2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)
Sept 14-16, 2015. San Diego, CA, USA

Wavelet Based Sensor Fusion for Tool Condition Monitoring of Hard to Machine Materials*

Farbod Akhavan Niaki, Durul Ulutan, and Laine Mears

***Abstract*— Tool condition monitoring in modern manufacturing systems is gaining more attention due to the fact that excessive tool damage can cause workpiece surface deterioration and increase idle time. Therefore, monitoring tool condition from the initial to final stages of tool life is a task that is critical yet difficult, especially in hard-to-machine materials. In this work, Wavelet Packet Decomposition is used for extracting statistical features in the time-frequency domain of two low cost sensing technologies, *i.e.* vibration and power, in addition to Principal Component Analysis to reduce the dimensionality of feature vectors. A Recurrent Neural Network is then trained with Bayesian regularization backpropagation method and the estimated tool wear is compared to the actual measured wear. Results show a maximum of 13% relative error in estimating tool wear which proves the effectiveness of implemented sensory data fusion method to be used in automated control of manufacturing processes.**

I. Introduction

In automated manufacturing systems, accurate estimation and monitoring of important states is a critical factor to reduce downtime and avoid catastrophic failure. Tool wear in difficult-to-machine materials (those of excessive strength compared to tooling material, poor thermal conductivity or rapid work hardening) is playing an important role in machining downtime and the quality of the resultant product. For these materials, the wear rate of inserts during the machining process is relatively high compared to conventional materials. High tool wear rate while machining makes establishing an accurate tool wear model a challenging task because only a limited number of experiments can be completed before the tool fails. Moreover, tool wear is a highly complex phenomenon that consists of different mechanisms such as abrasion, adhesion, chipping and plastic deformation [1]. The model for each separate mechanism is either not developed yet or developed under certain machining conditions for a particular type of material, which cannot be generalized for other materials or machining conditions.

Tool wear studies can be divided into two major categories: First, is the model-based studies that use either pre-developed analytical models for certain types of tool wear; these include the Takeyema model of flank wear [2] and the Usui model of crater wear [3], or numerical models derived based on finite element analysis [4-5]. The second category of tool wear studies is data-driven modeling that relies on empirical interpretation of the input data, and therefore requires a training set. The advantage of these techniques is observed mostly where a process model is not available. This feature is particularly useful for studying tool wear of hard to machine materials due to the lack of models proposed in understanding it [6].

Cutting forces, spindle power consumption while cutting, vibration, and acoustic emission (AE) are all shown to be correlated to the tool wear [7]. The force signal is the most widely used measurement signal in tool wear studies. Despite being precise, the measurement device (*i.e.* dynamometer) needs continuous calibration, limits the workpiece size, and is expensive that makes its implementation in industrial machining shops very difficult [8-10]. AE signals have been shown to have a good capability in studying tool wear; however acquiring these signals requires powerful data acquisition systems with enough memory storage [8]. On the other hand, Hall effect sensors that are used in power sensing technology and accelerometers are relatively inexpensive, they can be easily mounted in the machine, and they do not limit the size of the workpiece, which make them suitable candidates for machining performance assessment in industrial applications. Due to the differences in the nature of sensors, each can capture different information from the machine. Thus, by using sensor fusion methods, it is possible to extract more informative information from the signals [11-12]. Moreover, it has been shown that time-frequency analyses such as Continuous Wavelet analysis, Discrete Wavelet analysis, or Wavelet Packet Decomposition can provide valuable information about the state of the tool in different machining operations [13].

The objective of this study is to investigate the applicability of sensory information fusion using low-cost sensing technology, particularly spindle power and vibration for tool condition monitoring in milling hard-to-machine alloys. The organization of this work is as follows: theoretical background of Wavelet Packet Decomposition, Principle Component Analysis, and Recurrent Neural Network are briefly introduced in section II. Experimental setup and selected cutting conditions are explained in section III. In section IV, the processing of the data with WPD, feature reduction and results are discussed, and conclusions are provided in section V.

II. Mathematical Background

A. Wavelet Packet Decomposition

Depending on the type of application, spectral information of the measured signal may be useful. Time series analysis of a given signal only provides information

*Research supported by National Science Foundation (NSF).

Farbod Akhavan Niaki is with the Clemson University, International Center for Automotive Research (CU-ICAR), Greenville, SC 29607 USA (phone: 864-569-4396; e-mail: fakhava@ clemson.edu).

Durul Ulutan, is with the Bucknell University, Mechanical Engineering Department, Lewisburg, PA 17837 USA (e-mail: du005@bucknell.edu).

Laine Mears is with the Clemson University, International Center for Automotive Research (CU-ICAR), Greenville, SC 29607 USA (e-mail: mears@clemson.edu).

978-1-4799-7772-7/15/$31.00 ©2015 IEEE

over different time horizons, while spectral content of the signal is lost in the time domain. On the other hand, the Fourier transform converts the signal from time domain into frequency domain by multiplying it with an infinite series of sine functions. While being a powerful tool for extracting the spectral content of the signal, Fourier Transform does not provide any information about the evolution of frequencies in time. This is particularly important when studying non-stationary signals such as tool wear, where sudden peaks or anomalies emerge. To study a signal in the time-frequency domain, the wavelet transform has been proposed as an alternative. In wavelet analysis, a small wave (wavelet) function is stretched and scaled to capture the time-frequency content of the signal. The continuous wavelet transformation is shown in (1), where ψ is the mother wavelet function, the superscript star sign "*" denotes the complex conjugate, τ is the shift in time, s is the scale factor (which can be interpreted as the change in frequency content of the mother wavelet), and X_{WV} is the magnitude of transformed signal in time frequency domain. The term $1/\sqrt{s}$ is the normalizing constant used for different scales.

$$X_{WV}(\tau,s)=\frac{1}{\sqrt{s}}\int x(t)\psi^{*}\left(\frac{t-\tau}{s}\right)dt \qquad (1)$$

Since any signal is measured as a series of discrete events, (1) should be reshaped in discrete format. So, special digital filters known as Quadrature Mirror Filters (QMF) combined with downsampling and upsampling by a factor of 2 are used for discrete wavelet analysis. To decompose the signal, it is first passed through the low/high pass filter, then every other sample is dropped to avoid aliasing [14]. The output of the downsampled signal from the low pass filter is a denoised version of the actual signal, which is called the *approximation*, and the output of the downsampled signal from the high pass filter contains noise information of the actual signal, which is called the *detail*. The combination of all four filters used for decomposition and reconstruction is called the *filter bank*. Wavelet Packet Decomposing (WPD) decomposes the signal into multiple levels of approximation and detail packets (Figure 1). WPD is particularly useful when an anomaly such as tool breakage and tool chipping occurs at high frequency ranges. In that case, high frequency coefficients (detail coefficients) are as important as low frequency coefficients (approximation coefficients).

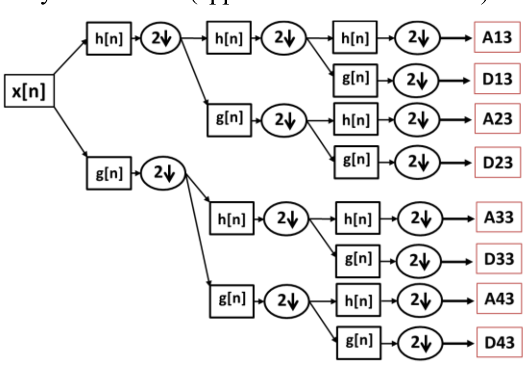

Figure 1. Cascade diagram for 3-level WPD with $g[n]$ as high pass and $h[n]$ as low pass filter

B. Principal Component Analysis

In any machine-learning algorithm, selection of features is the most critical part. If the features cannot capture the essential characteristics of the system, the trained algorithm will not be able to represent the actual system. Moreover, the selection of appropriate features reduces the cardinality of the training set and may reduce the run-time of the algorithm depending on the complexity and degree of nonlinearity of the problem. Selecting as many features as possible seems intuitively appealing; however, increasing the number of features can increase the computational cost significantly. Moreover, new features do not necessarily bring additional information and may be collinear and therefore redundant. The undesired effect of a large feature set, which is known as curse of dimensionality [15], can be overcome by conducting a feature reduction method called Principal Component Analysis (PCA), where the most related features with the highest variation can be identified.

Considering the $m \times n$ training set matrix X, the covariance of X is represented by the $n \times n$ matrix Σ. The eigenvector matrix of Σ, known as modal matrix M in (2), is used to decouple the features of the training set. Eigenvalues of the system are defined as the diagonal elements of the matrix E in this equation. Each eigenvalue corresponded to each eigenvector represents the variation of transformed feature in the new decoupled coordinate system. Therefore, dimensionality reduction can be achieved by simply dropping the lowest eigenvalue and the corresponding eigenvector from matrix M. The new modal matrix is now an $n \times j$ matrix where $j<n$. By multiplying the $m \times n$ training set by the new $n \times j$ transformation matrix, training set with reduced features is found.

$$E=M^{-1}\Sigma M \qquad (2)$$

C. Recurrent Neural Network Architecture

Neural Network (NN) algorithms are designed based on the learning behavior of neurons in the human brain [15]. Recognition with NNs has the advantage that it eliminates the need for the model. Therefore, Neural Networks are the best (and maybe last) resort where a model of a particular system is not available or hard to develop. The fact that no model is required for NN algorithms makes the role of input data very important, so that the trained NN with a weak training set cannot produce accurate results. In this study, the Nonlinear Auto Regressive with eXogenous input (NARX) model was selected as the NN architecture. In the NARX model, there exists a feedback with time delay from the outputs to the inputs of the NN, so that outputs are added as new features to the inputs. As shown in Figure 2, the NARX model consists of 3 layers: (a) The first layer is called the input layer, (b) the middle layer is called the hidden layer which may consist of more than one layer with different number of neurons, and (c) the last layer is the output layer. The number of hidden layers and hidden neurons are design parameters of the NN, and should be determined by trial and error to reach the minimum error without overfitting the training set.

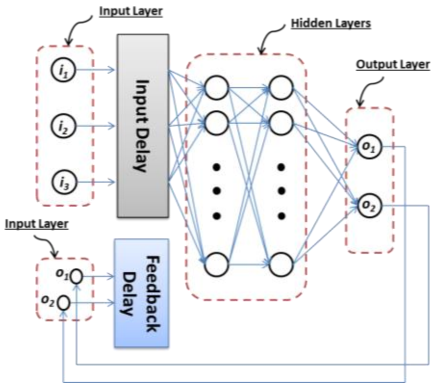

Figure 2. NARX Neural Network Schematic

The backpropagation method is the most widely used training technique for adjusting the weights of NN. The aim of backpropagation is minimizing the mean squared errors of NN's output using gradient descent method. However, overfitting of data can happen when using backpropagation. The Bayesian regularization training method was proposed to avoid overfitting [16]. In this method, an additional term related to sum of square of NN's weight is added to the sum of square error of outputs and targets as shown in (3), where E_T is the total sum of errors weighted by parameters α and β. E_1 is the sum of square errors of targets and outputs, and E_2 is the sum of square of NN weights. The objective is to minimize the new error term E_T. In Bayesian regularization algorithm, it is assumed that NN weights are random variables with known prior, which is updated at each iteration.

$$E_T = \alpha E_1 + \beta E_2 \tag{3}$$

III. Experimental Setup

As shown in Figure 3, a total of 8 experiments with 3 replications were conducted to study the tool wear when machining the gamma-prime strengthened alloy Rene-108 (R-108). Among these three replications, the first two were chosen as training sets, and the last replication was chosen for validating the results (*i.e.* the testing set). An OKUMA GENOS M460-VE 3-axis CNC machine was used to end mill (in the down-milling direction) rectangular blocks with dimensions of 60×80×25 mm, using a water soluble coolant. A 2-flute indexable tool holder with a diameter of 15.875 mm was used, and the width of cut was chosen to be 9.5 mm that corresponds to 60% tool engagement, as this was the maximum manufacturer recommendation for the particular tool holder. Full length of the blocks (60 mm) was utilized for machining, which was considered as a "pass." Depth of cut, cutting velocity, and feed are kept constant for all passes at 0.5 mm, 25 m/min and 0.1 mm/rev, respectively. The cutting conditions are determined based on the industrial applications targeted by this study, and keeping the cutting conditions constant is due to the fact that a change in the cutting conditions can induce abrupt changes in the behaviour of hard-to-machine alloys. Two different sensors were utilized in this work. The Hall effect sensor was mounted on the electric panel of the CNC machine to capture the change in current consumed by the spindle, and accelerometer is mounted on the spindle case to capture the induced vibration between tool and workpiece during the milling operation.

A data acquisition device (DAQ) was programmed to capture the spindle power consumption and vibration at a high sampling rate during cutting. To measure spindle power, the output of the sensor transducer was fed into the NI9215 analog input module, and to measure vibration, the output of the Kistler 8772A accelerometer was fed into the NI9234 analog input module mounted on an NI-cRIO9103 chassis programmed with LabVIEW (Figure 4). Data were collected at a sampling frequency of 10.24 kHz.

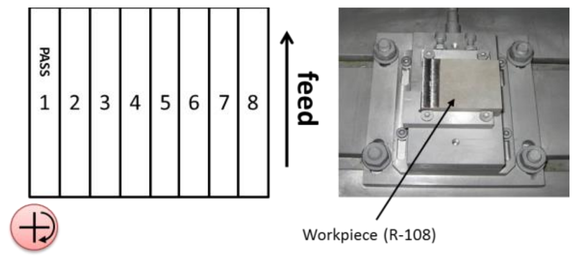

Figure 3. Schematic of milling experiments

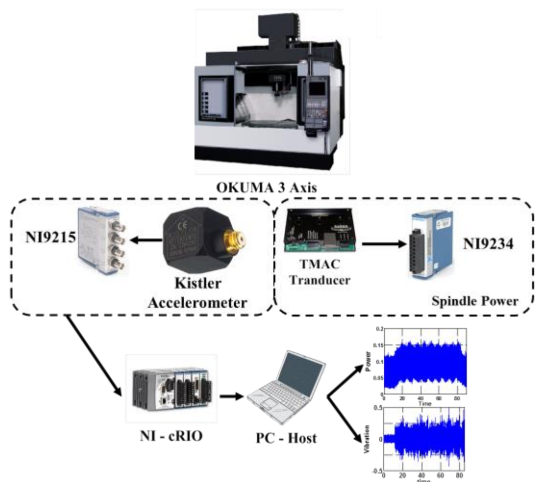

Figure 4. Data Acquisition with NI-cRIO9103

Inserts used in this work were Sandvik Coromill R390-11 T3 08M-PM 1030, hereafter referred to as "1030." The 1030 grade (TiAlN PVD coated) is recommended by Sandvik for milling R-108 and similar difficult-to-machine materials due to its resistance to material build-up on the cutting edge and plastic deformation [17]. Fresh (unworn) inserts were used at the beginning of each set of replication, with Olympus optical microscope to measure flank wear on the bottom edge of the insert. Measured tool flank wear of all 3 replications are shown in Table 1, where test "*m.n*" refers to.the m^{th} replication of the test and the n^{th} pass of that replication. The progress of tool flank wear is shown in Figure 5 for tests 1.2, 1.4, 1.6 and 1.8.

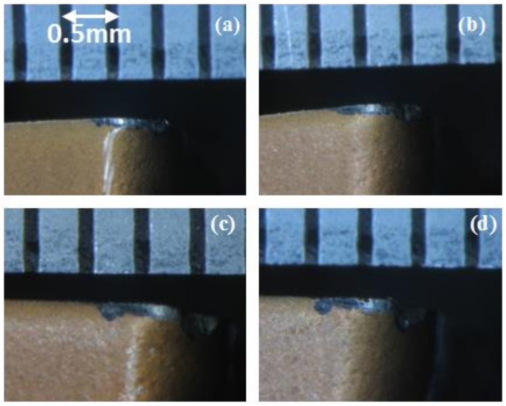

Figure 5. Measured flank wear for (a) Test 1.2, (b) Test 1.4 (c) Test 1.6, (d) Test 1.8

Table 1. Spindle power and flank wear measurement for (a) Training sets (b) Testing set

(a) Training Sets				(b) Testing Set	
Replication 1		Replication 2		Replication 3	
Test #	*VB* (μm)	Test #	*VB* (μm)	Test #	*VB* (μm)
1.1	84	2.1	83	3.1	81
1.2	89	2.2	87	3.2	87
1.3	100	2.3	103	3.3	99
1.4	108	2.4	107	3.4	103
1.5	111	2.5	109	3.5	109
1.6	116	2.6	116	3.6	115
1.7	119	2.7	125	3.7	116
1.8	125	2.8	127	3.8	120

IV. Results and Discussions

A. Data Processing with WPD

The captured spindle power and vibration signals for each pass is divided into 5 regions around 10, 20, 30, 40 and 50 mm distance of the tool travel on the workpiece, each containing 2^{11} sample points (*i.e.* 11 levels of WPD is feasible) as shown in Figure 6 and Figure 7. With an increase in tool wear, the magnitude and the spectral content of spindle power and vibration change; therefore, extracting proper features from wavelet packet coefficients makes tool wear estimation possible [13].

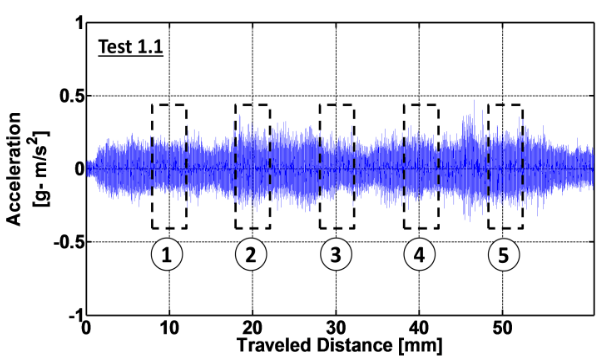

Figure 6. Measured vibration for pass 1.1

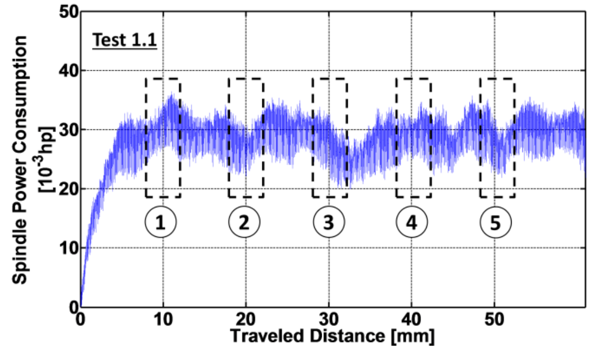

Figure 7. Measured power for pass 1.1

B. Feature Selection

Statistical features of the coefficients of wavelet packets at different levels are selected as the candidates that could be related to tool wear. Total of 8 different features, namely mean, standard deviation, root mean square (rms), skewness, Kurtosis factor, crest factor, peak to valley value, and energy of the coefficients, were selected as candidate features. Depending on the level of decomposition, different number of packets consisting of approximation and detail coefficients were generated. Therefore, the challenge is to determine (1) what type of wavelet and how many decomposition levels to use, and (2) which packet to select. Different researchers chose different wavelets and corresponding levels of decomposition; however, there is no clear explanation why those particular wavelets have been selected [18].

In this work, 3 different wavelets, namely Daubechies 4 (Db4), Daubechies 6 (Db6), Coiflet2 (Coif2), and Coiflet3 (Coif3) are selected, and corresponding features of level 3 and 4 decomposition on all the packets were calculated. Two factors were introduced to find the most correlated features to the output, correlation factor, and statistical overlap factor, which determines how much difference exists between a sharp and a dull insert [19]. After computing the two above factors for the features of each packet, the packet with the most correlated features was determined. For spindle power signal, features derived from approximation coefficients (A31 and A33) of 3-level WPD with Db3 mother wavelet showed the highest correlation to the tool wear. The same process was repeated for the vibration signal, and features derived from detail coefficients (D46 and D410) of 4-level WPD decomposition with Db4 as mother wavelet were selected. From these findings, it can be concluded that the features in lower frequency ranges of spindle power signal and features in higher frequency ranges of vibration signal are mostly correlated to the tool wear.

C. Feature reduction with PCA

Following the above procedure, a total of 32 different features were available. As stated in section III, redundancy exists between some of the features; on the other hand, some other features are not as much correlated as others to tool wear. Therefore, to avoid redundancy and decrease run-time of neural network, use of feature reduction is inevitable. After applying PCA, number of features decreased from 32 to 6. The cumulative sum of the first 6 highest eigenvalues in (4) shows that 85% of feature variations in the training set are captured by 6 of the transformed features.

cumulative sum of e-values=

$$\begin{bmatrix}0.29 & 0.49 & 0.63 & 0.72 & 0.80 & 0.85 & \dots & 1\end{bmatrix}^T_{32\times 1} \tag{4}$$

D. Training the NARX Neural Network

Before training, all of the six features and target values (measured tool wear) were normalized to avoid any inconsistency between dimensions of features. Mean square error (MSE) between outputs of neural network and targets are selected as the performance measure in training. Moreover, to avoid overfitting of the noise, cross validation in training set was conducted with 30% of the training inputs. Training iterations continue until the mean squared error of outputs in validation set increases or iterations exceed the predetermined threshold (1000 epoch). There are 4 parameters for the NARX model that need to be tuned in order to minimize the NN error: input delay (τ_i), feedback delay (τ_f), number of hidden layers (H), and number of hidden neurons (h) in each hidden layer. In this work, only one hidden layer (H=1) was assumed for training the network and other 3 parameters were tuned by running the NARX several times to reach the lowest MSE.

After training the NARX for several iterations, input delay τ_i =6, feedback delay τ_f =1 and hidden neurons h=1 were selected as the best NN parameters, which minimizes the testing set error. The estimation results for both testing set and training set are shown in Figure 8 and Figure 9 with corresponding 95% confidence intervals assuming enough samples for the training set. As expected, the confidence intervals when testing the network with untrained data are wider than when network is trained with the testing set.

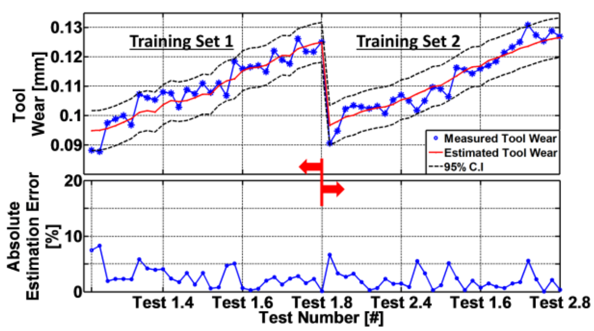

Figure 8. Neural Network results for training set 1st and 2nd

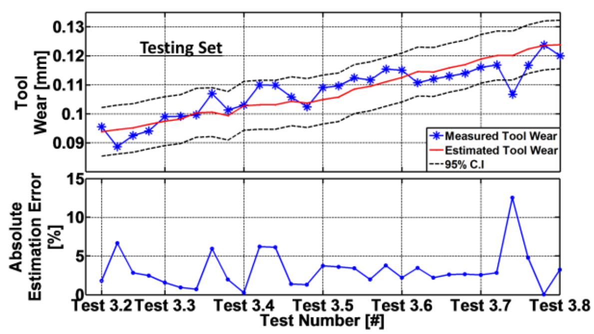

Figure 9. Neural Network results for testing set

Root Mean Squared Error (rsme), as the measure of performance, was also calculated for training and testing sets as 3 mm and 4.5 mm, respectively. The estimation result of the testing set (maximum of 13% relative error) shows the efficiency of using NARX model with low cost sensing methods for tool condition monitoring of difficult-to-machine materials. This method can be industrially implemented in CNCs with Open Architecture Controllers (OAC). A microprocessor with embedded signal acquisition and fast processing unit such Field Programmable Gate Array (FPGA) can programmed to flag an alarm when tool wear exceed the predefined threshold. Therefore, Improvement in productivity and quality and cost reduction per part can be guaranteed.

V. CONCLUSIONS

In this paper, recurrent neural network with statistical features of wavelet packets were trained for tool wear estimation of gamma-prime strengthened difficult to machine alloys. Sensory information from both spindle power and vibration were fused to increase the estimation performance of NN. The approach in this work was to implement sensor fusion idea with low cost sensing technologies that can be easily installed and maintained in machining operation plants. Since Hall effect sensors and accelerometer that were used in this work are relatively inexpensive, combined power and vibration method can be considered a cost effective measurement method for monitoring manufacturing processes. The main conclusions of this study are given as below:

- Wavelet packet decomposition was used for analyzing the measured signals in time-frequency domain. Eight statistical features of the wavelet coefficients were extracted and it was shown that decomposition of level 3 with Daubechies3 (Db3) wavelet and level 4 with Daubechies4 (Db4) wavelet produced the most correlated features to tool wear.
- Principal component analysis was utilized to reduce the dimensionality of extracted feature vectors from 32 features to 6 new features that had the largest variation in training set.
- Recurrent Neural network with NARX architecture selected and trained with Bayesian regularization backpropagation method to avoid overfitting of noise in training set. It was shown that NN architecture with 1 input delay, 6 feedback delays, and 1 hidden neuron is able to estimate the tool wear with maximum of 72 iterations.
- To validate the performance of the neural network with unseen data, new set of data was fed after training NARX, and it was shown that, the system is able to estimated tool wear with 4.5 RMS error and maximum of 13% point error.

Since Neural Network modeling is categorized as a black box method, the selection of proper features plays an important role for a meaningful training. Moreover, although tool wear estimation techniques are beneficial for modern manufacturing plants, they are passive methods that only give warning before catastrophic failure rather than being able to prevent it. Active methods, which can identify the state of the tool, can also prevent the tool reaching into catastrophic failure regions at the same time; so there is a need to investigate active tool condition monitoring. Investigating and identifying those methods is the subject of future studies.

VI. ACKNOWLEDGMENTS

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

REFERENCES

[1] D. Zhu, X. Zhang and H. Ding. Tool wear characteristics in machining of nickel-based superalloys. *Int. J. Mach. Tools Manuf. 64*pp. 60-77. 2013.

[2] H. Takeyama and R. Murata. Basic investigation of tool wear. *Journal of Manufacturing Science and Engineering 85(1),* pp. 33-37. 1963.

[3] E. Usui, T. Shirakashi and T. Kitagawa. Analytical prediction of three dimensional cutting process—Part 3: Cutting temperature and crater wear of carbide tool. *Journal of Manufacturing Science and Engineering 100(2),* pp. 236-243. 1978.

[4] Y. Yen, J. Söhner, B. Lilly and T. Altan. Estimation of tool wear in orthogonal cutting using the finite element analysis. *J. Mater. Process. Technol. 146(1),* pp. 82-91. 2004.

[5] A. Attanasio, E. Ceretti, S. Rizzuti, D. Umbrello and F. Micari. 3D finite element analysis of tool wear in machining. *CIRP Annals-Manufacturing Technology 57(1),* pp. 61-64. 2008.

[6] R. Teti, K. Jemielniak, G. O'Donnell and D. Dornfeld. Advanced monitoring of machining operations. *CIRP Annals-Manufacturing Technology 59(2),* pp. 717-739. 2010.

[7] D. E. Dimla. Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. *Int. J. Mach. Tools Manuf. 40(8),* pp. 1073-1098. 2000.

[8] L. Xiaoli. On-line detection of the breakage of small diameter drills using current signature wavelet transform. *Int. J. Mach. Tools Manuf. 39(1),* pp. 157-164. 1999.

[9] P. D. Alaniz-Lumbreras, R. A. Gómez-Loenzo, R. d. J. Romero-Troncoso, R. d. R. Peniche-Vera, J. C. Jáuregui-Correa and G. Herrera-Ruiz. Sensorless detection of tool breakage in milling. *Mach. Sci. Technol. 10(2),* pp. 263-274. 2006.

[10] P. D. Alaniz-Lumbreras, R. A. Gómez-Loenzo, R. d. J. Romero-Troncoso, R. d. R. Peniche-Vera, J. C. Jáuregui-Correa and G. Herrera-Ruiz. Sensorless detection of tool breakage in milling. *Mach. Sci. Technol. 10(2),* pp. 263-274. 2006.

[11] W. Wang, G. Hong, Y. Wong and K. Zhu. Sensor fusion for online tool condition monitoring in milling. *Int J Prod Res 45(21),* pp. 5095-5116. 2007.

[12] T. Segreto, A. Simeone and R. Teti. Sensor fusion for tool state classification in nickel superalloy high performance cutting. *Procedia CIRP 1*pp. 593-598. 2012.

[13] K. Zhu, Y. San Wong and G. S. Hong. Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results. *Int. J. Mach. Tools Manuf. 49(7),* pp. 537-553. 2009.

[14] D. L. Fugal. *Conceptual Wavelets in Digital Signal Processing: An in-Depth, Practical Approach for the Non-Mathematician* 2009Available: http://books.google.com/books?id=IUFAAQAAIAAJ.

[15] C. M. Bishop. *Pattern Recognition and Machine Learning* 2006 Available: http://books.google.com/books?id=kTNoQgAACAAJ.

[16] N. Baba and M. Kozaki. An intelligent forecasting system of stock price using neural networks. Presented at Neural Networks, 1992. IJCNN., International Joint Conference On. 1992, .

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

[18] R. Teti, K. Jemielniak, G. O'Donnell and D. Dornfeld. Advanced monitoring of machining operations. *CIRP Annals-Manufacturing Technology 59(2),* pp. 717-739. 2010.

[19] C. Scheffer and P. Heyns. An industrial tool wear monitoring system for interrupted turning. *Mechanical Systems and Signal Processing 18(5),* pp. 1219-1242. 2004.