitations of data-driven prognostics

1) From data to robustness: Data-driven prognostic methods rely on the assumption that the statistical characteristics of data are relatively unchanged unless a malfunctioning occurs. These methods aim thereby at transforming raw monitoring data into relevant information and behavior models (including the degradation) of the system. Availability of data is therefore a critical point.

However, real industrial systems are intrinsically not "perfect" and the usefulness of gathered data is highly dependent on the variability of phenomena, sensor nonlinearity, etc. Also, the degradation of an asset can not always be directly measured, so that indirect observations must be imagined. This all makes difficult for understanding (and modeling) of complex and uncertain behavior of real systems. Following that, it is obviously difficult to provide a prognostic model that is insensitive to uncertainty, and is capable of capturing dynamics of degrading asset in an accurate manner. Robustness of prognostics models appears to be an important aspect [14], and still remains a critical issue [15].

2) From robustness to reliability: Besides that, even if an approach appears to be robust enough to tolerate uncertainty, it also should be reliable enough to be used in a context that is slightly different from this one that has been considered during the training phase. In other words, the prognostic model should be able to face with variations related to the context, such as, multiple operating conditions, geometric scale or materials differences of components etc. Robustness and reliability of a prognostic approach appear to be closely related [16], and both should be considered as important to ensure the accuracy of RUL estimates.

3) From reliability to applicability: Even if a prognostic model fits with robustness and reliability criteria, it also has to be chosen accordingly to implementation requirements and constraints that restrict the applicability of the approach. Mainly, these constraints can be related to the quality and quantity of data, the generalization capability that is expected, the complexity and computational time required by the model, the human assumptions that clearly impact accuracy of results, etc. [9], [17]. This applicability problem still remains a technical challenge.

4) Synthesis of challenges addressed: According to all above, the accuracy of a prognostic model is related to the ability to face with input variations, unknown data, as well constraints that limit its applicability [9], [10], [14]. That is, under such challenges performance of the prognostic approach should be acceptable enough to accurately approximate and predict the degradation of equipment. However, for prognostics modeling, challenges like robustness, reliability and applicability are not very clear, and practitioners still face it hard to identify their links or to define them. Some definitions are presented in next section.

B. Defining robustness, reliability and applicability

According to author's knowledge, robustness, reliability and applicability of prognostic approaches are still not clearly defined in literature. Therefore, according to underlying concepts of afore mentioned works, we define them as follows.

- Robustness can be defined as the "ability of a prognostic approach to be insensitive to inherent variations of input data". In practice, a robust approach should be accurate enough to capture dynamic behavior of a degrading system, whatever the subset of the total learning frame is used.
- Reliability can be defined as the "ability of a prognostic approach to be consistent in situations when new or unknown data are presented". Reliability depicts how informative is the model, either when the estimated data has not been learned, either when data with different geometrical scale are presented.
- Applicability can be defined as the "ability of a prognostic approach to be practically applied under industrial constraints". Applicability can refer either to implementation requirements like the computation time, either to theoretical limits of the approach, either to plausibility of hypothesis.

Fig. 1 depicts the underlying dependencies among the three concepts of robustness, reliability and applicability. Enhancing prognostics should be obtained by trying to deal with all of them simultaneously. However, it is difficult to *a priori* ensure that an accurate prognostic model can be built. Also authors recognize that the definitions proposed in this paper for robustness, reliability and applicability, as well as their underlying relationships, are still not enough clear. This must be further investigated.
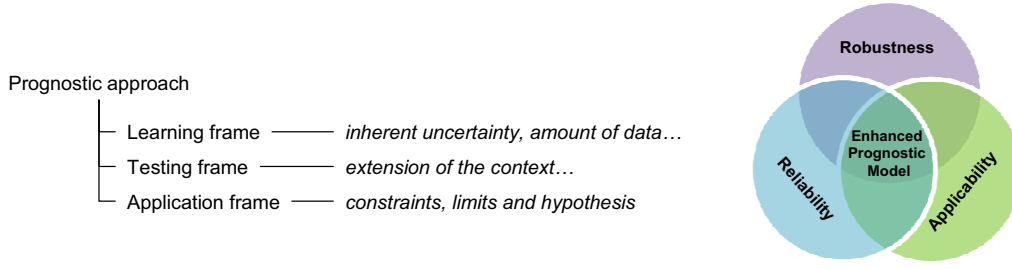
Fig. 1. Enhancing prognostics - frame and expected performances

III. ENHANCING PROGNOSTICS - THE IMP-ELM APPROACH

A. Data driven approaches for prognostics

Data driven approaches can also be referred as machine learning or data mining methods. For such approaches, input-output data is considered as the major source of information, so that behavior models can be learned directly from routinely monitored data. Indeed, data-driven approaches learn by examples and have good capability to capture complex relationship among data that are hard to describe. Also such methods are suitable for situations where physics based modeling is not favorable to replicate behavior model, or the cases for which there is an absence of prior knowledge or human expertise. Regarding prognostics, these methods are increasingly applied since they can efficiently capture nonlinear and complex relation among actual survival condition and measured condition monitoring information.

Within this category of approaches, artificial neural networks (ANNs) are currently the most widely used machine learning methods in prognostic domain [18]. In brief, an ANN is a parallel structure of data processing nodes (called as neurons), and weighted connections. In simple words, an ANN can be addressed as a multiple inputs and multiple output (MIMO) function approximator that appears to a user as a black box. In addition, this kind of systems must be tuned in order to fit the studied problem thanks to a learning phase of parameters. Although, ANNs have several advantages like simple network structures or good approximation capabilities, practitioners still have to face with several issues that may limit the applicability of ANNs for real industrial scenario [19]. As for examples, such problems can be like: hidden neurons selections, slow iterative tuning, imprecise learning rate, presence of local minima, overfitting, generalization ability, etc. However, such limitations can be overcome by recently proposed new single layer feed forward neural network algorithm called as the Extreme Learning Machine ELM [11]. Recent survey show the superiority of ELM over previous ANNs approaches [20]. Note that ELM has not been explicitly employed before in prognostic domain.

B. The Extreme Learning Machine

1) Benefits from ELM: ELM introduced by [11], is a new learning algorithm for single hidden layer feed forward neural networks (SLFNs), which is basically a structure composed of three fully connected layers namely, the input layer, the hidden layer and the output layer (see Fig.2). Previously, almost all learning algorithms for such SLFNs required adjustment of parameters that resulted dependence between different layers of parameters like, weights and biases. Therefore, many iterative tuning steps were required by learning algorithms to achieve better performance [21]. However, unlike traditional algorithms for SLFNs, ELM algorithm avoids iterative slow learning procedure and thus simply performs learning in a single-step operation. According to that, ELM algorithm requires less human involvement, because, it does not have any control parameters to be manually tuned, except the number of hidden neurons, which makes it suitable for real applications [22]. Moreover, it has been shown that, ELM has the capability to perform good enough in domain where data are scarce [23].

Fig. 2. Basic architecture of ELM [11]

2) Architecture and mathematical perspective: Consider a basic ELM architecture as shown in Fig.2, given N distinct learning data samples (x_i, t_i), where $x_i \in \Re^n$ and $t_i \in \Re^m$, therefore standard SLFN with $"L"$ hidden nodes and an activation function $g(x)$, are mathematically expressed as:

$$\sum_{k=1}^{\tilde{N}} \beta_k g(w_k . x_j + b_k) = o_j \ , j = 1, 2, ..., N \qquad (1)$$

where the coloum vector $w_k \in \Re^n$, an input weight vector to connect neurons of the input layer to k^{th} neuron of hidden

layer. In addition, w_k and x_j represent the inner product of weights and the inputs and $b_k \in \Re$ denotes the bias of kth neuron of hidden layer. Lastly, $\beta_k = [\beta_{k1}, \beta_{k2}, ..., \beta_{km}]^T \in \Re^m$, is the weight vector to connect the kth neuron of hidden layer and output neurons. According to that, to minimize the difference between network output o_j and given target t_j , there exist β_k, w_k and b_k such that:

$$\sum_{k=1}^{\tilde{N}} \beta_k g(w_k.x_j + b_k) = t_j \ , j = 1, 2, ..., N \qquad (2)$$

which can be expressed in matrix form as,

$$H\beta = T \qquad (3)$$

where

$$H\left(w_1, \ldots, w_{\tilde{N}}, x_1, \ldots, x_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}}\right)$$
$$= \begin{bmatrix} g(w_1.x_1 + b_1) & \ldots & g(w_{\tilde{N}}.x_1 + b_{\tilde{N}}) \\ \vdots & \ldots & \vdots \\ g(w_1.x_N + b_1) & \ldots & g(w_{\tilde{N}}.x_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \qquad (4)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} O = \begin{bmatrix} O_1^T \\ \vdots \\ O_{\tilde{N}}^T \end{bmatrix}_{N \times m} T = \begin{bmatrix} t_1^T \\ \vdots \\ t_{\tilde{N}}^T \end{bmatrix}_{N \times m} \qquad (5)$$

The least square solution of the linear system defined in eq. 3, with minimum norm (magnitude) of output weights β is given as:

$$\hat{\beta} = H^\dagger T \qquad (6)$$

where $H^\dagger$ represents the Moore-Penrose generalized inverse solution.

3) Learning principles: As explained before, the ELM is composed of three layers (input, hidden and output layers). It should be noted that, in ELM structure, there is no bias for the output layer neuron, which is the only difference between basic architecture of SLFN and ELM.

To begin single-step learning, the network parameters (weights and biases) are randomly assigned without any training procedure, so that, there is no dependence among parameters and input data. Consequently, after random assignment of input weights and hidden layer biases, SLFN becomes a system of linear equations. In this manner, the unknown weights linking the hidden layer to the output layer can be obtained analytically by simply applying Moore-Penrose generalized inverse operation on hidden layer output matrices [24]. Moreover, it was argued that although there are several methods to obtain generalized inverse matrix. However, singular value decomposition (SVD) is preferred for all cases [25]. As a synthesis, with such methodology ELM approach for SLFNs results faster learning speed and good generalization ability, and shows enhanced applicability as compared to other approaches for SLFNs [21].

C. Proposition of an Imp-ELM for enhanced prognostics

1) Problem formulation: As mentioned in previous section, the learning step of the ELM is quiet efficient as compared to traditional methods for SLFNs. Also it only requires the hidden neurons to be tuned. However, it has been shown that the performances of the ELM algorithm are dependent from the parameter initialization, and that this impacts the network complexity. Accordingly to the expected performances of a prognostics approach presented in section II, all this needs to be addressed.

a) The ELM avoids slow learning procedures and randomly initializes its input to hidden layer parameters. This, random parameter initialization step may affect the performance of ELM and is closely related to that [22]. Consequently, such initialization procedure of ELM may require high number of neurons in hidden layer for better performance [26]. Moreover, random initialization of network parameters, and the training of ELM with large number of hidden neurons may lead to ill-condition. That is, in such situations, the algorithm is not robust enough to capture variations in data and therefore, the estimated solution is nowhere closer to actual solution [27]. Thereby, the parameters initialization is of key importance and needs to be improved.

b) Considering the network complexity factor, i.e the number of hidden neurons, it is necessary to carefully select hidden processing node activation functions that have good capacity to handle complexity, and that can result a compact architecture of network [28]. In this context, the proposed improvements of ELM algorithm are explained in next topic.

2) Improvement: To improve the performance of ELM approach without compromising its advantages of faster speed and generalization ability, the random weight initialization procedure is replaced with well known Nguyen Widrow (NW) method [29]. Here, the assumption is to provide a good starting point by applying NW parameter adjustment procedure. This method is a simple modification of random initialization for input weights and hidden biases based on their network architecture. It has be found that, in comparison to other methods for random parameter initialization, NW algorithm has shown better performance [30]. Following that, to further optimize the performance of hidden processing layer, a complex activating function is applied, rather than traditional activation functions for ELM, i.e. sin, sigmoid, etc. [11]. It has been shown in previous works that, with complex activation functions, neural networks can perform well in real and complex domains [25], [28]. In addition, in these works, parameters of the network are also complex values. However, in our case real valued parameters are used for learning and a complex activation function in hidden processing layer to achieve better performance. In brief, such activation functions were initially proposed by Kim and Adali for a complex neural network model [31]. As for examples, consider the inverse circular function $arcsinz = \int_0^z \frac{dt}{(1-t)^{1/2}}$ or the inverse hyperbolic functions $arcsinhz = \int_0^z \frac{dt}{(1+t^2)^{1/2}}$. With real-valued input data, this type of complex activation

functions return real outputs.

Finally, the improved algorithm is based on the ELM approach with NW algorithm for random initialization of real values of input weights and hidden biases, and with a complex activation function for hidden processing layer. The learning procedure of the Imp-ELM can be synthesized as follows.

Learning requires:
- N distinct learning data samples (x_i, t_i), where $x_i \in \Re^n$ and $t_i \in \Re^m$
- A complex activation function $g_c(x)$
- $\tilde{N}$ number of hidden neurons

Imp-ELM algorithm:

Step 1: Parameter initialization
- Initialize small (random) input weights w_t and hidden neuron biases b_t
- Adjust parameters by applying NW method
  - calculate $\beta_{factor} = 0.7 \times h^{\frac{1}{n}}$
  where h represent *no.* of hidden neurons and n represent *no.* of input neurons
- Adjust parameters by dividing with their norms and multiplying with $\beta_{factor}, i.e.$
  - $w_{t+1} = \beta_{factor} \times \frac{w_t(random)}{\|w_t(norm)\|}$

Step 2: Obtain the hidden layer output matrix $\mathbf{H}$ using eq. 3

Step 3: Find the output weight matrix β using eq. 6

IV. APPLICATION ON REAL INDUSTRIAL CASE: CNC MILLING MACHINE

A. Industrial perspective

To achieve better productivity and cost-effective means to manufacture products with high surface quality, high speed machining process has become most important in manufacturing industry [32]. In this manner, CNC high speed mchining tools are most widely used in manufacturing process by industry. However, it has been found that, statistically around *20%* of the down-time of such machines is related to failure of cutting tool components, which affects productivity and result economic losses [33]. Mainly, during the cutting process, the tool-wear increases due to discontinuous and varying loads on flutes of a cutter that is repeatedly engaged and disengaged with the cutting surface [10]. More precisely, the effect of cutting tool flute wear may result imperfections in cutting surface finish and dimensional accuracy of finished parts, and can also cause damage to the work piece or machine [34]. Besides that, due to complex nature of flute wear phenomena, inconsistency and variability in cutting tool geometry/dimensions as well as the uncertainties of machining tool condition, the modeling of cutting tool performance is a difficult task to achieve [35]. Hence, it is desirable to accurately estimate the remaining life of degrading cutting tool to ensure better productivity and failure avoidance. In this paper, tool wear estimation in a high speed CNC machine is carried to evaluate our proposed Imp-ELM approach modeling performances.

B. Experimental Setup

Tests were made thanks to real data provided by Simtech Institute in Singapore, where a high speed CNC milling machine (Roders Tech RFM 760) with spindle speed 10360 rpm was used as a test-bed [12], [34]. In the machining test, the material of work-piece used was Inconel 718. Mainly, 3 cutting tools of tungsten carbide with 6 mm ball-nose/3-flutes were used in a high speed milling operation. The work-piece was cut-off from the original stock and then to get rid of its original skin layer with hard particles, the surfaces were prepared via face milling process. While milling the feed rate was 1.555 mm/min. The Y depth of cuts generated was 0.125 mm and Z depth of cuts was 0.2 mm. During the experimental phase of cutting process, the authors of experiments recorded monitoring data in terms of tool wear represented by force, acoustic and vibration properties of cutting operation. This task was performed after each cut by using dynamometer, acoustic emission and accelerometer sensors. However, the cutting force signal was used instead, because of its highly sensitivity to cutting tool wear and good measurement accuracy [34].

Therefore, to measure force signals during cutting, the dynamometer was placed between the work-piece and machining table to obtain cutting forces in form of charges and converted them to voltages. The obtained voltage signal was sampled by PCI 1200 board at 50 kKHz and directly streamed to the hard drive of a computer. The collected date was composed of 315 cuts made by 3 individual cutters during experiments. The wear patterns of 315 cuts made by 3 cutting tools namely C33, C18 and C09 are show in Fig. 3 in which the max wear of all cutters is indicated by dotted line and a wear limit know in advance is supposed. It should be noted that, cutters C33 and C18 are of same geometry but different coatings, while cutter C09 has its own geometry design and coating. Lastly, a total of 16 main features were extracted using force signals against each cut for learning of prognostic models [12], [34].
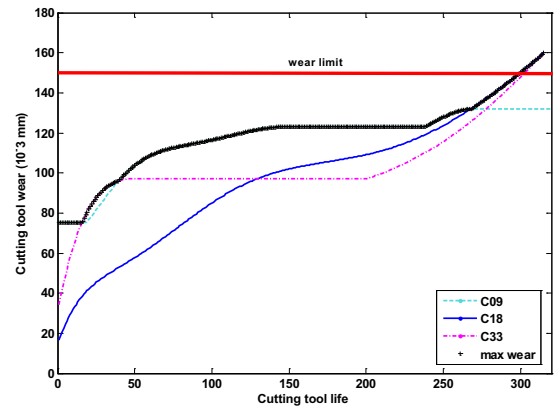

Fig. 3. Wear patterns with max wear and provided limit

C. Methodology of tool wear estimation

The growing imperfection in cutting surface finish form worn cutters can lead to rework or scrapping the work-piece or

even cause damage to the machine. Therefore, the main aim of cutting tool wear estimation is to accurately predict the remaining useful life of cutter during the machining process. That is predicting the cutting tool condition (wear) in terms of remaining cuts that could made successfully for a give wear limit. Thereby, for this application the limit of 150.10^-3 mm is considered (see Fig. 3). Further, in order to conduct tool wear estimation task, selected feature data and measured wear from cutting tool are stored into database. Thus, dataset is composed with a total of 3 cutters with 315 wear patterns from each cutting tool. Due to inherent complexity of tool wear phenomena, the simulation task is performed in a comprehensive manner by considering inherent challenges of prognostic modeling. In addition, the simulations aim to validate our proposed Imp-ELM algorithm for a considered case, and to benchmark it with ELM and ANFIS approaches. Therefore, learning of prediction models is performed in order to analyze the performances of these tools in terms of robustness, reliability and applicability (as proposed in section II). Fig. 4) depicts the whole procedure. Same data sets from cutting tool wear ware used for the three models. Simulation settings of each model are shown in Table I. D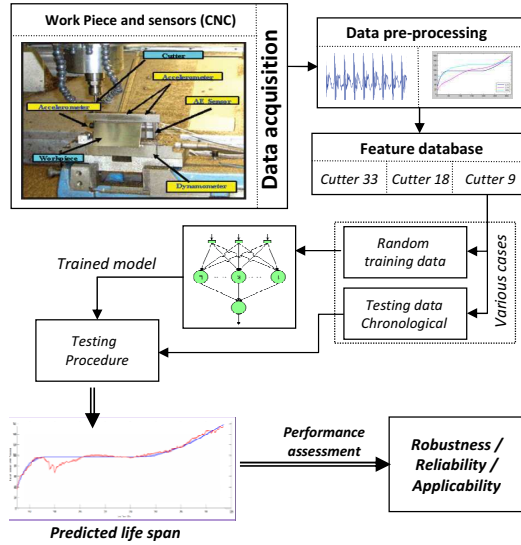uring simulation phase, different tests are performed using learned models in order to analyze their robust and reliable performance in terms of accuracy, as well as to assess their applicability by considering time and network complexity. The obtained results are mainly divided into two different cases for better understanding.

D. Case A: Single tool model - Results and discussion

This case aim at evaluating the prediction accuracy of each model by assessing their robust performance with a single cutting tool to establish a model i.e., C33, C18 and C09. Therefore, for learning purpose, training set from a particular cutting tool is created by randomly selecting half of the data samples, while rest of the data are presented in chronological order to test the accuracy of the learned model. Performances are assessed by computing the Mean Average Percent Error of prediction ($MAPE$) that should be as low as possible, the coefficient of determination (R^2) that should be closer to 1, and the computation time ($Time$), that is the required time to learn dataset. This procedure is repeated 10 times for each model-cutting tool couple, that is creating random training input data set and assessing its prediction accuracy against varying input patterns. Further, the results obtained from 10 repetitions are averaged to accumulate model performance. A comparison of tool wear prediction robustness and complexity from all model-cutting tool couples is shown in Table II. One

Fig. 4. Methodology for cutting tool wear prediction

TABLE I
PREDICTION MODELS - SETTINGS

Imp-ELM Parameters	Settings
In. / Hidden / Out. layer neurons	4 / 15 to 25 / 1
Hidden Act. function.	arcsinh
ELM Parameters	**Settings**
In. / Hidden / Out. layer neurons	4 / 70 to 80 / 1
Hidden Act. function.	sin-sig
ANFIS-Parameters	**Settings**
Input / Output layer neurons	4 / 1
Number / type of input memb. ftn	2 / Pi-shaped
FIS	First order Sugeno
Training Algorithm	Hybrid Method

TABLE II
TEST CASE A: ANALYZING ROBUSTNESS PERFORMANCE FOR PREDICTIONS WITH SINGLE TOOL MODEL

Cutter 33					
			Accuracy		
$Method$	$Feature$	$Train/Test$	$MAPE$	R^2	$Time(s)$
$Imp-ELM$	Force	150/165	**0.047**	**0.89**	**0.02**
$ANFIS$	Force	150/165	0.050	0.83	0.70
ELM	Force	150/165	0.064	0.41	0.03
Cutter 18					
			Accuracy		
$Method$	$Feature$	$Train/Test$	$MAPE$	R^2	$Time(s)$
$Imp-ELM$	Force	150/165	**0.064**	**0.95**	**0.02**
$ANFIS$	Force	150/165	0.081	0.92	0.69
ELM	Force	150/165	0.121	0.75	0.03
Cutter 09					
			Accuracy		
$Method$	$Feature$	$Train/Test$	$MAPE$	R^2	$Time(s)$
$Imp-ELM$	Force	150/165	**0.035**	**0.82**	0.02
$ANFIS$	Force	150/165	0.040	0.77	0.71
ELM	Force	150/165	0.040	0.78	**0.01**

can note that, for all cutters with varying input data, Imp-ELM model with compact architecture show robust performance as indicated by its testing accuracy. Moreover, with given data set ELM and Imp-ELM models can learn faster as compared to ANFIS model. A comparative plot of tool wear prediction from all models using cutter C18 are shown as an example in Fig 5.

E. Case B: Multi-tools model - Results and discussion

This case investigates multi-tools model performances with two different testing methodologies. **Step 1.** In this step, model
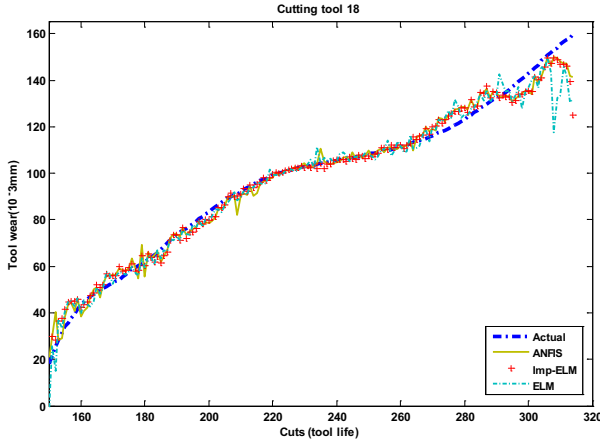
Fig. 5.   Tool wear predictions for a single tool model - robustness illustration
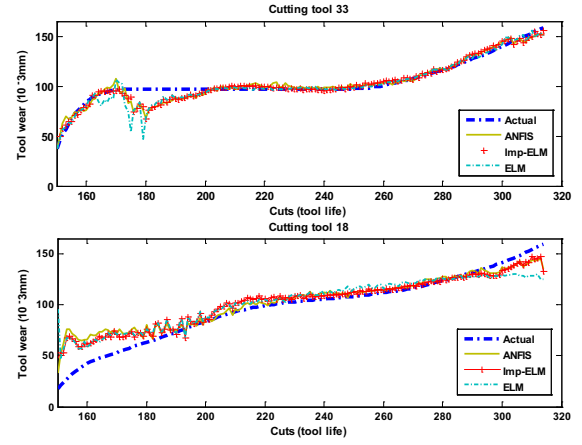

Fig. 6.   Tool wear prediction for reliability performance with 2-3 cutters

reliability performance is investigated by establishing a multi-tools model. That is, half of the data sets for training are selected from different cutters and rest of the data from one of these cutters are presented in chronological order to test the model. The main aim for this procedure is to analyze the reliability prediction capabilities of the particular model when it is learned with multiple cutting tools. The simulations are repeated 10 times in similar manner like the previous case and averaged to see accumulated response of the model. A comparative analysis on reliability performance from all models is summarized in Table III.

It can be observed from the above results that, although the accuracy factor decreases for all models, but still the performance of Imp-ELM model is good enough. In addition, it can learn much faster as compared to ANFIS model, even if the data size is increased or doubled. Tool wear prediction for reliability performance by learning a model with 2-3 cutters and testing on single cutting tool are shown in Fig 6.

performances on another cutter that has not been learned. The learning set is composed of all randomly selected data samples from two different cutters, while testing data from a third cutter is presented in a chronological order. Performances of each multi-tools model build by learning cutter 33 and 09 and testing on cutter 18 are shown in Table IV.

Imp-ELM shows better results on unknown data as compared to other models. However, when cutters 33 and 09 are presented as unknown data to the prediction model (these results are not given in the paper), the model decreases to a poor level. Indeed, results of Imp-ELM for cutter 18 are better because the scale of its wear pattern exist between cutters 33 and 09 as wear increases (Fig. 3). In addition, as stated before, cutters 33 and 18 have same geometry while cutter 09 has its own geometry. Eventually, comparative plots form this case using cutter 18 are shown in Fig 7 that enables to see that reliability has still to be improved.

TABLE IV
TEST CASE B: ANALYZING RELIABILITY PERFORMANCE FOR
PREDICTIONS ON UNKNOWN DATA

Cutters 33-09 (train) / Cutter 18 (test)					
			Accuracy		
$Method$	$Feature$	$Train/Test$	$MAPE$	R^2	$Time(s)$
$Imp-ELM$	Force	630/315	**0.23**	**0.66**	**0.02**
$ANFIS$	Force	630/315	0.29	0.473	3.01
ELM	Force	630/315	0.30	0.45	0.02

V. CONCLUSION

In this work the inherent challenges of prognostic modeling like robustness, reliability and applicability are explicitly addressed and defined according to literature. On this basis, a new Imp-ELM algorithm is developed for predicting the cutting tool condition from high speed CNC machine. The viability of Imp-ELM approach is evaluated using force-based approach to build a prediction model. Simulations are performed with limited data of three worn cutters form cutting process. Comparative analysis show that with less complex

TABLE III
TEST CASE B: ANALYZING RELIABILITY PERFORMANCE FOR
PREDICTIONS WITH A MULTI-TOOLS MODEL

Cutters 33-18(train)/Cutter 33(test)					
			Accuracy		
$Method$	$Feature$	$Train/Test$	$MAPE$	R^2	$Time(s)$
$Imp-ELM$	Force	300/165	**0.052**	**0.83**	**0.02**
$ANFIS$	Force	300/165	0.063	0.77	1.4
ELM	Force	300/165	0.069	0.43	0.02
Cutters 33-18-09 (train)/ Cutter 18 (test)					
			Accuracy		
$Method$	$Feature$	$Train/Test$	$MAPE$	R^2	$Time(s)$
$Imp-ELM$	Force	450/165	**0.16**	**0.84**	**0.02**
$ANFIS$	Force	450/165	0.18	0.79	2.01
ELM	Force	450/165	0.20	0.66	0.02

Step 2. In this step, reliability performance of each model is evaluated by establishing a reference model through learning on two different cutters and testing its wear prediction
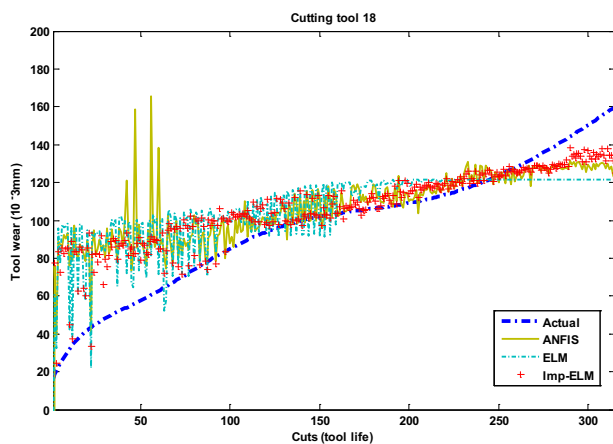
Fig. 7. Tool wear predictions for a multi-tools model - poor reliability illustration on an unknown cutter

network architecture, Imp-ELM performs better than ELM and ANFIS algorithms in predicting tool condition with single and muli-tool models. However, performances of learned models were not good enough when unknown data is presented. Thus, model reliability still needs to be improved.

REFERENCES

[1] C. Byington, M. Roemer, G. Kacprzynski, and T. Galie, "Prognostic enhancements to diagnostic systems for improved condition-based maintenance," in *IEEE Aerospace Conf., Big Sky, USA*, 2002.

[2] M. Dong, "A tutorial on nonlinear time-series data mining in engineering asset health and reliability prediction: concepts, models, and algorithms," *Mathematical Problems in Engineering*, vol. 10, 2010.

[3] A.-K.-S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Systems and Sig. Processing*, vol. 20, pp. 1483–1510, 2006.

[4] A. Heng, S. Zhang, A.-C.-C. Tan, and J.Matwew, "Rotating machinery prognostic:state of the art,challenges and opportunities," *Mech. Systems. and Sig. Processing.*, vol. 23, no. 3, pp. 724–739, 2009.

[5] M. Lebold and M. Thurston, "Open standards for condition-based maintenance and prognostics systems," in *5th Annual Maintenance and Reliability Conference*, 2001.

[6] G. Vachtsevanos, F.-L. Lewis, M. Roemer, and A. Hess, *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. New Jersey, Hoboken: Wiley & Sons, 2006.

[7] ISO, 13381-1, *Condition monitoring and diagnostics of machines - prognostics - Part1: General guidelines*. Int. Standard, ISO, 2004.

[8] A.-S.-Y. Heng, "Intelligent prognostics of machinery health utilizing suspended condition monitoring data," Ph.D. dissertation, Queensland university,Austrailia, 2009.

[9] M. El-Koujok, R. Gouriveau, and N. Zerhouni, "Reducing arbitrary choices in model building for prognostics: An approach by applying parsimony principle on an evolving neuro-fuzzy system," *Microelectronics Reliability*, vol. 51, pp. 310–320, 2011.

[10] S. Das, R. Hall, S. Herzog, G. Harrison, and M. Bodkin, "Essential steps in prognostic health management," in *IEEE Conference on Prognostics and Health Management*, Denver, CO, USA, 2011.

[11] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *International Joint Conference on Neural Networks*, Budapest, Hungary, 2004.

[12] X. Li, B.S. Lim, J.H. Zhou, S. Huang, S.J. Phua, and K.C. Shaw, "Fuzzy neural network modelling for tool wear estimation in dry milling operation," in *Annual Conference of the Prognostics and Health Management Society*, San Diego, CA, USA, 2009.

[13] A.-J. Vallejo, R. Morales-Menendez, and J.-R. Alique, "On-line cutting tool condition monitoring in machining processes using artificial intelligence," *Robotics Automation and Control*, pp. 143 – 166, 2008.

[14] L. Liao, "An adaptive modeling for robust prognostics on a reconfigurable platform," Ph.D. dissertation, University of Cincinnati, 2010.

[15] F. Camci and R.-B. Chinnam, "Health-state estimation and prognostics in machining processes," *IEEE Transactions on automation science and engineering*, vol. 7, pp. 581–597, 2010.

[16] Y. Peng, M. Dong, and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," *Int J Adv Manuf Technology*, vol. 50, pp. 297–313, 2010.

[17] J.Z. Sikorska, M.Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Journal of Mechanical Systems and Signal Processing*, vol. 26, no. 5, pp. 1803 – 1836, 2011.

[18] R. Zemouri, R. Gouriveau, and N. Zerhouni, "Combining a recurrent neural network and a pid controller for prognostic purpose:a way to improve the accuracy of predictions," *WSEAS Transactions on Systems and Control*, vol. 5, 2010.

[19] R. Singh and S. Balasundaram, "Application of extreme learning machine method for time series analysis," *International Journal of Intelligent Technology*, vol. 2, pp. 256–262, 2007.

[20] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: a survey," *International Journal of Machine Learning and Cybernetics*, vol. 2, no. 2, pp. 107–122, 2011.

[21] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, pp. 489–501, 2006.

[22] A. U. Bhat, S. Merchant, and S. S. Bhagwat, "Prediction of melting point of organic compounds using extreme learning machines," *Industrial and Engineering Chemistry Research*, vol. 47, no. 3, pp. 920–925, 2008.

[23] E. A. El-Sebakhy, "Extreme learning machine as a new framework in predicting material properties: Methodology and comparison," in *12th Int. Conf. of International Association for Computer Methods and Advances in Geomechanics (IACMAG)*, Goa, India, 2008.

[24] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and its Applications*. John Wiley and Sons, New York, 1971.

[25] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully complex extreme learning machine," *Neurocomputing*, vol. 68, pp. 306–314, 2005.

[26] R. Rajesh and J. S. Prakash, "Extreme learning machines - a review and state-of-the-art," *International journal of wisdom based computing*, vol. 1, pp. 35–49, 2011.

[27] G. Zhao, Z. Shen, C. Miao, and Z. Man, "On improving the conditioning of extreme learning machine: a linear case," in *7th international conference on Information, communications and signal processing ICICS'09*, Piscataway, NJ, USA, 2009.

[28] H. A. Jalab and R. W. Ibrahim, "New activation functions for complex-valued neural network," *IEEE Trans. Systems, Man, Cybernetics*, vol. 6, pp. 1766–1772, 2011.

[29] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *Int. Joint Conf. on Neural Networks IJCNN*, San Diego, CA , USA, 1990.

[30] A. Pavelka and A. Prochazka, "Algorithms for initialization of neural network weights." [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.115.2073

[31] T. Kim and T. Adali, "Fully complex multi-layered preceptron network for nonlinear signal processign," *Journal VLSI Signal Processing*, vol. 32, pp. 29–43, 2002.

[32] L.Y. Zhai, M.J. Er, X. Li, O.P. Gan, S.J. Phua, S. Huang, J.H. Zhou, S. Linn, and A.J. Torabi, "Intelligent monitoring of surface integrity and cutter degradation in high-speed milling processes," in *Annual Conference of the Prognostics and Health Management Society*, Portland, Oregon, USA, 2010.

[33] D.A. Tobon-Mejia, K. Medjaher, and N. Zerhouni, "CNC machine tool health assessment using dynamic bayesian networks," in *18th IFAC World Congress*, Milano, Italy, 2011.

[34] J. Zhou, X. Li, O.P. Gan, S. Han, and W.K. Ng, "Genetic algorithms for feature subset selection in equipment fault diagnostics," *Engineering Asset Management*, vol. 10, pp. 1104–1113, 2006.

[35] A.J. Torabi, M.J. Er, X. Li, B.S. Lim, L.Y. Zhai, S.J. Phua, L. San, S. Huang, and J.T.T. Tijo, "A survey on artificial intelligence technologies in modelling of high-speed end-milling processes," in *International Conference on Advanced Intelligent Mechatronics*, Singapore, 2009.

Kamran Javed received his BS. computer engineering degree from COMSATS University Abbottabad (Pakistan) in 2006. He was honored by a fully funded scholarship to complete his Masters in computer engineering from COMSATS, 2009. During his MS. he also worked as a research associate for developing Power system Planning Suit for WAPDA organization Pakistan, which mainly works for Hydel Power and Water Sector projects. In 2009 he joined COMSATS for a lecturer position in electrical engineering department. Currently he performs his doctoral studies at FEMTO-ST (Franche-Comte University), in Besançon (France) that is funded by Ministry of France since 2011. His research activities are concerned with developing intelligent prognostics systems using connectionist methods like extreme learning machine, neural networks, and neuro-fuzzy methods.

Dr. **Rafael Gouriveau** is an IEEE member. He received his engineer degree from National Engineering School of Tarbes (ENIT) in 1999. He then got his MS (2000) and his Ph.D. in Industrial Systems in 2003, both from the Toulouse National Polytechnic Institute (INPT). During his PhD, he worked in the field of risk management and dependability analysis. In September 2005, he joined the national high school of mechanics and microtechniques of Besanon (ENSMM) as Associate Professor. His main teaching activities are concerned with production, maintenance, manufacturing, and informatics domains. As for investigation, Rafael Gouriveau is with the AS2M (Automatic Control and Micro-Mechatronic Systems) department of Femto-st Institute. He is currently at the head of PHM team from this department. His research interests concern nowadays the development of industrial prognostics systems by using connexionist approaches like neuro-fuzzy methods, and the investigation of reliability modeling by using possibility theory.

Prof. **Noureddine Zerhouni** received his engineer degree from National Engineers and Technicians School of Algiers (ENITA) in 1985. After a short period in industry as engineer, he received his Ph.D. Degree in Automatic Control from the Grenoble National Polytechnic Institute in 1991. In September 1991, he joined the National Engineering School of Belfort (ENIB) as Associate Professor. At this time, his main research activity was concerned with modelling, analysis and control of manufacturing systems. Since September 1999, Noureddine Zerhouni is Professor at the national high school of mechanics and microtechniques of Besanon. As for investigation, he is with at the head of AS2M (Automatic Control and Micro-Mechatronic Systems) department of Femto-st Institute. His main research activities are concerned with intelligent maintenance systems and e-maintenance. Professor Noureddine Zerhouni has been and is involved in various European and National projects on intelligent maintenance systems like FP5 European Integrated Project of ITEA program (Information Technology for European Advancement) PROTEUS, NEMOSYS (Naval E-Maintenance Oriented SYStem) with DCNS, and AMIMAC-FAME (Reliability Improvement of Embedded Machines) with ALSTOM.

Dr. **Ryad Zemouri** received his engineer degree at Mouloud Mammeri University of Tizi-Ouzou(Algeria) in 1998 and his Ph.D. in automatic control and computer science in Franche-Comte University in2003. He joined the Conservatoire National des Arts & Mtiers (CNAM) of Paris in September 2003 where he works actually as Associate Professor. His main research interests are computational intelligence, pattern recognition, neural networks, and industrial prognostics systems.

Dr. **Xiang Li** received her Ph.D. in 2000 from Nanyang Technological University of Singapore. She now forms part of the Singapore Institute of Manufacturing Technology and has more than 15 years of experience in research on artificial intelligence, data mining, machine learning and statistical analysis, such as neural networks, fuzzy logic systems, fuzzy neural networks, unsupervised data clustering and regression modeling. She has been jointly awarded together with NTU several grants under the A*STAR SERC. She led several projects to develop intelligent systems for tool condition monitoring, equipment/process performance prediction, multivariate anomaly detection for product quality control, manufacturing execution & decision support system, data warehousing, intelligent optimization & simulation, web-based intelligent forecasting and discovering customer demand for new product design. Her research interests include data mining and knowledge discovery, decision support systems, prognostic health management. She is a member of the Decision Sciences Institute, USA and member of IEEE.