INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

On-line tool wear estimation in turning through sensor data fusion and neural networks

Kamarthi, Sagar Vidya, Ph.D.

The Pennsylvania State University, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

The Pennsylvania State University

The Graduate School

Department of Industrial and Manufacturing Engineering

ON-LINE TOOL WEAR ESTIMATION IN TURNING THROUGH SENSOR DATA FUSION AND NEURAL NETWORKS

A Thesis in

Industrial Engineering

by

Sagar V. Kamarthi

© 1994 Sagar V. Kamarthi

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 1994

We approve the thesis of Sagar V. Kamarthi.

Date of Signature

Soundar R-T Kumara

oct 6, 1994

Soundar R. Tirupatikumara
Associate Professor Industrial Engineering
Thesis Co-Advisor
Chair of Committee

Paul H. Cohen

October 6, 1994

Paul H. Cohen
Professor of Industrial Engineering
Thesis Co-Advisor

Russell R. Barton

October 5, 1994

Russell R. Barton
Associate Professor Industrial Engineering

Asok Ray

October 4, 1994

Asok Ray
Professor of Mechanical Engineering

Allen L. Soyster

October 19, 1994

Allen L. Soyster
Professor of Industrial Engineering
Head of the Department of
     Industrial and Manuacturing Engineering

ABSTRACT

This work investigates a method which uses sensor data fusion and neural networks for on-line flank wear estimation. This method, unlike most of the existing ones, predicts the gradually increasing flank wear on a cutting tool continually.

The proposed method involves four important steps: sensor data acquisition, sensor data preprocessing, sensor data representation, and flank wear estimation from sensor data using a neural network. In this approach proper sensor data representation is crucial for obtaining accurate flank wear estimates. The purpose of sensor data representation in this application is to extract the information that is sensitive to gradually increasing flank wear.

During cutting, force, vibration, and acoustic emission signals are monitored. These sensor signals are digitized and preprocessed using a set of bandpass filters to improve the signal to noise ratio. Force and vibration signals are represented by univariate or multivariate autoregressive moving average models. Acoustic emission signals are represented by discrete wavelet transforms. The patterns created from the coefficients of either autoregressive moving average models or discrete wavelet transforms are input to a specially designed neural network architecture to compute flank wear estimates. This neural network is designed by combining Kohonen's feature maps, radial basis functions, and recurrent neural networks. This network satisfies two important design considerations: it reduces demand on supervised training time and data, and it remains fault tolerant to fluctuations in metal cutting.

The performance of the flank wear estimation method is studied by conducting a set of turning experiments on AISI 6150 steel. The results indicate that the proposed estimation method provides accurate flank wear estimates for the range of operating conditions that were used during neural network training. These results also indicate that the development of crater wear on a cutting tool adversely affects the accuracy of flank wear estimates. The root mean square of flank wear estimation errors is in the range of 0.0010 to 0.0017 inches. This makes the proposed flank wear estimation method very attractive for real-world applications.

This work stands unique in representing force and vibration signals by autoregressive moving average models, representing acoustic emission signals by discrete wavelet transforms, developing a novel neural network architecture, and testing the proposed flank wear estimation method with extensive experimentation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my sincere thanks to Dr. Soundar Kumara and Dr. Paul Cohen for initiating and guiding my research work in neural networks, sensor data fusion and machining processes. I am also thankful to them for their constant support in academic and administrative matters, and for their insightful guidance for my professional development. I am thankful to my committee members Dr. Russell Barton and Dr. Asok Ray for their enthusiasm, useful suggestions and critical comments on my research work.

I am grateful to all the faculty and staff members of IE department who have taught me and assisted me over the course of my studies at Penn State.

My heartfelt thanks to all of my student colleagues, particularly to the IDDRL members, for being so friendly, supportive, and helpful in several ways. Some of these benevolent friends are Mazin Al-Hamando, Nina Berry, Satish Bukkapatnam, Karthik Chittayil, Alex Daltrini, Niu Duan, Ching-Yao Kao, Taioun Kim, Jinwhan Lee, Kalyanaraman Madhavan, Yongjip Park, Ramani Pilla, Sayeed Quazi, Dongmok Sheen, Shabriya Theodore, Brian Sheldon, and Tim Thomas.

I thank my wife Geeta Kamarthi for her patience and understanding during the final and the critical phase of my Ph.D. work.

Finally, I thank all those who have honored me with their friendship, guided me with their wisdom, helped me with their kindness, and tolerated me out of their love.

I dedicate this work to my friend and guru, Bhagavan Sreeramulu Gajula, and my parents, Hampaiah Kamarthi and Narasamma Kamarthi.

Chapter 1

INTRODUCTION

1.1 Tool Wear Estimation: Preliminaries

On-line tool wear estimation is one of the main problems in automating turning processes. Several attempts have been made previously to develop on-line tool wear monitoring methods in the metal cutting domain.

Most of the previous approaches depend, directly or indirectly, on analytical models to address this problem. Most of the analytical models used in these approaches are inadequate in modeling the tool wear mechanisms. This in turn leads to inaccurate tool wear estimation. Inaddition, these analytical models use several empirical constants which are difficult and expensive to determine for each combination of tool and workpiece materials at various operating conditions. More recently approaches to tool wear estimation use neural networks. These neural network-based methods are effective in tool wear monitoring, but most of them are devised for either the prediction of tool breakage or for the detection of tool wear states such as fresh or worn.

Recent research findings show that no single process variable (such as force, temperature, acoustic emission, or vibration) by itself is sufficient to estimate tool wear states under all conditions. For example, acoustic emission is good for tool failure detection but not that good for continuous tool wear estimation. On the other hand, temperature is good for the tool wear estimation under certain conditions but not for tool failure detection. Similarly, estimation of tool wear

1

from force signals alone is difficult because of the opposing effects of flank and crater wear on cutting forces. While flank wear tends to increase cutting forces, the crater wear serves to decrease the forces due to its sharpening effect at the cutting edge.

The nature of the problem suggests that to develop a method for practical and accurate tool wear estimation, the following two criteria should be considered:

1. The tool wear estimation method should follow a data-dependent approach.

2. The method should use multiple process variables to contain sufficient tool wear related information at varying operating conditions.

This necessitates the development of techniques for using multiple sensor data (data fusion) and nonlinear estimation methods (neural networks). Therefore, in the present research a method based on sensor data fusion and neural networks for continuous tool wear estimation is developed. The present method is based on the understanding that a metal cutting process produces dynamic signals which contain information about the changing process states such as tool wear, and that this information can be extracted and related to tool wear. The proposed method relies on the possibility that these dynamic signals from the metal cutting process can be measured and processed in real-time to obtain on-line tool wear estimates. This assumption is realistic due to the existence of sophisticated sensors, high speed data acquisition systems, and fast computers.

1.2 Overview of Tool Wear in Turning

Generally, in a turning process the term *tool wear* refers to the following three changes:

1. Degradation of the cutting tool due to gradual wear on its flank face and rake face.

2. Fracture of the cutting edge.

3. Deterioration of the tool tip mechanical properties due to high cutting temperature and pressure.

Figure 1.1 shows the regions of tool wear in a turning process. It occurs in various forms, predominantly due to three wear mechanisms: adhesion, abrasion, and diffusion.
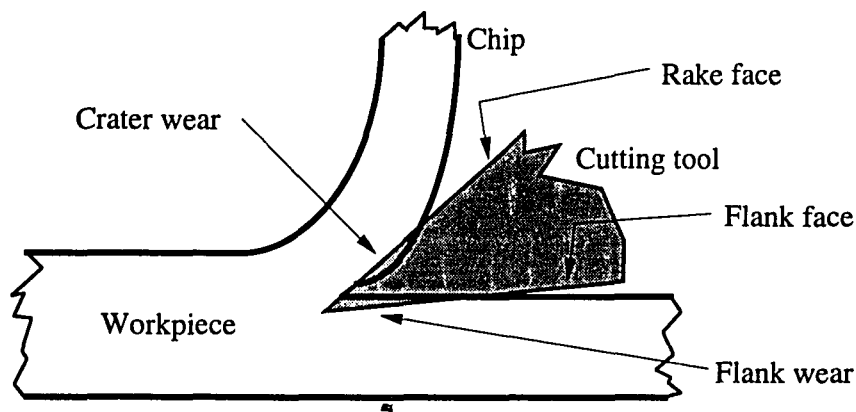
Figure 1.1: Regions of tool wear on a cutting tool in turning process

Adhesion wear on a cutting tool is caused by the fracture of welded asperity junctions between the tool and workpiece materials. In a cutting process, as new material surfaces uncontaminated with oxide films are created, welded asperity junctions are formed. When these junctions are fractured, small fragments of tool material are torn out and carried away with the chip sliding on the workpiece surface.

3

Abrasion wear on a cutting tool surface occurs when hard particles on the underside of the chip pass over the tool face and remove the tool material by mechanical action. These hard particles may be highly strain-hardened fragments of an unstable built-up edge, fragments of the hard tool material removed by adhesion wear, or hard constituents in the workpiece material.

Diffusion wear occurs when atoms move from the tool material to the workpiece material as the tool and workpiece materials come in close contact with each other at high pressure and temperature. This movement of atoms from the tool material to the workpiece material weakens the surface structure of the tool. The diffusion wear may also occur when some of the constituent atoms in the tool material diffuse from surface layers to deeper layers.

While cutting is in progress, various types of wear may develop on a cutting tool. These can be classified into regular and irregular types of wear. Figure 1.2 depicts the regular types of wear and the dimensions that characterize them. The regular types of wear are major flank wear, minor flank wear, nose wear, notch wear, and crater wear. These are usually present in a machining operation and have regular growth characteristics with respect to cutting time. The flank, nose, and notch wear are formed by the rubbing action of the newly generated workpiece surface with the major and minor flank faces of the tool. The crater wear is created by the chip sliding on the rake face at high pressure and temperature.

The irregular types of wear are plastic deformation, chipping, cracking, and breakage. Though these are unpredictable in occurrence, they can be avoided by proper selection of tool materials and cutting conditions.

The general undesirable effects of tool wear are loss in dimensional accuracy of the finished product, reduction in the workpiece surface-finish quality, and damage

4
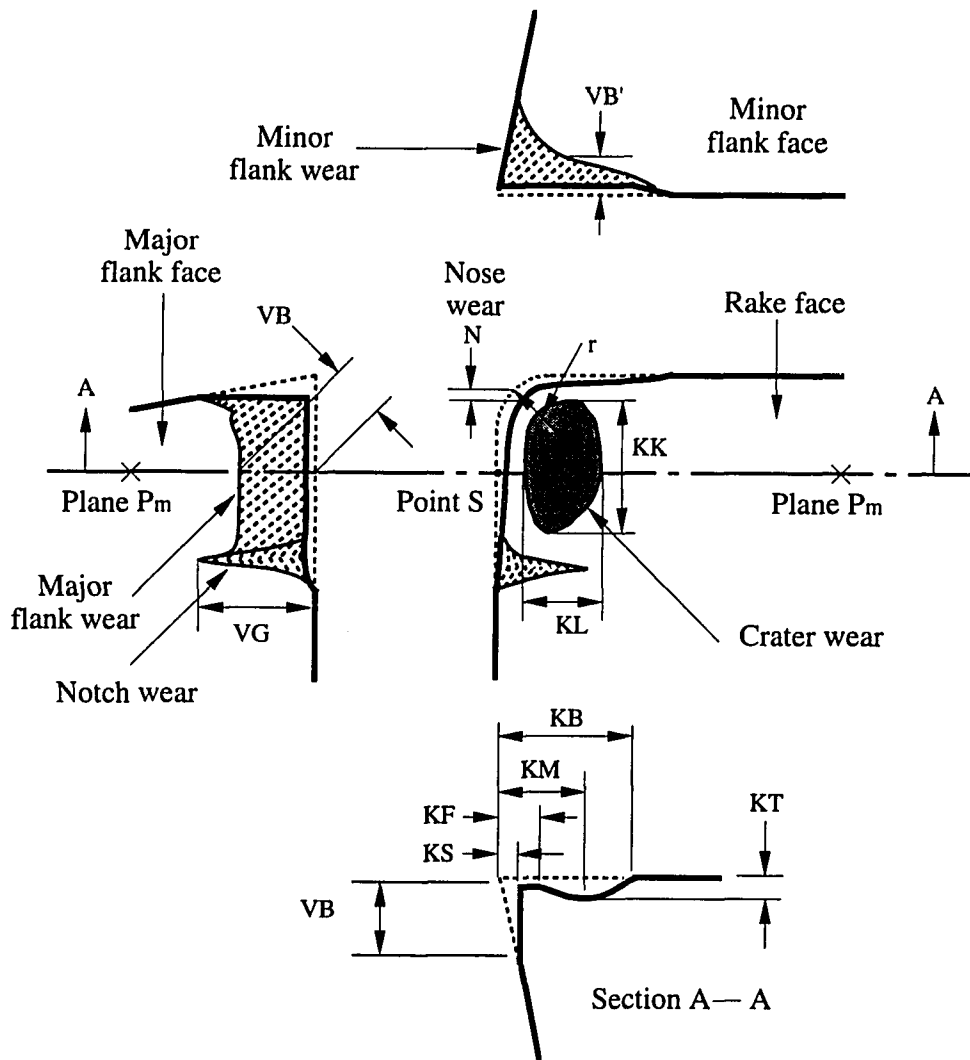
S: point on the original major cutting edge at middle of flank/crater wear
Pm: plane perpendicular to the major cutting edge through point S
VB: major flank wear
VB': minor flank wear
KT: crater depth on the rake face
KK: crater length on the rake face
KL: crater width on the rake face
KS: retract of the cutting edge
N: nose wear

Figure 1.2: Regular types of tool wear and the dimensions that characterize them

5

to the workpiece. Any of these effects may result in scrapping the product.

Of the regular types of tool wear, only major flank wear and nose wear directly affect workpiece dimension and surface roughness. Major flank wear and nose wear together are referred to by simply flank wear in this thesis. The effect of flank wear on the quality of a workpiece is illustrated in Figure 1.3. As the flank wear on the flank face progresses, the recession of the cutting edge increases. This results in gradual increase of the workpiece diameter. As both the major and the minor flank wear progress simultaneously, the nose radius (r in Figure 1.2) decreases and consequently the quality of the workpiece surface finish deteriorates. Among all types of wear, flank wear has the most deleterious influence on the quality of the workpiece (Boothroyd and Knight 1989).

Crater wear, unlike flank wear, does not affect the quality of the workpiece when it gradually increases with the cutting time. As the crater wear progresses towards the cutting edge, it will weaken the tool edge. Eventually when it touches either major or minor cutting edge of the tool, the tip of the tool will break catastrophically. This may badly damage the workpiece.

1.3 On-Line Flank Wear Estimation in Turning

In this thesis, on-line flank wear estimation refers to the prediction of the height of the flank wear land (VB) on a cutting tool in a turning process. The height of the flank wear land (VB), as shown in Figure 1.4, is considered an important parameter that characterizes the flank wear land on a cutting tool.

6
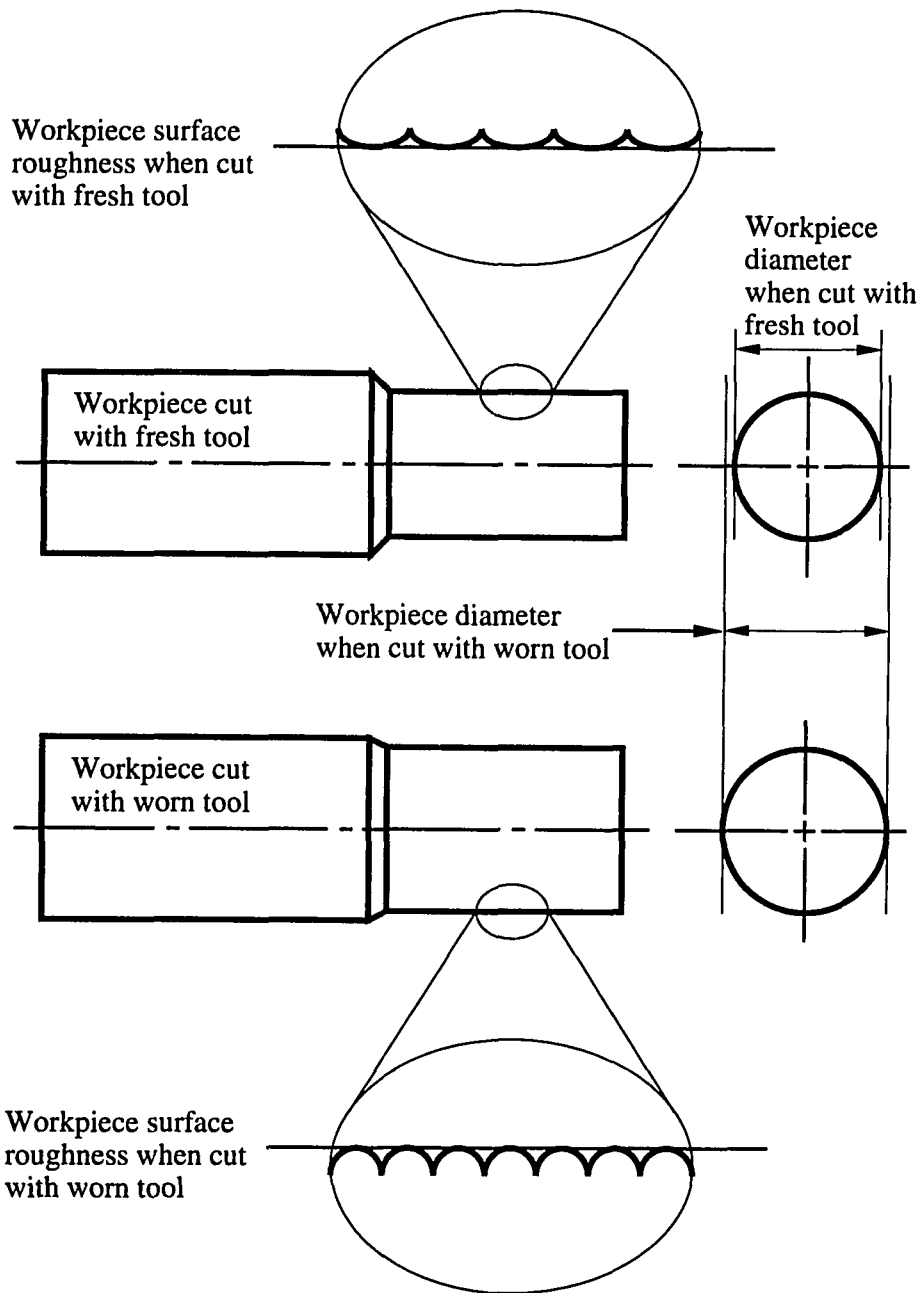
Figure 1.3: Effect of flank wear on workpiece diameter and surface roughness

7

1.3.1 Problem Statement

Statement: *Develop a method for on-line continuous flank wear estimation in a turning process.*

Hypothesis: *The metal cutting process, in turning, generates some dynamic signals which contain information about the gradually progressing flank wear on the cutting tool. This flank wear-related information can be extracted from these dynamic signals and used for a continuous on-line flank wear estimation.*
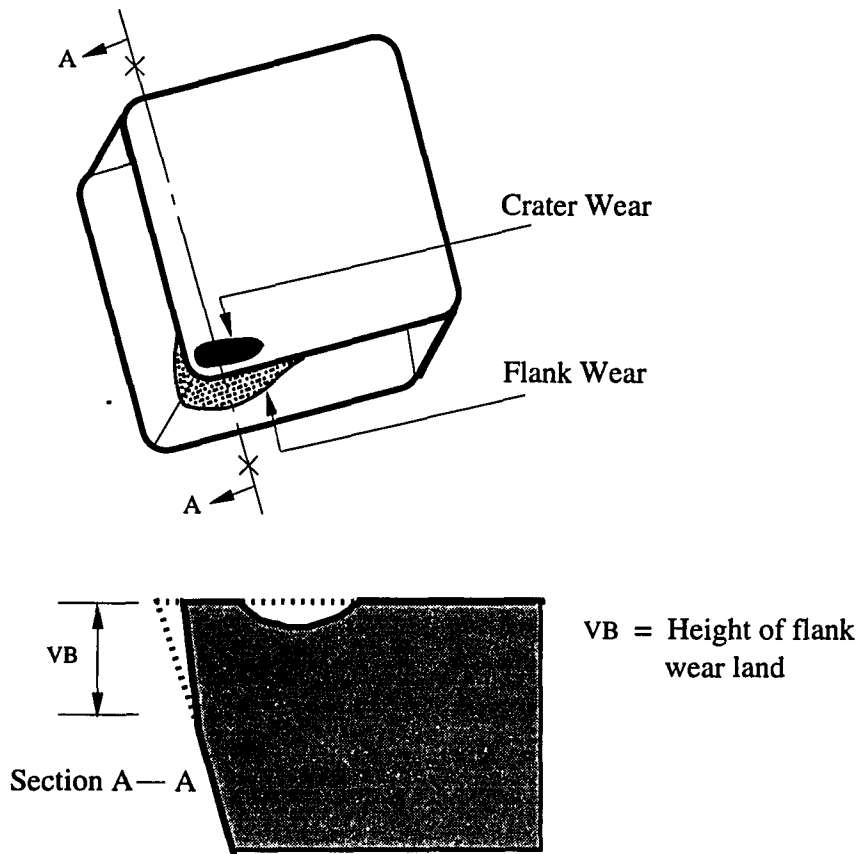
Figure 1.4: Height of flank wear land

8

1.3.2 Research Objectives

The following are the objectives of the current research:

1. Determine the process variables in turning processes that contain useful information related to flank wear.

2. Determine the appropriate sensors that can be used for monitoring the process variables that are related to flank wear.

3. Design and develop a technique to enrich sensor signals so that the information related to flank wear is retained and the information irrelevant to flank wear is eliminated.

4. D velop the most appropriate sensor data representation schemes.

5. Design and develop the effective mathematical or statistical methods that can be used to relate sensor measurements to flank wear.

6. Design, develop and implement the experimental procedures to verify and validate the on-line flank wear estimation method.

7. Study the performance of different sensor combinations for flank wear estimation.

1.4 Motivation

While the continued development of machines and cutting tools capable of high cutting speeds contributed to the higher production rates, it also resulted in the increased downtime due to frequent failure of cutting tools either by wear or by breakage (Balakrishnan 1989). It was reported that tool failure contributes, on an average, 6.8% to the downtime of a machining center (Kegg 1984). In order to avoid catastrophic tool failures, the current tool change strategies are based

9

on conservative tool life estimates. This results in the underutilization of tools (Mackinnon et al.1986, Dan and Mathew 1990). To automate the tool change operations and maximize the cutting tool utilization, it is necessary to develop methods for on-line tool wear estimation and tool failure detection.

Continuous tool wear estimation is also important in adaptive control of turning processes in CNC machine tools. Systems for Adaptive Control with Optimization (ACO) either maximize or minimize a given performance index subjected to the process and system constraints. These performance indices are usually the economic functions of material removal and tool wear rates (Centner 1964). Without the measurements of tool wear and wear rate, the economic optimization of a machining process cannot be achieved. Particularly in milling and turning, the ACO systems require on-line tool wear measurements. Therefore, the progress of ACO systems depends on the availability of accurate on-line tool wear estimation methods (DeFillippi et al. 1969, Takeyama et al. 1970, Mathias et al. 1980, Koren 1989).

A Geometric Adaptive Control (GAC) system maintains a workpiece quality by achieving the required dimensional accuracy and consistency in the surface finish of machined parts. Both the dimensional accuracy and the surface finish are affected by the flank wear which develops during turning operation. If accurate on-line flank wear estimates are available, dimensional accuracy can be achieved by adjusting the tool-workpiece distance, and the desired surface quality can be maintained by varying the machining feed rate (Peklenik 1970, Koren 1989).

In addition to maintaining the required dimensional accuracy and surface finish, it is desirable to avoid tool fracture during a turning operation. A method capable of anticipating tool fracture will avoid damage to the workpiece, partic-

10

ularly during finishing cuts. Even a method that can detect a broken tool after breakage occurs offers benefit of reducing wasted time.

In order to automate the tool change operations and implement the adaptive control systems, it is important to develop methods that can provide reliable and accurate on-line tool wear estimates under practical operating conditions. This forms the fundamental motivation for this research work.

1.5 Thesis Outline

Chapter 2 reviews the important tool wear estimation methods based on analytical models, observers, and empirical models. This chapter discusses the relative merits of these approaches.

Chapter 3 reviews the tool wear estimation methods based on neural networks. This chapter classifies the neural network-based methods and identifies their essential strengths and differences.

Chapter 4 studies the possible sensors that are useful for measuring flank-wear-sensitive process parameters such as cutting forces, vibration acceleration, acoustic emission, temperature, etc. This chapter sets the criteria for selecting the most suitable sensors for on-line flank wear estimation in turning processes

Chapter 5 describes the methodology to address the on-line flank wear estimation problem. This chapter also describes the experimentation strategy to verify and validate the performance of the proposed methodology.

11

Chapter 6 studies the characteristics of cutting forces and vibration signals in turning process. This chapter develops a representation scheme for force and vibration signals using univariate or multivariate autoregressive moving average models.

Chapter 7 studies the characteristics of acoustic emission signals in turning processes. This chapter develops a representation scheme for acoustic emission signals using discrete wavelet transforms.

Chapter 8 describes a specially designed neural network architecture for flank wear estimation. This chapter deals with the design, implementation, and application of the proposed architecture to the continuous flank wear estimation problem.

Chapter 9 presents the experimental results and analysis. Flank wear estimation results given by different combinations of force, vibration, and acoustic emission sensors are compared.

Chapter 10 presents the conclusions and uniqueness of the work. It also discusses the directions for future research.

Chapter 2

A REVIEW OF IMPORTANT TOOL WEAR ESTIMATION
METHODS IN TURNING

2.1 Introduction

This chapter reviews some of the most important tool wear estimation methods in turning. This includes the tool wear estimation methods based on analytical models, observers, and empirical models. The scope of the chapter is limited to the methods that provide continuous tool wear estimation. This chapter summarizes the selected methods and discusses their merits and drawbacks. The concluding section of the chapter discusses the comparative merits of analytical models, observers, and empirical models.

On-line tool wear estimation in the turning process is primarily required for predicting tool failures, devising optimal tool change strategies, and implementing Adaptive Control with Optimization (ACO) and Geometric Adaptive Control (GAC) schemes. A reliable tool failure prediction method will maximize the utilization of cutting tools. An ACO scheme will achieve optimal metal removal rates by adjusting the process inputs according to the changes in the process states such as flank wear and crater wear. These process states are usually estimated on-line by monitoring the process variables such as force, vibration, temperature, and acoustic emission (Centner and Idelson 1964, Hinds 1977). A GAC scheme will maintain the desired dimensional tolerances and surface finish requirements on a workpiece by adjusting the process variables or the cutting tool position with

13

respect to the workpiece. These adjustments are necessary to compensate for tool wear growth, tool deflection, workpiece vibration, etc., (Peklenik 1970, Koren 1989).

The optimal tool change strategies and adaptive control schemes can be implemented if reliable and consistent on-line tool wear estimation methods are available. Both direct and indirect methods for on-line tool wear monitoring have been investigated for over two decades. Direct methods make an assessment of tool wear by evaluating the worn surface using optical methods, measuring the material loss of the tool by radiometric techniques, measuring electrical resistance at the tool-workpiece contact, or measuring the distance between tool and workpiece during machining. Indirect methods, on the other hand, estimate tool wear from a model relating tool wear and the measurable process variables such as force, acoustic emission, vibration, temperature, power or motor current, sound, or workpiece size and surface roughness. Though both direct and indirect methods are effective in tool wear monitoring to some degree, direct methods are difficult to implement due to:

1. Limited accessibility to the cutting region during machining,

2. Long processing times in optical methods, and

3. Potential hazards and special tool preparations in radiometric techniques.

Indirect methods are more suitable for on-line tool wear monitoring. The advances in sensor technology have made indirect methods more attractive. For relating tool wear to the measurable process variables, indirect methods follow three main approaches: analytical models, observers, and neural networks. Traditionally, analytical models are used for tool wear estimation; the current trend is to use either observers, empirical models, or neural networks. There are some

14

other methods that strictly cannot be classified into any of these approaches. A sensor-based method developed by El Gomayel and Bregger (1986) and a wear index-based method develop by Rao (1986) are examples of such methods.

This chapter reviews the important contributions of the investigations into the application of analytical models, observers, empirical models, and the other miscellaneous methods for tool wear estimation. Section 2.2 discusses the analytical models. Section 2.3 discusses the observer based methods. Section 2.4 discusses the empirical model-based methods and section 2.5 discusses the miscellaneous methods. In all these categories, the scope of the chapter is limited to the methods that provide continuous tool wear estimates. This chapter summarizes the selected methods and discusses their merits and drawbacks. Wherever available, the experimental details of the select methods have been presented. A critical cross comparison of different methods is present in section 2.6.

2.2 Analytical Models For Tool Wear Estimation

In metal cutting, analytical methods are developed based on the understanding of the physics of the cutting process. Analytical models of tool wear are based on the mathematical descriptions of the physical mechanisms such as diffusion, adhesion, and abrasion that are responsible for tool wear in a cutting process. These models are developed by closely studying the physics of a cutting process. The following sections briefly describe three important analytical models of tool wear developed by Koren and Lenz (1972), Koren (1978), and Usui et al. (1984).

2.2.1 Analytical Model for Flank Wear (Koren and Lenz 1972)

In an effort to optimize the turning operation, Koren and Lenz (1972) first

15

developed a mathematical model for a turning process. This model yields a relationship between wear components and process variables. The wear components considered in the model are the flank wear due to abrasion and the flank wear due to diffusion. Process variables in the model are cutting speed, feed, depth of cut, and rake angle. This model uses cutting force and temperature as internal variables. The model is given by

$$\frac{l_0}{v} \dot{w}_{f1} + w_{f1} = K_F \cos \gamma \left(\frac{F}{fd} \right) ,$$
(2.1)

$$\dot{w}_{f2} = K_D \sqrt{v} \exp \left(\frac{-A}{273 + \theta} \right) ,$$
(2.2)

where

A	constant coefficient,
d	depth of cut,
F	cutting force,
K_F	constant coefficient,
K_D	constant coefficient,
l_0	length constant,
v	cutting speed,
w_{f1}	width of wear on the tool flank due to abrasion,
w_{f2}	width of wear on the tool flank due to diffusion,
θ	mean temperature at the tool-work interface, and
γ	rake angle.

16

This model is developed specifically for carbide tools. The wear-time curves generated from these equations are reported to agree with those observed in general practice.

Comments: This is the first model of its type. The tool wear estimation results using this model are not sufficiently explored. Like most of the analytical models, this model needs extensive experimentation to evaluate the various model coefficients for different tool and workpiece materials at different process conditions.

2.2.2 Analytical Model for Flank Wear (Koren 1978)

Koren (1978) developed a physical model of flank wear based on a feedback mechanism. It is assumed that the flank wear has two principal components: a thermally activated one and a mechanically activated one. The overall flank wear is considered equal to the sum of these two components. This model describes the growth of flank wear with time:

$$w_{f1}(t) = A_1 \left[\exp\left(\frac{v^n t}{C_1}\right) - 1 \right] + B_1 \left[1 - \exp\left(\frac{-vt}{C_2}\right) \right], \qquad (2.3)$$

$$w_{f2}(t) = A_2 \left[\exp\left(\frac{v^n t}{C_1}\right) - 1 \right] + B_2 \left[1 - \exp\left(\frac{-vt}{C_2}\right) \right], \qquad (2.4)$$

where

A_i constant coefficients,

B_i constant coefficients,

C_i constant coefficients,

n constant exponent, and

v cutting speed.

17

The parameters A_1, A_2, B_1, B_2, C_1, C_2, and n are dependent on feed, depth of cut, and tool and workpiece materials. These parameters are to be determined from experimental wear curves for selected tool and workpiece materials and force measurements. It is reported that wear curves generated from this model agree with the practical wear curves.

Comments: Sufficient evidence regarding the accuracy of tool wear estimation using this model is not available. Because of the difficulty in obtaining the model coefficients and exponents, the usefulness of this model in a practical situation is limited.

2.2.3 Analytical Model for Flank and Crater Wear (Usui et al. 1984)

Usui et al. (1984) developed an analytical method for the prediction of crater and flank wear of tungsten carbide tools. This method is designed for wear prediction for a wide variety of tool shapes and cutting conditions. This method has four main steps:

1. Identification of a wear characteristic equation which involves wear rate, normal stress, temperature, and two material constants.

2. Prediction of chip formation and three cutting force components from orthogonal cutting data using the energy method. The tool edge shape and cutting conditions are taken as inputs to this method.

3. Computation of the normal stress and the temperature on the wear faces from the force components obtained from the energy method.

4. Computation of crater and flank wear from the characteristic equation applying normal stress and temperature computed from the previous step.

18

Wear Characteristic Equation: The proposed wear characteristic equation was developed based on the assumption that wear rate is determined predominantly by an adhesion type wear mechanism in the tungsten carbide tools. The proposed wear characteristic equation originates from the following equation for adhesive wear:

$$dW' = A_r \frac{c}{b} Z dL, \qquad (2.5)$$

where dW' is the wear volume for the sliding distance dL, A_r is the real area of contact, c is the height of the postulated plate-like wear particles, b is the mean spacing of the asperities, and Z is the probability of producing a wear particle per asperity encountered. After introducing some substitutions and approximations, the final wear characteristic equation takes the following form:

$$\frac{dW'}{\sigma_t V_c dt} = C_1 \exp\left(-\frac{C_2}{\theta}\right), \qquad (2.6)$$

where σ_t is the normal stress, V_c is the chip velocity, θ is the temperature of the chip surface, and C_1 and C_2 are the material constants. In this equation, the constants C_1 and C_2 are assumed to account for the adhesion wear mechanism as well as for the effects of any abrasive particles involved at the asperities in the material. For simplicity in computing crater and flank wear, it was assumed that the variation in the parameters σ_t, V_c, and θ with the progress of wear is negligible.

Results: This wear characteristic equation was verified by using both experimental and predicted values of V_c, θ and σ. It was found that the wear characteristic equation holds well for steels of different carbon content at higher temperature ranges. The wear characteristic equation is insensitive to the variations in cutting

19

force, temperatures, tool face stresses, and crater wear caused by the variation of carbon content in steels. However, the wear characteristic equation was reported to give incorrect results for steels with high carbon content at low temperature ranges.

The wear characteristic equation was found to estimate the correct wear rates for both crater and flank wear. It was also found that the effect of carbon in carbide wear is reflected in the characteristic constants C_1 and C_2. The difference in wear rates for steels of different carbon contents is attributed to the difference in their abrasive nature.

The predicted contour lines of crater wear from the wear characteristic equation showed an overall agreement with those obtained in actual turning experiments. However, the deepest portion of the real crater was located closer to the front cutting edge, than the crater predicted by the equation. This discrepancy was attributed to the assumption that σ_t is uniform over the area of contact even though it is nonuniform in reality. The predicted flank wear progress with time showed close agreement with that obtained from the experimentation.

Comments: In this work, the physics underlying the tool wear mechanisms is studied in detail. The proposed wear characteristic equation is fairly simple and provides good tool wear estimates for both crater and flank wear. The wear characteristic equation gives good results for steels of varying carbon content, different tool geometries, and a wide range of operating conditions.

The accurate evaluation of characteristics constants C_1 and C_2 and the equation parameters σ_t, V_c, and θ poses a major difficulty to using this wear characteristic equation in practice. Nevertheless, this study enhances the understanding of the cutting process and wear mechanisms in turning processes.

20

2.3 Observers for Tool Wear Estimation

An observer is an auxiliary dynamic system driven by available system inputs and outputs to reconstruct the missing state variables of a dynamic system. If some observer parameters are unknown, a parameter estimator is usually combined with the state observer to estimate the model parameters and the observer states simultaneously. Such an observer is called an adaptive observer. If a dynamic system is described by a set of nonlinear equations, state variables of the system can be obtained in either of the following two ways. The nonlinear equations can first be linearized and then a state observer can be designed for the linearized set of equations. Or a nonlinear observer can be designed for reconstructing the missing state variables. Danai and Ulsoy (1988a, 1988b) investigated an adaptive observer for estimating the flank and crater wear components. Park and Ulsoy (1989a, 1989b, 1992) investigated a nonlinear observer for estimating flank wear components. Further, Park and Ulsoy (1990, 1993a, 1993b) investigated the adaptive observer integrated with a computer vision system for flank wear estimation. The following sections review the literature on these important applications of observers for on-line tool wear estimation.

Ralston et al. (1988) proposed a linear observer for tool wear estimation and Arulalan et al. (1988) extended it to an adaptive observer for optimal control of tool wear. However, no experimental work or practical results have been reported on the performance of these observers. These observers are not reviewed in this chapter because the details of their practical application are not available.

2.3.1 Adaptive Observer (Danai and Ulsoy 1988a, 1988b)

Danai and Ulsoy (1988a, 1988b) developed an adaptive observer for on-line

21

tool wear estimation in the turning process. The observer, which estimates the process states, is designed based on a dynamic state model that captures mechanistic and empirical relationships of process variables in the turning process. The parameter estimator, which estimates the observer parameters, is designed according to an output error method. Once the observer states are estimated, they are transformed into physically meaningful tool wear states for the original state space model from which the state observer had been derived.

State Observer: In this work, the turning process is represented by the following linear time-invariant single-input single-output 3rd order state space model:

$$\mathbf{x}(k+1) = \mathbf{P}\mathbf{x}(k) + \mathbf{q}u(k), \tag{2.7}$$

$$y(k) = \mathbf{c}^T\mathbf{x}(k) + Du(k). \tag{2.8}$$

The state vector $\mathbf{x}$ contains three components: flank wear due to abrasion w_{f1}, flank wear due to diffusion w_{f1}, and crater wear w_c. The scalar input u is the feed rate, and the scalar output y is the force. The model structure, states, input, and output are identified by studying the properties of the linearized state space model of the following nonlinear dynamic state model of tool wear in the turning process (Danai and Ulsoy 1987).

State Equation

$$\frac{l_0}{v} \dot{w}_{f1} + w_{f1} = K_1 \cos \alpha_r \left(\frac{F}{fd} \right), \tag{2.9}$$

$$\dot{w}_{f2} = K_2 \sqrt{v} \exp \left(\frac{-K_3}{273 + \theta_f} \right), \tag{2.10}$$

22

$$\dot{w}_c = K_4\, F\, v\, \exp\left(\frac{-K_5}{273 + \theta_c}\right). \qquad (2.11)$$

Output Equation

$$F = \left[K_6\, f^{n1}\,(1 - K_7 \alpha_r) - K_8 - K_9\, v\right] d + K_{10}\, d\, w_f - K_{11}\, w_c, \qquad (2.12)$$

where

$$w_f = w_{f1} + w_{f2}, \qquad (2.13)$$

$$\theta_f = K_{12}\, v^{n2}\, f^{n3} + K_{13}\, w_f^{n4}, \qquad (2.14)$$

$$\theta_c = K_{14}\, F\, v^{n5}\, f^{n6}\, d^{n7}, \qquad (2.15)$$

and

d depth of cut,

F cutting force,

K_i constant coefficients,

l_0 constant coefficient,

n_i constant exponents,

v cutting speed,

w_f total width of wear on the tool flank,

w_{f1} width of wear on the tool flank due to abrasion,

w_{f2} width of wear on the tool flank due to diffusion,

w_c maximum depth of crater wear,

θ_f mean temperature at the tool-work interface,

23

θ_c mean temperature at the tool-chip interface, and

α_r rake angle.

This nonlinear model is the combination of tool wear models developed by Koren and Lenz (1972) and Usui et al. (1978). Before using this nonlinear model, its suitability to on-line tool wear estimation was verified. Its qualitative and quantitative behavior were verified by simulation studies. Its stability, controllability, and observability properties were studied by linearizing the nonlinear model. The model was found to be controllable with feed input and observable with force output. However, the model was found to be unstable for large values of wear.

The state space model expressed in Eq. (2.7) and (2.8) was transformed into the observer form through a transform matrix $\mathbf{T}$ such that $\mathbf{x} = \mathbf{T}\mathbf{x}_0$. Thus, the observable canonical form for a third order linear model is

$$\mathbf{x}_0(k+1) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} \mathbf{x}_0(k) + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} u(k), \qquad (2.16)$$

$$y(k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{x}_0(k) + Du(k). \qquad (2.17)$$

An estimate of the state $\hat{\mathbf{x}}_0(k)$ is obtained from a state observer of the form

$$\hat{\mathbf{x}}_0(k+1) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} \hat{\mathbf{x}}_0(k) + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} u(k) + \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} [y(k) - \hat{y}(k)], (2.18)$$

$$\hat{y}(k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \hat{\mathbf{x}}_0(k) + Du(k). \qquad (2.19)$$

24

In the above observer, a_i and b_i are the model parameters, and g_i are the observer gains and are set to $g_1 = -\hat{a}_1$, $g_2 = -\hat{a}_2$, and $g_3 = -\hat{a}_3$ to achieve a faster convergence of the observer error.

Adaptive Observer: The state observer described in Eq. (2.18) and (2.19) represents the machining process by a deterministic linear time-invariant, single-input, single-output, observable model. Since such a linear time-invariant model is rarely adequate to describe the changes in a complex process such as machining, the observer's constant parameters will not hold well over the entire process conditions. To circumvent this problem, parameters of the linear state observer are adaptively updated using a parameter estimator. A least square recursive parameter estimation algorithm based on an *output error* method or an *equation error* method is used to estimate the observer parameters. It is assumed that the adaptive observer parameters vary slowly compared with the observer states, and also that the initial observation error due to poor initial parameter convergence is tolerable.

Reconstruction of Tool Wear States: Once the states x_0 in the observer form are estimated, they are transformed into the tool wear states x of the original system state space model form which the observer from had been derived. To reconstruct the tool wear components from the observed states, a transformation matrix or a transformation function is determined from off-line measurements. In the case of a first order linear model, the scalar transformation parameter is easily determined from off-line measurements. In the case of a higher order model, a transformation function is defined as $\hat{x} = f(\hat{x}_0)$, where the function $f(\hat{x}_0)$ is determined by regression analysis on off-line tool wear measurements and observer states.

25

Experimentation: To check the performance of the adaptive observer in a single-wear case and also to investigate the behavior of the estimated model parameters, two sets of experiments were conducted on a CNC lathe: one causing flank wear and the other causing crater wear as the single dominant wear component. During these experiments, cutting force data, flank wear, and crater wear were measured at several points.

Results: The estimated flank wear from the first experiment showed good agreement with the actual values of tool wear measured at several points during the cut. However, the estimated crater wear from the second test did not match well with the measured wear. Also, this method did not perform well in the multi-wear situation.

The estimated parameters in these experiments showed a drastic change due to tool breakage. These results suggest that tool failure can be detected from the sudden change in the tool wear parameters.

Comments: The adaptive observer proposed in this method produced close estimates for the model parameters and observer states. However, the difficulty in reconstructing the real wear values from the observer states limits the method's usefulness for on-line tool wear estimation. The inaccuracies in tool wear estimation results are attributed to the inaccuracies in the wear model, linearization of the model, and the high degree of accuracy required for the transformation of wear components of small magnitude.

2.3.2 Adaptive Observer (Park and Ulsoy 1990, 1993a, 1993b)

Park and Ulsoy (1990, 1993a, 1993b) proposed an adaptive observer for on-

line tool wear estimation based on force measurements and intermittent flank wear measurements. The intermittent flank wear measurements are obtained using a computer vision system. The proposed adaptive observer is based on the simplified version of the flank wear model developed by Koren and Lenz (1972). This model was described earlier in this chapter (See Eq. (2.1) and (2.2)). The original model of Koren and Lenz (1972) assumes that flank wear is caused by abrasion and diffusion mechanisms. Park and Ulsoy (1990, 1993a, 1993b) simplified the original model by assuming that flank wear is caused predominately by diffusion. The discrete time version of the observer is described by the following equations:

$$\hat{w}_f(t+1) = \bar{w}_f(t) + TK_2\sqrt{v}\,\exp\left[\frac{-K_3}{273 + \theta_f(t)}\right] + K\left[F(t) - \hat{F}(t)\right], \quad (2.20)$$

$$\hat{F} = F_0 + K_{10}\,d\,(w_f(t) - w_{f0}), \quad (2.21)$$

where

$$\hat{\theta}_f(t) = K_{12}\,v^{n2}\,f^{n3} + K_{13}\,w_f^{n4}\,. \quad (2.22)$$

In the above model, F_0 and w_{f0} are the initial cutting force and flank wear at the beginning of a cutting operation respectively. T is the sampling interval, and K is the discrete time observer gain between 0 and $2/K_{10}d$. It is assumed that initial flank wear w_{f0} and initial cutting force F_0 are available through direct measurements. Also, it is assumed that parameter K_3, the activation energy for diffusion for the given tool-work combination, is found from the diffusion data handbooks. The gradient term $K_{10}d$ is approximated from the direct flank wear measurements. The only unknown parameter, $\theta = K_2\sqrt{v}$, is recursively estimated

27

by the *Recursive least square* parameter estimator. The estimated parameter θ is used by the state observer to estimate the flank wear $\hat{w}_f$. Since the parameters of the state observer are adjusted intermittently by a parameter estimator, the state observer acts as an adaptive observer.

Integration of the Adaptive Observer and Computer Vision: The integration of the adaptive observer with the computer vision system eliminates the need for preliminary experiments to determine the parameter K_{10}. Whenever the tool is not in contact with the workpiece, the computer vision system directly measures the actual amount of flank wear. Using the current and the previous flank wear measurements and the cutting force data, the observer gradient $K_{10}d$ is calculated and updated. The updated K_{10} is used by the adaptive observer during the subsequent machining operation until it is updated again. $w_{f(i)}$ and $w_{f(i-1)}$ are the directed flank wear measurements at t_i and t_{i-1} respectively. The gradient $K_{10}d$ is calculated using

$$K_{10}d \approx \bar{C} \left[\frac{t_i - t_{i-1}}{w_{f(i)} - w_{f(i-1)}} \right] , \qquad (2.23)$$

where, $\bar{C}$ is a parameter. $\bar{C}$ is estimated using the cutting force data and the least squares estimation method.

Experimentation: Experiments were conducted on AISI 4340 medium carbon steel cylinders with 223 BHN hardness. Valenite VC55 TNMA-432 uncoated cemented carbide inserts were used to cut these workpieces on a CNC lathe. The operating conditions were as follows:

<div>

Speed   656 sfpm (33.33 m/s),

Feed   0.003–0.006 ipr (0.08–0.16 mm/rev),

</div>

28

Depth   0.05 in. (1.27 mm).

During the cutting experiments, the flank wear was intermittently measured using a computer vision system and a microscopic lens. The flank wear on the tool was also measured several times in between successive computer vision measurements using a toolmaker's microscope. The estimated flank wear values between any two computer vision measurements were compared with the microscope measurements. While cutting was in progress, feed was repeatedly changed between 0.003 and 0.006 ipr (0.08 and 0.16 mm/rev) and the flank wear was estimated on-line, using the adaptive observer.

Results: The results of the experiments indicated that flank wear estimates obtained in between the computer vision measurements were in close agreement with the microscope measurements. The maximum estimation error over the entire cutting operation was approximately 15.8% of the measured wear. During the experiments, it was observed that the gradient $K_{10}d$ changed as the cutting process continued. Therefore, adaptively updating the observer gradient $K_{10}d$ using the computer vision measurements improves the accuracy of the flank wear estimates.

Comments: This method gave accurate flank wear estimates at both constant and time-varying cutting conditions. This method is insensitive to the development of crater wear. The effect of crater wear on the rate of cutting force increment is compensated for in the flank wear model by adaptively updating the observer gradient. This method has another advantage. By choosing the observer gain K smaller or larger, the sensitivity of the flank wear estimation to the fluctuation in force measurements can be reduced or increased. However, the use of this model is justified only in high speed cutting conditions where flank wear takes place mainly

due to diffusion mechanism. To determine the flank wear model completely, the activation energy for diffusion K_3 must be obtained for every tool-work combination from the preliminary experiments or diffusion data handbooks. The flank wear model assumes the availability of the tool-work temperature. Currently, on-line temperature measuring methods have not been practically feasible. To obtain the tool-work temperature using an empirical equation such as Eq. (2.22), the parameters of the equation have to be determined from preliminary experiments.

2.3.3 Nonlinear Observer (Park and Ulsoy 1989a, 1989b, 1992)

Park and Ulsoy (1989a, 1989b, 1992) developed a nonlinear observer based on the dynamic state model of tool wear proposed by Danai and Ulsoy (1988a). In the previous section, this dynamic model was described by Eq. (2.9) to (2.15). However, Park and Ulsoy (1989a, 1989b, 1992) simplified the model by considering only the flank wear equations. Equations describing the crater wear were eliminated from the model.

The nonlinear observer is designed to eliminate the drawbacks encountered with the adaptive observer investigated by Danai and Ulsoy (1988a, 1988b): the nonlinear observer eliminates the problems associated with the linearization of the dynamic state model and with transformation of the linearized model to observer canonical form. Assuming that the structure and parameters of the nonlinear model are correct and the model exactly describes the actual cutting process, the expression for the nonlinear observer becomes

$$\dot{\hat{w}}_{f1} = -\frac{v}{l_0}\hat{w}_{f1} + \frac{v}{l_0}K_1\cos\alpha_r\frac{F}{fd} + G_1(F - \hat{F}), \qquad (2.24)$$

30

$$\dot{\hat{w}}_{f2} = K_2 \sqrt{v} \exp\left(\frac{-K_3}{273 + \theta_f}\right) + G_2(F - \hat{F}), \tag{2.25}$$

and

$$\hat{F} = \left[K_6 f^{n1} (1 - K_7 \alpha_r) - K_8 - K_9 v\right] d + K_{10} d\hat{w}_f, \tag{2.26}$$

where

$$\hat{w}_f = \hat{w}_{f1} + \hat{w}_{f2}, \tag{2.27}$$

$$\hat{\theta}_f = K_{12} v^{n2} f^{n3} + K_{13} \hat{w}_f^{n4}. \tag{2.28}$$

G_1 and G_2 in Eq. (2.24) and (2.25) are the observer gains. The values of G_1 and G_2 are selected such that the observer error dynamic system satisfies the stability condition.

Experimentation: Experiments were conducted on a CNC lathe by cutting AISI 4340 medium carbon steel cylinders of 223 BHN hardness. Valenite VC55 TNMA-432 uncoated cemented carbide inserts were used to cut the workpieces. Two sets of experiments were conducted. The first set of experiments were conducted to evaluate some of the nonlinear model coefficients: l_0, K_1, K_6, and K_{10}. The rest of the model coefficients (K_2, K_3, K_7, K_8, K_9, K_{11}, and K_{12}, K_{13}, n_1, n_2, n_3, and n_4) were taken from the published literature. The second set of experiments were conducted under the following operating conditions to verify the accuracy of the estimated flank wear:

Speed   656 sfpm (3.33 m/s),

Feed   0.003 ipr (0.08 mm/rev),

31

Depth   0.05 in. (1.27 mm).

During the experiments, flank wear was measured regularly using a tool-maker's microscope. To determine the observer gains G_1 and G_2, the estimated flank wear $\hat{w}_f$ was assumed to be bounded by the limits 0.0 mm and 0.3 mm.

Results: Experiments showed that flank wear estimates given by the nonlinear observer are accurate and less sensitive to the poor initial state estimates and random fluctuations in cutting force measurements. It was found that the selection of observer gains is influenced not only by the stability condition but also by the level of noise in the force measurements. When force measurements contain high levels of noise, lower observer gains are required to get the correct tool wear estimates.

Comments: This method generated accurate on-line flank wear estimates. However, a major set back with this method is the difficulty in determining the model parameters. To evaluate the complete set of model parameters for each combination of workpiece and tool materials, a large number of experiments have to be conducted. It is possible that model parameters vary over time due to varying operating and process conditions. This would lead to inaccuracy in tool wear estimates. In addition, inaccuracies in the model structure will influence the accuracy of tool wear estimates.

2.4 Empirical Models for Tool Wear Estimation

In this approach, a model that describes a process output or state is developed. The model parameters are determined from either the experimental data or from some underlying theory related to the process. A merit function that measures the agreement between the data and the model with a particular choice of parameters

32

is chosen. The parameters of the model are then adjusted to achieve a minimum in the merit function, yielding best-fit parameters. Beyond the parameter estimation, the appropriateness of the selected model is tested using a statistical measure of goodness-of-fit.

In a cutting process, an output or a state variable R is expressed in the following general nonlinear multiplicative form (Oraby and Hayhurst 1991):

$$R = c \left[\prod_{j=i}^{p} x_j^{\beta_j} \right] \bar{\varepsilon}. \qquad (2.29)$$

where x_j are the machining variables such as speed, feed and depth of cut; c and β_j are the model parameters; and $\bar{\varepsilon}$ is a multiplicative random error. In general, R is an output variable such as a force component, or a state variable such as a tool wear component. This equation is then linearized using natural logarithm:

$$\ln R = \ln c + \sum_{j=1}^{p} \beta_j \ln x_j + \ln \bar{\varepsilon}. \qquad (2.30)$$

This equation is a linear polynomial of the first order, and either a simple linear regression or a multiple linear regression technique can be applied to estimate the model parameters using the experimental data. Sometimes the model parameters may be estimated from the underlying theory of the process.

The following sections describe two important models for tool wear estimation proposed by Koren et al. (1991) and Oraby and Hayhurst (1991).

2.4.1 Models for Flank Wear Estimation (Koren et al. 1991)

Koren et al. (1991) designed a model-based methodology for on-line tool wear estimation using cutting force measurements under conditions where one of

33

the cutting variables—speed or feed or depth of cut—exhibits persistent stepwise variations. A model relating force and flank wear is combined with an on-line parameter estimation method to obtain tool wear estimations. In this method, the direct effect of varying cutting conditions on cutting force is separated from its indirect effect on force through tool wear.

Force and Tool Wear Model: Initially, a model for force increment (ΔF) in a cutting process is developed as follows:

$$\Delta F = C_p v^{\beta_1} f^{\beta_2} d^{\beta_3} \dot{W}(v^{\gamma_1} f^{\gamma_2} d^{\gamma_3})t, \qquad (2.31)$$

where

C_p constant coefficient,

d depth of cut,

f cutting feed,

t cutting time,

v cutting speed,

β_i constant exponents, and

γ_i constant exponents.

In this equation the wear rate is expressed as a function of cutting variables. β_i determines a cutting variable's direct effect on force increment. γ_i determines a cutting variable's indirect effect on force increment through tool wear. Assuming that at any time only one of the cutting variables is varied, the above equation is simplified to the following expression:

$$\Delta F = C b^{\beta} \dot{W}(b^{\gamma})t, \qquad (2.32)$$

34

where b represents a cutting variable that varies in a stepwise manner: v or f or d. The coefficient C depends on the tool and workpiece materials and on the other cutting variables that do not vary. Using this equation, flank wear is estimated in three different cases: (1) $\beta \simeq 0$ and $\gamma \neq 0$, (2) $\beta \neq 0$ and $\gamma \simeq 0$, (3) $\beta \neq 0$ and $\gamma \neq 0$. A separate flank wear estimation method is designed for each one of these cases. A typical example of the first case is when $b = v$ (cutting speed). In this case, changes in the cutting speed strongly affect the flank wear rate (i.e., $\gamma \neq 0$) but their direct effect on ΔF is small (i.e., $\beta \simeq 0$). Similarly an example of the second case is when $b = d$ (depth of cut). In this case, the effect of the depth of cut on flank wear is very small (i.e., $\gamma \simeq 0$) but ΔF is linearly proportional to the depth of cut (i.e., $\beta \neq 0$). An example of the third case is when $b = f$ (cutting feed) and sometimes when $b = d$.

Simulation and Experimentation: Three simulation studies were conducted to verify the accuracy of tool wear estimation in the three cases. Also, a cutting experiment was conducted to verify the effectiveness of the second method in a practical environment. This experiment was performed on a 25 HP CNC lathe by cutting AISI 4340 annealed steel. TNWA-431F and TNMA-433E inserts were used to cut the workpiece. These experiments were conducted at the following operating conditions:

Speed   1200 sfpm (61 m/s),

Feed   0.001 ipr (0.0254 mm/rev),

Depth   0.043–0.118 in. (1.1–3.00 mm).

The cutting conditions were selected to generate rapid flank wear without significant crater wear. During the cutting operation, the depth of cut was changed in steps of 0.025 in. (0.64 mm) or 0.05 in. (1.28 mm).

Results: Simulation results showed that in all three methods, the flank wear estimates were close to the true flank wear values. Even the results from the cutting experiments showed that the estimated flank wear values were close to the microscope measurements.

Comments: This model-based method gave accurate flank wear estimates. This method assumes that the wear curve intercept with wear axis W_0, and the parameter C are available. If these values are not available, which is often the case, preliminary experiments have to be conducted to determine them. For this method to work, cutting variables must vary one at a time in a stepwise constant manner and the effect of crater wear must be negligible.

2.4.2 Models for Wear and Force (Oraby and Hayhurst 1991)

Oraby and Hayhurst (1991) investigated an empirical approach to flank wear estimation from force measurements. In this approach, empirical models were developed to describe the wear-time, force-time, and wear-force relationships in the turning process. The empirical models include three force components and three wear elements. The three force components are the cutting (F_1), feed (F_2), and radial (F_3) forces. The three wear elements are nose (W_1), flank (W_2), and notch (W_3) wears. In addition to these three wear elements, the average wear (W_4) on the flank face is also considered in the empirical models. The constants of the models were estimated from the experimental data using linear and nonlinear regression procedures. The significance of each model was examined using several statistical diagnostics such as correlation factor R^2, standard error, confidence interval, and lack-of-fit test. The adequacy of the model was verified through the its linearity, independence of residuals, normality of residuals and equality of variance.

36

Wear-Time Empirical Models: In these models, the development a wear element on a tool flank is expressed as

$$W_i = W_0 + \Delta W_i, \tag{2.33}$$

where W_0 is the initial wear, and ΔW_i is the wear increase with time and $W_i \ni \{W_1, W_2, W_3, W_4\}$.

Since wear increase ΔW_i at a particular cutting time is dependent on the cutting conditions employed, it is expressed as

$$W_0 = a_0 V^{p_{01}} f^{p_{02}} d^{p_{03}} D^{p_D}, \tag{2.34}$$

$$W_i = W_0 + a_i V^{p_{i1}} f^{p_{i2}} d^{p_{i3}} t^{p_{i4}}, \tag{2.35}$$

where

a_i	constant coefficients,
d	depth of cut,
D	workpiece diameter,
f	feed,
p_D	constant exponents of workpiece diameter,
p_{i1}	constant exponents of depth of cut,
p_{i2}	constant exponents of feed,
p_{i3}	constant exponents of time,
p_{i4}	constant exponents of cutting speed, and
V	cutting speed.

The constants of the above models were determined from the experimental data using nonlinear and linear regression procedures. The details of experiments and results of the model are discussed in the later sections.

37

Force-Time Empirical Models: These empirical models describe the effect of progressive tool wear on the force components. In these models, force-time trend is expressed by the following relationship:

$$F_j = F_{0j} + b_0 (W_i - W_0)^{b_1} , \qquad (2.36)$$

where F_{0j} is the initial value of a force component, F_j is the progressive force component, and b_0 and b_1 are contents. The subscript j refers to a force component such that $F_j = \{F_1, F_2, F_3\}$. Since both the initial wear W_0 and the initial force F_{0j} are dependent on cutting conditions the above expression is re-written as

$$F_j = c_0 V^{q_{j1}} f^{q_{j2}} d^{q_{j3}} + b_0 (W_i - a_0 V^{p_{01}} f^{p_{02}} d^{p_{03}})^{b_1} , \qquad (2.37)$$

where

b_0 constant coefficient,

b_1 constant exponent,

c_0 constant coefficient,

q_{j1} constant exponents of depth of cut,

q_{j2} constant exponents of feed, and

q_{j3} constant exponents of cutting speed.

The details of the experiments used for finding the model constants are discussed in the later sections of this chapter.

Wear-Force Empirical Models: The force-time models expressed in Eq. (2.36) and (2.37) contribute to understanding the effect of wear on force components. However, they are not useful in obtaining tool wear estimates. Instead, such tool

38

wear estimates are obtained from wear-force empirical models. The functional relationship between wear elements and force components is approximated by

$$W_i = W_0 + d_0 \left(F_j - F_{0j}\right)^{d_1} , \qquad (2.38)$$

where d_0 and d_1 are constants.

Since W_0 and F_{0j} are basically dependent on the cutting conditions, Eq. (2.38) is simplified to

$$W_i = e_0 V^{r_{01}} f^{r_{02}} d^{r_{03}} + d_0 \left(F_j\right)^{d_1} , \qquad (2.39)$$

where

e_0 constant coefficient,

r_{01} constant exponents of depth of cut,

r_{02} constant exponents of feed, and

r_{03} constant exponents of cutting speed.

In the above expression, coefficients e_0 and d_0 are usually dependent on the tool geometry, the workpiece hardness, and the type of lubricant. While the effect of tool geometry and the type of lubricant are considered constant for the same set of tool and cutting conditions, the effect of the variation in the workpiece hardness is eliminated by using the ratio of the feed component F_2 or the radial component F_3 to the cutting component F_1. Thus, Eq. (2.39) becomes

$$W_i = e_0 V^{r_{01}} f^{r_{02}} d^{r_{03}} + e_1 \left(\frac{F_2}{F_1}\right)^{e_2} , \qquad (2.40)$$

39

or

$$W_i = e_0 V^{r_{01}} f^{r_{02}} d^{r_{03}} + e_3 \left(\frac{F_3}{F_1} \right)^{e4}, \qquad (2.41)$$

where e_1, e_2, e_3 and e_4 are constants.

The model constants were evaluated using the experimental data. The details of experiments are explained in the following section.

Experimentation: Experiments were conducted on a Mascot 1600 lathe by cutting 709M40 alloy steel bars with Sandvik GC 435 triple-coated carbide tool inserts. The experiments were conducted under the following operating conditions:

Speed   164–676 sfpm (0.833–3.433 m/s),

Feed   0.0024–0.024 ipr (0.06–0.6 mm/rev),

Depth   0.06–0.12 in. (1.5–3.0 mm).

To develop models of force-wear characteristics which have statistical significance over a desired range of operating conditions, the experimental data was collected from the experiments conducted according to the central composite design. The experiments were conducted in 24 trails. In each trail, the cutting process was interrupted at intervals of approximately two minutes and the tool wear elements were measured on the flank face of the cutting inserts. Each trail was terminated by either catastrophic failure of the tool or by the upper tool wear limit.

Results: By studying the wear-time, force-time, and wear-force models, several results were observed. The wear-time model of the average wear (W_4) satisfied the statistical criteria better than the wear-time models of the other wear elements. The wear-time model of the average wear provided accurate tool estimates in the

40

initial and the constant wear phases of the tool life cycle. However, the model did not give satisfactory estimates in the final phase of the tool life cycle where the wear rate is high.

From force-time curves it was found that among the force components, cutting force (F_1) was least affected by tool wear elements. Among the different wear elements, notch wear had the least influence on force components. While nose wear affected radial component (F_3) more than it did feed component (F_2), flank wear affected them equally. The global effect of tool wear as represented by the average wear (W_4) was higher on radial force than on the other two force components. It was also found that the effect of excluding the cutting speed on the predictive capability of the models was negligible.

From wear-force models it was found that the models are able to give accurate tool wear estimates for average and nose wear in all three (initial, constant, and final) phases of tool life.

Comments: The measurement of the variation of the ratio between the radial and cutting force components (F_z/F_y) seems to provide a practical method for an on-line approach to tool wear estimation and failure detection. The only drawback with this approach is that it requires a considerable amount of experimentation to determine the model coefficients.

2.5 Miscellaneous Tool Wear Estimation Methods

This section reviews two useful tool wear estimation methods. The first one is a sensor-based method developed by El Gomayel and Bregger (1986). This method monitors workpiece diameter to get flank and nose wear estimates. The second

41

one, developed by Rao (1986), is a method for flank wear estimation using the wear index, which is the measure of the resistance at the tool-workpiece interface to the forced oscillations of the cantilever portion of the tool holder.

2.5.1 Tool Wear Estimation by Monitoring Workpiece Diameter

El Gomayel and Bregger (1986) developed an electromagnetic sensing device to measure flank wear and nose wear indirectly by monitoring the change of the workpiece diameter. This electromagnetic device works on the principle that an electromotive force is generated in a conductor when a magnetic flux changes with time. The voltage output of the device varies with the distance between the probe and the workpiece, due to the variation in the electromagnetic field. When the workpiece diameter increases with the development of flank or nose wear, the gap between workpiece and the probe decreases and it, in turn, decreases the voltage output.

Experimentation: Experiments were conducted on workpiece made of AISI 4140 alloy steel with an average hardness of 190 BHN. Kennametal K6 and K21 inserts were used for cutting the workpiece. The operating conditions were:

    Speed   350–550 sfpm (1.778–2.794 m/s),

    Feed   0.0208–0.0232 ipr (0.528–0.589 mm/rev),

    Depth   0.03–0.10 in. (0.762–2.54 mm).

The electromagnetic sensor was placed 0.75 inches from the tool tip. The tools were engaged in cutting until a significant amount of wear was developed. In most cases, the workpiece diameter was measured after every pass at both ends of the workpiece. Increase in workpiece diameter was indicated by the voltage difference at the two ends of the workpiece. Curves were plotted relating the accumulated

42

flank or nose wear with the accumulated voltage differences.

Results: The results of the experiments confirmed that monitoring the changes in the workpiece diameter through voltage difference is a feasible means for on-line tool wear sensing. The flank and nose wear curves generated in the experiments agreed with the generally accepted trends. It was found that the sensing device was capable of detecting even minute wear changes.

Comments: This method is effective in on-line tool wear sensing provided that the cutting operation is not influenced by disturbances caused by various kinds of misalignments, deflections, and vibrations of spindle of the headstock, live center of the tailstock, workpiece, cutting tool, and the tool holder.

2.5.2 Flank Wear Estimation Using Dynamics of Stable Turning

Rao (1986) investigated a tool wear estimation method based on Wear Index (WI). This WI is the measure of the resistance, at the tool-workpiece interface, to the forced oscillations of the cantilever portion of the tool holder. This method is based on the observation that increasing flank wear results in an increasing area of contact between the tool tip and the workpiece and this in turn results in increasing WI. The definition of WI and its relationship to flank wear are as follows:

$$\text{WI} = \frac{\text{Force amplitude at first natural frequency of toolholder}}{\text{Vibration amplitude at the same frequency}}, \quad (2.42)$$

$$W_{avg} = K_1 + K_2(\text{WI}), \quad (2.43)$$

43

where W_{avg} is the average of the flank and the nose wear, and K_1 and K_2 are the empirical constants to be determined from experiments.

In this investigation, Rao (1986) found that even under stable turning conditions, oscillations around the natural frequencies of the toolholder dynamometer system existed. The lower frequencies dominated when the tool was new, and these frequencies shifted to the first natural frequency of the cantilever portion of the toolholder as tool wear increased. This shift in dominant frequency as opposed to pure increase or decrease in vibration or force amplitude commonly occurs as the tool wear increases.

Rao (1986) further described the reasons for the aforementioned observations. When the tool is subjected to random disturbances on its rake face due to chip breaking, stick-slip processes occurring between the chip and the tool rake face, etc., these random disturbances set the toolholder-dynamometer system in oscillations at its natural frequencies. Since the tool is in very intimate contact with the workpiece along its flank edge, these oscillations are resisted by the work material, and the resisting force is registered by the dynamometer carrying the tool. In particular, oscillations at the natural frequency of the cantilever portion of the toolholder result in indentation at the nose and cutting edge. When the tool edge is new, this contact between the tool insert and the workpiece is restricted to an almost line contact, thus offering a small amount of resistance to the oscillations. However, as the flank wear-land develops, the increasing contact area increases the amount of workpiece material that must be elastically deformed. This increase in resistance to the oscillations causes the dynamic force amplitude to increase more rapidly than the vibration amplitude.

44

In view of these observations, Rao (1986) defines the WI as the measure of the reaction to flank interference under stable cutting conditions, and he uses it as an indirect measure of the flank wear width.

Experimentation: Experiments were conducted on three different work piece materials: AISI 1020, 1045, and 4340. All these workpieces were hot rolled. These workpieces were cut using tungsten carbide tool of grade C2(K420), C6(k421), and C7(K45). The tools had a negative rake angle of $-3°$, a clearance angle of $3°$, and a side cutting edge angle of $0°$. The experiments were run at the following ranges of operating conditions:

Speed   300–700 sfpm (1.52–3.56 m/s),

Feed   0.005–0.02 ipr (0.127–0.508 mm/rev),

Depth   0.025–0.1 in. (0.635–2.54 mm).

The force and vibration signals in the feed direction were collected using a signal acquisition system in the frequency rang of 0–6 KHz. The vibration signal from the accelerometer was double integrated to obtain the vibration amplitude. The WI was computed several times a second from the Root Mean Square (RMS) values of force and vibration data. The computed values of WI were used to estimate the average wear on the flank face from Eq. (2.43). The constants of this equation, K_1 and K_2, were initially calibrated from the experimental data.

Results: In this investigation it was found that WI is independent of cutting process variables: speed, feed, depth of cut, and nose radius. However, the model was found to be sensitive to the hardness of the work material. The best of the several experimental trails conducted had the mean and the standard deviation of errors 0.024 mm and 0.0074 mm respectively. By fine tuning the model constants K_1

45

and K_2 in Eq. (2.43), a monitoring accuracy of $\pm$ 0.0147 mm on flank wear-land width at a 95% confidence level was expected.

Comments: This investigation reveals the dynamic characteristics of vibration and force in stable turning. The proposed WI seems to correlate well with flank wear elements. Since the method was verified with extensive experimentation, the results obtained in this investigation can be assumed to be to be reliable even at the practical operating conditions. However, there are a few difficulties in applying this method to a practical situation. The dependency of WI on the hardness of the material affects the accuracy of the tool wear estimates from one layer to the other layer of the workpiece. The interference of the electrical characteristics of the dynamometer amplifier system with the natural frequencies of the toolholder-dynamometer system will also affect the results. The change in the accelerometer gain due to the temperature rise at the tool tip will affect the accuracy of vibration measurements and, thus, the flank wear estimation results.

2.6 Conclusions

This chapter reviewed important tool wear estimation methods based on analytical models, observers, empirical models, and other techniques. When considered in a practical situation, all these methods have merits as well as limitations.

Analytical methods are developed based on the understanding of the physics of the cutting process. The cutting process is inherently very complex. The complete understanding of the physics of the process is practically difficult. For this reason analytical models fail to accurately describe the cutting process. Since the coefficients and exponents of analytical models change with tool-workpiece

46

combinations and cutting conditions, expensive off-line experiments are required in each case to evaluate model coefficients and exponents. Appropriate sensors for monitoring some of the process variables used in analytical models are not always available. For example, most models emphasize the relationship between tool wear and temperature. Currently robust shop-floor sensors for temperature monitoring are not available. This adds another limitation to the applicability of the analytical models for on-line tool wear estimation.

Tool wear estimation using an observer is more practical than using an analytical model. From the review of the observers for tool wear estimation, it appears that these observers can be designed in three possible ways:

Approach 1: In this approach, an observer is designed considering the cutting process as a black box. In this case, all parameters of the observer are adaptively estimated using a parameter estimator. The adaptive observer designed by Danai and Ulsoy (1988a and 1988b) belongs to this class. The advantage of this approach is that no preliminary experimentation is necessary to determine the observer parameters. Since observer parameters are adaptively updated, the time varying cutting process can be modeled correctly by the observer parameters. However, the method requires off-line measurements to transform observer states into meaningful tool wear states. It is not always easy to reconstruct tool wear states from observer states. This is a limitation of the method.

Approach 2: In this approach, an observer is designed totally based on a dynamic state model of tool wear. This approach assumes that the structure and parameters of the process model are correct and the model exactly describes the actual cutting process. Park and Ulsoy (1989a, 1989b, 1992) followed this approach

47

while designing the nonlinear observer for tool wear estimation. This approach is practical only if a correct process model is available. Otherwise, it suffers from the disadvantages associated with an analytical model approach.

Approach 3: This is a hybrid approach which combines the above two approaches. In this approach, the design of an observer is partially based on a dynamic state model of the cutting process. Some of the observer parameters are taken from the dynamic state model, while other parameters are adaptively estimated or updated using off-line or on-line measurements. The design of the adaptive observer in Park and Ulsoy (1990, 1993a, 1993b) follows this approach. An observer designed using this approach is more practical to implement and is more likely to give accurate tool wear estimates. Since only a few observer parameters are taken from the process model, fewer preliminary experiments are necessary to determine them. Because some of the observer parameters are adaptively updated, the observer will model the time-varying nature of the process more closely.

Empirical models are the mathematical models that relate the desired process states with process inputs and measurable process outputs. Initially, a model of a specific structure is assumed and then the model parameters are determined from the experimental data. Unlike analytical models, empirical models do not depend much on the physical mechanisms underlying the process. They are more data-driven then mechanism-driven models. Since the parameters of empirical models are determined from the experimental data of a specific problem, those parameters usually apply to the specific problem only. While empirical models seem attractive for tool wear estimation, they also have some drawbacks. It is not always easy to come up with the correct structure of the model that gives statistically acceptable results. The model parameters need to be determined from the experimental data

48

of specific applications.

This chapter mainly concentrated on the methods that are geared towards continuous tool wear estimation. There are several other contributions that are designed for the classification of tools into two or more classes depending on their wear levels. Some of the most important methods are listed below:

1. A technique based on a linear discriminant function for the detection of tool wear, tool failure, and chip disturbance events using the spectra of acoustic emission signals (Emel and Kannatey-Asibu, Jr. 1988).

2. A technique based on time series analysis of acoustic emission signals for tool wear detection (Liang and Dornfeld 1989).

3. A pattern recognition and multisensor strategy using acoustic emission and force signals for the detection of tool breakage and classification of tools based on their wear levels (Balakrishnan et al. 1989).

An alternative approach to analytical models, observers, and empirical models is the neural network approach. This approach does not rely on the understanding of the physics of the cutting process. Instead, neural networks learn the relationship between measurable process variables and tool wear from the experimental data. Neural networks, in that respect, are similar to empirical models. However, in the neural network approach, unlike in empirical models, no assumption regarding the structure of an initial model is necessary. Recently, several neural networks methods for tool wear estimation were investigated. Some of the most important neural network-based methods are reviewed in the next chapter.

49

Chapter 3

A REVIEW OF NEURAL NETWORK BASED METHODS FOR TOOL WEAR MONITORING IN TURNING

3.1 Introduction

This chapter reviews the important neural network methods for tool wear monitoring in the turning processes. It summarizes each important contribution reported in the literature and discusses its strengths and weakness. The concluding section of the chapter presents an overall review of all the methods.

On-line tool wear estimation in the turning process is primarily required for predicting tool failures, devising optimal tool change strategies, and implementing geometric adaptive control and adaptive control with optimization schemes. These optimal tool change strategies and adaptive control schemes can be implemented only if reliable and consistent on-line tool wear estimation methods are available.

Traditionally, analytical models are used for tool wear estimation. The current trend is to use either observers or neural networks. Neural network-based methods are among the latest ones in the evolution of tool wear monitoring methods. These methods have resulted from advanced developments in the areas of neural networks, computers, sensors, and high frequency data acquisition systems. Often, sensor data fusion is a feature associated with neural network applications in tool wear monitoring.

This chapter reviews the literature on neural network-based methods for tool wear monitoring in turning processes. Three main types of neural networks have

50

been investigated in connection with tool wear monitoring problems:

- Multilayer Neural Networks

- Adaptive Resonance Theory Networks

- Kohonen's Feature Maps

Though researchers used different neural network models and different methods, they all followed a common investigation strategy which involves the following steps:

1. Select process parameters to be measured. Usually, these parameters are forces, acoustic emission, vibration, spindle-motor current, and temperature.

2. Select a data acquisition method.

3. Select a feature extraction or a data reduction method.

4. Select a neural network model to relate sensor measurement to tool wear.

5. Conduct experiments to collect training and test data.

6. Train the neural network with experimental data.

7. Analyze the performance of the neural network with test data.

This chapter is organized into five main sections. Section 3.2 describes the application of multilayer neural networks for tool monitoring. Section 3.3 describes the application of adaptive resonance networks. Section 3.4 covers the application of Kohonen's feature maps. Section 3.5 presents the overall review of all the methods and the author's comments and suggestions.

3.2 Application of Multilayer Neural Networks (MNNs)

Multilayer Neural Networks (MNNs) have been investigated more frequently than other neural network models for tool wear monitoring applications. The

51

following sections review the application of multilayer neural networks for tool wear monitoring investigated by Rangwala and Dornfeld (1990), Dornfeld and DeVries (1990), Wang and Dornfeld (1992), Emel (1991), Elanayar et al. (1990), and Chryssolouris et al.(1992).

3.2.1 Application of MNNs (Rangwala and Dornfeld 1990)

Rangwala and Dornfeld (1990) investigated the application of a single-layer neural network and a two-layer neural network for integrating information from Acoustic Emission (AE) and force sensors in order to recognize the occurrence of tool wear in a turning process. In their work, a multilayer neural network was used as a learning and pattern recognition device to associate sensor signals with appropriate decisions on tool wear status. It was shown that a multilayer neural network is able to filter out noise in sensor signals and also able to successfully associate sensor signals with tool wear states.

Experimentation: Experiments were conducted by cutting hardened AISI 1060 bars with Kennametal TPGF-322 inserts of grade K68. The nominal diameter and the length of the bars were 2.0 in. (50.8 mm) and 12.0 in. (305 mm) respectively. The process variables were varied in the following ranges:

 Speed 278–556 sfpm (1.42–2.83 m/s),

  Feed 0.002–0.008 ipr (0.05–0.20 mm/rev),

 Depth 0.01–0.03 in. (0.25–0.75 mm).

Sensor signals were collected when flank wear levels were less than or equal to 0.01 in. (0.25 mm) and were assumed to belong to fresh tool cutting. Next the signals were collected when flank wear levels were greater than or equal to 0.02 in. (0.50 mm) and were assumed to belong to worn tool cutting.

52

During the actual cutting process, AE and force signals were amplified and recorded respectively on video and audio channels of a modified Sony recorder. During post-processing, the signals were played back, filtered, and digitized on a HP waveform recorder. A digitized AE signal was sampled at 5 MHz for a record length of 1024 points. A digitized force signal was sampled at 1 kHz for a record length of 512 points. A total of 65 samples of fresh tool cutting and 58 samples of worn tool cutting were collected for training and testing the neural networks.

Signal Processing and Feature Selection: Initially, the power spectra of AE and force signals were individually determined and then put together to create a vector of dimension 768. This vector was treated as a measurement vector. Each component of the vector represented the signal power at a discrete frequency in either the AE or the force spectrum. The dimension of the measurement vectors was reduced by retaining only those components of the vectors which are most sensitive to tool wear and less sensitive to noise and process parameters. This feature selection was done by maximizing the following criterion function which represents the signal to noise ratio of the selected features:

$$J = trace(S_w^{-1} S_b),$$
(3.1)

where S_w is the within class scatter matrix and S_b is the between class scatter matrix of feature vectors (Devijver and Kittler 1982). Using this criterion function and the heuristic Sequential Forward Search algorithm (Whitney 1971), measurement vectors of dimension 768 were reduced to feature vectors of dimension 6. To these feature vectors, speed and feed rate were added as two additional features. These feature vectors were used for training and testing the neural networks. The training set contained 30 samples equally divided between fresh and worn tool

53

cutting. The trained neural networks were tested on 50 fresh tool samples and 43 worn tool samples.

Results: In this work, the performance of a single-layer neural network (8-1, i.e., eight input nodes and one output node) and a two-layer neural network (8-3-1, i.e., 8 input nodes, 3 hidden nodes, and 2 output nodes) were studied. These studies indicated that the single-layer neural network yields classification success rates between 80-88% and the two-layer neural network between 84-97%. It was found that the accuracy of tool wear classification is maintained for a wide range of operating conditions. It was also observed that the classification success rates are higher when information from both AE and force sensors are used together than when information from the AE sensor alone is used.

Comments: This is the pioneering work employing neural networks for tool wear identification through sensor data fusion. This work has set a new trend in tool wear monitoring methods by using multiple sensors, feature extraction techniques, and neural networks. The results of this research indicate that neural networks and multisensor data fusion offer good promise for effective tool wear monitoring and warrant further investigation.

In this work, sensor signals were collected when flank wear was less than or equal to 0.01 in. (0.25 mm) and also when flank wear was greater than or equal to 0.02 in. (0.50 mm). In between these two limits no samples were collected. This will adversely affect the performance of a neural network because it is not exposed to the data samples when flank wear is in between 0.01 in. (0.25 mm) and 0.02 in. (0.50 mm). It is equivalent to skipping a valid part of input space while drawing training samples. However, this problem may be solved by narrowing the

54

gap between the flank wear limits for classifying fresh and worn tools.

To some extent, the off-line selection of features contributed to the success of this method. Input features were specifically chosen off-line so that the neural networks could adequately distinguish fresh tools from worn tools. Since the input features were extracted using a fixed set of training data, the same set of features over a wide range of operating conditions and for different workpiece or tool material may not give satisfactory tool wear classification. One needs to repeat the entire feature selection process whenever a new situation is encountered. Another question to be addressed in this method is how to select an optimal number of features.

3.2.2 Application of MNNs (Dornfeld and DeVries 1990)

Dornfeld and DeVries (1990) and Choi et al. (1990) investigated the application of a two-layer neural network for on-line tool wear monitoring. Input patterns to this neural network were generated from the multichannel autoregressive (AR) coefficients and the power spectral components of AE, force, and spindle motor current signals. The performance of the neural network was studied by varying machining conditions, number of input features, number of hidden layer nodes, and the combination of sensors.

Experimentation: Experiments were conducted on a Tree 1000 lathe by cutting AISI 1060 bars with Kennametal TPGF-322 inserts of grade K68. The nominal diameter of the bars was 4.0 in. (102 mm). The process variables were varied in the following ranges:

Speed   250–350 sfpm (1.27–1.78 m/s),

Feed   0.01–0.07 ipr (0.254–1.778 mm/rev),

55

Depth   0.005–0.008 in. (0.127–2.032 mm).

Signals collected during cutting when flank wear width was less than or equal to 0.039 in. (1.0 mm) were labeled as belonging to fresh tool cutting, and those collected when flank wear width was above 0.039 in. (1.0 mm) were labeled as belonging to worn tool cutting.

Signal Processing and Feature Selection: Initially, the power spectra of AE, force, and spindle motor current signals were individually determined and then combined to form a single vector. To this vector, the coefficients of the multichannel AR model of AE, force, and current signals were added as additional components. The dimension of this measurement vector was reduced by retaining only those components of the vector which are more sensitive to tool wear and less sensitive to noise and process parameters. This feature selection was done by maximizing a criterion function J representing the signal to noise ratio of the selected features:

$$J = \left[\frac{trace(S_b)}{trace(S_w)} \right] , \tag{3.2}$$

where S_w is the within class scatter matrix and S_b is the between class scatter matrix of feature vectors. Using this criterion function and the heuristic Sequential Forward Search algorithm (Whitney 1971), a measurement vector was reduced to a feature vector of desired dimension.

Results: In this work, a significant performance improvement was observed by fusing information from multiple sensors. A classification accuracy as high as 95% was achieved. The accuracy of classification was maintained even for cutting conditions varying by 50% to 200% from the operating conditions used during training. Among the different combination of sensors tried, the AE–force combination gave

56

the best results. By varying the number of hidden layer nodes and the number of input features, it was found that there is an optimal number of hidden layer nodes and an optimal number of input features that give the best results. Choi et al. (1990) reported that the classification accuracy was at 95% when sensor signals were processed by multivariate AR modeling compared to a classification accuracy of 58% when sensor signals were processed by FFT.

Comments: The methodology of this work is similar to that reported by Rangwala and Dornfeld (1990). One major difference is that in this work, input patterns contain both frequency and time domain features.

In this work, a flank wear limit of 0.039 in. (1.0 mm) was used as a threshold for classifying fresh and worn tools. This limit is rather high compared with the ISO standard. According to the ISO standard dealing with tool life testing (Boothroyd and Knight 1989), the recommended threshold limit is around 0.015 in. (0.38 mm). Neural networks which discriminate fresh and worn tool signals at the threshold limit of 0.039 in. (1.0 mm) may not give equally good results at a lower threshold limit where the distinction between fresh and worn tools is less marked.

This method also involves the off-line selection of features to maximize the distinction between fresh and worn tool patterns. Similar to Rangwala and Dornfeld's (1990) method, the input features are extracted using a fixed set of training data, and the same set of features over a wide range of operating conditions and for a different workpiece or tool material may not give us good results.

3.2.3 Application of MNNs (Wang and Dornfeld 1992)

Wang and Dornfeld (1992) investigated the application of a two-layer neural network for on-line estimation of progressive tool wear in turning processes. Input

57

pattern features were selected from different AE parameters: coefficients of AR models of the RMS AE, integrated RMS AE, and the skewness and kurtosis of an assumed Beta distribution for the RMS AE.

Experimentation: Experiments were conducted on a Tree 1000 lathe by cutting AISI 1045 bars of 100 mm diameter with Kennametal TPGF-322 inserts of grade K21. The experiments were conducted at the following operating conditions:

Speed   328 sfpm (1.67 m/s),

Feed   0.001 ipr (0.0254 mm/rev),

Depth   0.079 in. (2.0 mm).

The AE sensor output was passed through a preamplifier and then through an RMS meter. The RMS AE signal was collected at a sampling rate of 1 kHz. For each cutting tool, a total of 300 RMS AE samples were collected at regular intervals of 3 seconds. During the approximate cumulative machining time of 900 seconds, an approximate flank land of 0.02 in. (0.5 mm) was developed. Over the span of each cutting experiment, the tool wear was measured 10 times at intervals of 90 seconds using a microscope.

Signal Processing and Feature Selection: In this work, the variations of different parameters derived from AE signals with progressive tool wear were studied. The parameters derived from the AE signals were coefficients of AR models of the RMS AE, integrated RMS AE, and the skewness and kurtosis of an assumed Beta distribution for the RMS AE. In this work, it was observed that as the number of parameters used for estimating tool wear increases, the accuracy of estimates will also increase. Based on this study, it was decided that all of these parameters would be used together to create input patterns to make the tool wear monitoring

58

system more effective and robust.

A multilayer neural network with one hidden layer was trained with the input patterns created from the selected AE parameters. As mentioned earlier, there were 10 tool wear measurements and 300 RMS AE samples available from each cutting tool experiment. To improve the training accuracy, two more data points were introduced between the measured tool wear levels using the linear interpolation method. The 10 measured tool wear levels and the 20 introduced values formed the 30 training input-output patterns for the neural network.

Results: It was found that RMS AE increases with increasing tool wear up to a certain point and then remains flat for the remaining tool wear progress. Whereas the integrated RMS AE increases monotonically with increasing tool wear, the skewness and kurtosis of an assumed Beta distribution for the RMS AE show increasing and decreasing trends respectively with increasing tool wear.

According to the results, when the number of features in input patterns was increased from 2 to 6, the convergence of the neural network training became faster and the output error became smaller. When only skewness and kurtosis were used as input features, the tool wear prediction error was over 100%. When two AR coefficients were added to the skewness and kurtosis, the prediction error decreased considerably. Finally, when mean RMS AE and integrated RMS AE were also added, the prediction error was reduced to less than 10%.

Comments: This chapter reports useful observations on the relationship between various AE parameters and tool wear levels. This method seems to provide good tool wear estimates for the experiments reported in the chapter. However, the extent of experimentation is very limited. The reliability and effectiveness of the

59

method need to be verified with a more extensive set of experiments.

3.2.4 Application of MNNs (Emel 1991)

Emel (1991) studied the application of multilayer neural networks of different structures for identifying tool wear levels using AE spectrum features as inputs. He also studied the effect of training parameters on the performance of multilayer neural networks.

Experimentation: Experiments were conducted on a CNC turning machine by cutting heat-treated AISI 4340 bars of 370 BHN with TNMA-432 carbide inserts of grade VC7. In this experiment, two tests were conducted with the following process parameters:

Test 1

      Speed   525.6 sfpm (2.67 m/s),

      Feed   0.008 ipr (0.2032 mm/rev),

      Depth   0.063 in. (1.60 mm).

Test 2

      Speed   734.25 sfpm (3.73 m/s),

      Feed   0.016 ipr (0.4064 mm/rev),

      Depth   0.063 in. (1.60 mm).

In these experiments, cutting signals collected at a flank wear level of 0.001 in. (0.025 mm) were assumed to belong to fresh tool cutting, and those signals collected at 0.012 in. (0.305 mm) were assumed to belong to worn tool cutting. Each test was replicated four times. In each replication, 20 fresh tool samples and 20 worn tool samples were collected. A digitized AE signal was sampled at 4 MHz for a record length of 4096 points.

60

Signal Processing and Feature Selection: The AE samples collected from an experiment were first transformed into vectors of spectral components. Next, the dimension of the vectors was reduced by retaining only a subset of the original vector components. The subset of components was selected using a class-mean-scatter criterion:

$$J(x_j) = \left[\frac{s_{w,jj}}{s_{b,jj}} \right] , \tag{3.3}$$

where $s_{w,jj}$ and $s_{b,jj}$ are the jth diagonal elements of the mixture-scatter matrix and the class-mean-scatter matrix respectively. According to this criterion the components of the original vectors were ranked in descending order, and then the first 5 components from the top of the list were selected to create the feature vectors of dimension 5. Further, the components of the feature vectors were normalized and converted to logarithmic scale. The feature vectors thus obtained were used as input patterns to train the neural networks.

Results: In this work, the performance of three multilayer neural networks of different structures were studied for different values of learning rate (η) and momentum (α) in backpropagation training. The three multilayer neural networks investigated were, a single-layer neural network (5-2), a two-layer neural network (5-20-2), and a three-layer neural network (5-20-5-2). The performance of these neural networks was compared to that of the Linear Discriminant Function (LDF) method.

It was found that compared to the LDF method, multilayer neural networks give better classification accuracy. The LDF method gave the correct sharp-worn classification at 66%-76% accuracy in test 1 and at 66%-83% in test 2; whereas the multilayer neural networks gave the accuracy of correct sharp-worn classification

61

at 80%–90% in test 1 and at 80%–91% in test 2. It was also observed that two-layer and three-layer neural networks perform better than a single-layer neural network.

Comments: This method is similar to the tool wear classification method proposed by Rangwala and Dornfeld (1990) and Dornfeld and DeVries (1990). Unlike the previous ones, this method did not use multiple sensors; it used a single AE sensor. In this work, extensive experimental data was collected to verify the results. Also, many variations in the form of neural network structures and training parameters were investigated. The results obtained in this work reconfirm the results reported in earlier works.

3.2.5 Application of MNNs (Elanayar et al. 1990)

Elanayar et al. (1990) investigated the application of hierarchically arranged two-layer neural networks for estimating flank wear, crater wear, and surface finish attributes from the measurements of three force components. In this approach, the following expressions that relate flank and crater wears with force components were developed:

$$w_f = G_{xy}^{-1} \left[\frac{F_{y0}}{F_{x0}} \cdot \frac{F_x}{F_y} - 1 \right] , \qquad (3.4)$$

$$w_c = G_{zy2}^{-1} \left[\frac{F_{y0}}{F_{z0}} \cdot \frac{F_z}{F_y} G_{zy1}^{w_f} - 1 \right] , \qquad (3.5)$$

where, F_x, F_y, and F_z are the force components in cutting, feed, and radial directions; F_{x0}, F_{y0}, and F_{z0} are the initial force values in the respective directions;

62

w_f and w_c are flank and crater wears; and G_{xy}, $G_{zy1}^{w_f}$, and G_{zy2} are the unknown functions of wear.

From the structure of this expression, a basis was formed for using force ratios F_x/F_y and F_z/F_y as inputs to the first level neural network for estimating the flank wear width R_f, the crater wear length parallel to flank face L_c, and the crater width D_c. The estimated values of R_f and L_c and the measured cutting force F_z were used as inputs to the second level neural network for estimating surface finish attributes: arithmetic average RA and ten-point height RZ.

Experimentation: To verify the method, experiments were conducted on a 20 HP LeBlond lathe by cutting AISI 6150 steel alloy bars with carbide inserts. The diameter of the bars was in the range of 3.0–3.3 in. (78.7–83.8 mm) and their length was 15.75 in. (400 mm). The experiments were conducted at the following operating conditions:

Speed   650 sfpm (3.3 m/s),

Feed   0.0074 ipr (0.188 mm/rev),

Depth   0.025 in. (0.635 mm).

While cutting was in progress, force signals in three primary directions were sampled at 100 Hz. The machining test was interrupted at predetermined intervals to measure the tool wear level using a toolmaker's microscope. After machining across the full length of the block, surface roughness attributes RA and RZ were measured at the same locations where the tool wears were measured. At the identical operating conditions, three tests were conducted to collect three sets of data.

Results: To test the performance of the proposed method, three identical but independent neural networks were trained with the data collected in test 1, test 2, and test 3. Each one of these three trained neural networks were tested for predicting tool wear and surface finish attributes for the data collected in all three tests. The results showed that the predicted values of tool wear and surface finish attributes approximately corresponded to the measured values.

Comments: This work followed a different approach compared to the previous methods. Unlike the other works reported in literature which classify the tools as fresh or worn, this method provides continuous tool wear estimates. To implement control strategies in a real-world situation, a method that provides continuous tool wear estimates is more useful than a method that identifies a tool as fresh or worn.

In this method, an initial empirical equation relating force components, flank, and crater wear was formulated. From the structure of the equation it was decided to use the force ratios as inputs to the first level neural network that predicts flank and crater wear. However, no explanation is given for the selection of inputs to the second level neural network that predicts surface finish attributes.

Chapter 2 discussed how operating conditions such as speed and feed rate affect the force components and surface finish attributes. Logically, it would have been appropriate to also use these operating conditions as inputs to the neural networks that predict wear values and surface finish attributes.

The results presented in this work show that neural networks are effective in predicting tool wear values and workpiece surface finish attributes from force measurements.

64

3.2.6 Application of MNNs (Chryssolouris et al. 1992)

In all the applications described in the previous sections, multilayer neural networks were used for tool wear prediction through sensor data fusion. In contrast, Chryssolouris et al. (1992) used multilayer neural networks for tool wear estimation through the synthesis of tool wear predictions given by different models of the turning process. The performance of the neural networks for tool wear synthesis was compared with that of statistical methods.

Experimentation: Experiments were conducted on a 20 HP Daewoo Puma 10 CNC lathe by orthogonally cutting the ends of ASTM A53 steel pipes. The wall thickness and the length of the steel pipes were 0.125 in. (3.18 mm) and 6.0 in. (157.4 mm) respectively. M2 HSS orthogonal tools by Enco Mfg., (with 6° relief angle and 30° rake angle) were used in all the tests.

Analysis: In this research, measurements were taken from AE, force, and temperature sensors and fed into AE-based, force-based, and temperature-based models respectively to get three independent tool wear predictions. The three tool wear predictions provided by these three independent process models were synthesized to obtain more reliable and accurate tool wear estimates. Two approaches were tried for tool wear synthesis: a neural network approach and a statistical approach. The performance of these two methods were studied using both the experimental data and the test data generated by simulation.

A two-layer neural network (5-4-1) and a three-layer neural network (5-8-8-1) were investigated for tool wear synthesis. Each neural network received wear estimates provided by the three process models and cutting velocity and feed rate as inputs.

65

Results: From the experimental results it was found that a neural network synthesis method is more effective than a statistical synthesis method for providing better estimates. When the process models adequately reflected the complexity of turning processes, the statistical synthesis methods gave better estimates; but when the process models inadequately described the turning processes, the neural network synthesis methods gave better results. Again, when one of the sensors failed, the statistical synthesis methods gave better results than neural network synthesis methods.

Comments: The accuracy of the tool wear estimates in this method depends more on the process models that predict the tool wear than on the neural networks or the statistical methods that synthesize the predicted tool wear values. As observed in this work, neural networks improve the method's tolerance for the inadequacy of process models that the predict tool wear values.

3.3 Application of Adaptive Resonance Theory Networks

Burke (1993) studied the performance of adaptive resonance networks for tool wear identification in turning processes. To investigate these adaptive resonance networks, Burke used the experimental data generated for investigating the performance of multilayer neural networks in Rangwala and Dornfeld (1990). Burke found that adaptive resonance networks performed almost at the same level as the two-layer neural networks investigated by Rangwala and Dornfeld (1990). Both adaptive resonance networks and two-layer neural networks gave a classification accuracy of about 95%.

66

Burke (1993) concluded that these two classes of neural networks differed in learning mode, labeled training data requirement, and training time requirement. A multilayer neural network learns in a supervised training mode while an adaptive resonance network learns in an unsupervised training mode. A multilayer neural network needs all its training sets to be labeled whereas an adaptive resonance network needs only some of its training sets to be labeled. Unlike a multilayer neural network, the label information of the training sets is not used for network learning; instead it is used for interpreting the clusters produced. A multilayer neural network is adaptive during training only. On the other hand, an adaptive resonance network is adaptive during training as well as post-training operations. Finally, an adaptive resonance network needs much less training time compared with a multilayer neural network. A detailed comparison of the performance of these two neural networks for tool wear classification is reported in Burke and Rangwala (1991).

Comments: The results of this work indicate that the performance of an unsupervised adaptive resonance networks is comparable to that of a multilayer neural network for tool wear identification. As reported in this work, the advantage of an adaptive resonance network is that it requires fewer labeled training sets and much less training time. Inaddition, an adaptive resonance network is adaptive even after the training is complete.

The feature extraction scheme in this method is not adaptive and the entire feature selection process has to be repeated for new situations. This will hinder the adaptive ability of an adaptive resonance network. Even if an adaptive feature selection process is used, an adaptive resonance network may not be able to maintain the accuracy of its performance across different situations over a period of

67

time because these networks are not as stable (in retaining what has been learnt) as multilayer neural networks.

3.4 Application of Kohonen's Feature Maps (KFMs)

Kamarthi et al. (1991) and Leem and Dreyfus (1992) investigated the application of Kohonen's feature maps for on-line tool wear monitoring in a turning process. In both these applications, a Kohonen's feature map was used for classifying cutting tools into three classes: fresh tools, medium tools, and worn tools. Jammu et al. (1993) investigated the application of a Single Category-Based Classifier (SCBC) for the detection of tool breakage. The performance of this classifier is compared to that of a Kohonen's feature map and ART2 network.

3.4.1 Application of KFMs (Kamarthi et al. 1991)

In their research, Kamarthi et al. (1991) investigated the application of a two-dimensional Kohonen's feature map for on-line classification of flank wear levels through vibration and force information. A Kohonen's feature map was trained with the coefficients of ARMA models of time series data collected from the vibration and force sensors at different tool wear levels and different operating conditions. The objective of the work was to investigate whether KFM could classy correctly the input patterns for fresh, medium, and worn tools.

Experimentation: Experiments were conducted on a 20 HP Leblond lathe by cutting a grey cast iron bar (199 BHN) with Kennametal SPG-422 carbide insert of grade K68. The initial diameter of the bar was 5.5 in. (140 mm) and the length was 24.0 in. (610 mm). In this work, four tests were conducted varying the process

68

variables in the following ranges:

Speed   600–650 sfpm (3.05–3.30 m/s),

Feed   0.0051–0.01 ipr (0.127–0.254 mm/rev),

Depth   0.10 in. (2.54–2.54 mm).

The time series data from both vibration and force sensors, in cutting and feed directions, was collected at time intervals of approximately 20 seconds using a Tektronix Fourier Analyzer. The data was collected at the sampling rate of 10 kHz.

These experiments were conducted under four different operating conditions, collecting a total of 70 data samples. Each time the data was sampled, time series data from four separate channels—vibration and force in cutting and feed directions—was collected at the same time. The data samples were labeled belonging to fresh, medium, or worn tools, depending on the level of wear at the time of sampling. These labeled data samples were used for training and testing a Kohonen's feature map. The following tool wear classification scheme was used:

wear level < 0.01 in (0.22 mm):   Fresh tool,

0.01 in (0.22 mm) ≤ wear level ≤ 0.015 in (0.37 mm):   Medium tool,

wear level > 0.015 in (0.37 mm):   Worn tool.

Signal Processing and Feature Selection: In this work, coefficients of ARMA models of individual time series channels were first determined and then put together to achieve the desired data fusion. Three combinations of data fusion were investigated. In case 1, coefficients of vibration signals in cutting and feed directions were put together. In Case 2, coefficients of force signals in cutting and feed directions were put together. In case 3, the coefficients of all four signals were used. The input patterns formed by combining the coefficients of the models of multiple

69

sensors were used as inputs to Kohonen's feature maps. The dimensionality of each input pattern in case 1 was 16, in case 2 was 14, and in case 3 was 30. The 70 input patterns in each case were used for training three separate Kohonen's feature maps. Each feature map had 400 output nodes arranged in a 20 x 20 grid. These three Kohonen's feature maps had all other training conditions identical.

Results: Ideally, if a feature map had learned to correctly discriminate the input patterns of three different classes, each node in the feature map should be activated by the input patterns of any one class only. In case 1 where only vibration sensor data was used, 2 out of 70 (about 3%) input patterns were classified into nodes of wrong classes. In case 2 where only force sensor data was used, all input patterns were classified into nodes of correct classes. In case 3 where both vibration and force sensor data were used, 3 out of 70 (about 5%) input patterns were classified into nodes of wrong classes. These results indicated that the feature map had learned to classify the input patterns correctly with an approximate reliability of 95%.

Comments: This work investigated the application of Kohonen's feature maps for sensor data fusion and tool wear classification. This method used ARMA model coefficients as a means of sensor data reduction. Since the identification of ARMA models from sensor data can be performed on-line, it eliminates the need for off-line feature selection. Like an adaptive resonance network, a Kohonen's feature map learns from unsupervised learning but a Kohonen's feature requires a longer training time than an adaptive resonance network. A Kohonen's feature map requires fewer labeled training sets than a multilayer neural network.

70

The results of this method indicated that this approach has good potential for tool wear classification. However, the experimental data used for verifying the performance of the method is too small. It was found that a Kohonen's feature map can separate training inputs into their respective classes: fresh, medium, and worn tools. However, this work did not verify if the network can classify new inputs during the operational phase.

3.4.2 Application of KFMs (Leem and Dreyfus 1992)

Leem and Dreyfus (1992) investigated the application of a one-dimensional Kohonen's feature map for integrating the information from AE and force sensors and assessing the tool wear levels from these sensor measurements. In this work, a Kohonen's feature map was trained in two stages. In the first stage, the standard unsupervised training method specified in Kohonen (1990) was used. In the second stage, the feature map was tuned using additional supervised training. This supervised training was conducted according to the Input Feature Scaling algorithm. This Input Feature Scaling algorithm was designed to scale the features in input patterns to improve the classification accuracy and to eliminate the pre-training feature extraction step.

Experimentation: Experiments were conducted on a Tree lathe by cutting hardened AISI 1060 bars with Kennametal inserts of grade K68. The nominal diameter of the bars was 2.0 in. (50.8 mm). The process variables were varied in the following ranges:

Speed   315–630 sfpm (1.60–3.20 m/s),

Feed   0.004–0.008 ipr (0.10–0.20 mm/rev),

Depth   0.01–0.03 in. (0.50–1.00 mm).

71

In this experiment, cutting signals were classified as fresh, medium, or worn tools using the following tool wear classification scheme: signals collected at a flank wear level of 0.0039 in. (0.10 mm) were assumed to belong to fresh tools, and those collected at flank wear level of 0.03 in. (0.75 mm) were assumed to belong to worn tools. In between these two limits, signals collected at flank wear level of 0.0197 in. (0.50 mm) were assumed to belong to medium worn tools.

$$\text{wear level} \leq 0.0039 \text{ in } (0.1 \text{ mm}): \quad \text{Fresh tool,}$$

$$0.0039 \text{ in } (0.1 \text{ mm}) < \text{wear level} < 0.03 \text{ in } (0.75 \text{ mm}): \quad \text{Medium tool,}$$

$$\text{wear level} \geq 0.015 \text{ in } (0.37 \text{ mm}): \quad \text{Worn tool.}$$

Signal Processing and Feature Selection: While cutting was in progress, AE RMS values and force signals were sampled at 1 kHz. The digitized time domain signals were sampled for a length of 128 points. By using the Welch method (Oppenheim and Schafer 1975), 32 point spectra of AE RMS values and force signals were produced. These spectra were put together to form vectors of dimension 64. To each vector, speed and feed rate were added as two additional components to create vectors of length 66. Also, four statistics—mean, standard deviation, skewness, and kurtosis—of amplitudes of AE RMS values and force signals were added to make the vectors of length 74 (66+4+4). Following this data acquisition scheme, 200 test samples were collected. Out of these 200 samples, 100 samples were used for training a Kohonen's feature map in unsupervised mode, 50 samples for training the feature map with the Input Feature Scaling algorithm, and the remaining 50 for testing the performance of the feature map.

Results: In this work, 12 simulation experiments were conducted. In these experiments, a one-dimensional Kohonen's feature map with three output nodes and

72

74 input nodes was used. The three output nodes of the feature map represented fresh, medium, and worn tools. The 74 input nodes of the feature map corresponded to the 74 features in an input vector. The Kohonen's feature map, after unsupervised training, gave a classification accuracy in the range of 64% to 76%. When the feature map was tuned with the Input Feature Scaling algorithm, the classification accuracy rose to an average of 92%.

Comments: This work investigated a Kohonen's feature map for classifying tools into three classes: fresh, medium, and worn tools. This chapter also investigated a new method for scaling the input features proportional to the degree of their correlation with tool wear. A new algorithm called the Input Feature Scaling algorithm was formulated to determine these input feature weights. This algorithm is a useful contribution that eliminates the off-line feature selection process prior to training a neural network. This method was verified by extensive experimentation. The results obtained by this work indicate that this method is effective in classifying tool wear levels.

In this work, flank wear of 0.0197 in. (0.50 mm) and 0.030 in. (0.75 mm) were considered respectively as medium and worn levels of wear. These levels are too high compared with what is actually followed in industry. The classification results obtained at these high levels of wear may not be obtained at the lower levels of wear because it is more difficult to discriminate between fresh, medium, and worn classes.

The Input Feature Scaling algorithm eliminates the need for off-line feature selection. Since these input feature weights are determined based on a batch of labeled training sets, the same weights may not be valid for the new situations created by different operating conditions and tool-workpiece combinations. There-

73

fore, it may be necessary to run the Input Feature Scaling algorithm to determine input feature weights whenever a new situation is encountered.

In this work it was assumed that a Kohonen's feature map will form topological ordering of its nodes according to ascending or descending order of tool wear levels. Though the training algorithm of a Kohonen's feature map is designed to form such topological ordering of its nodes, there is no existing theoretical proof that guarantees such a topological ordering in a feature map whose dimension is smaller than the dimension of input space. In this case, the dimension of the feature map is 1 and the dimension of input space is 74. It is quite possible that the Kohonen's feature map will not form topological ordering of its nodes.

The method proposed in this work is computationally very intensive. Training time of a Kohonen's feature map increases linearly with $(N \times n \times m \times U)$ and that of the Input Feature Selection algorithm with $(L \times n \times m^2 \times S)$. Here N is the number of unlabeled training sets, L is the number of labeled training sets, n is the number of input features, m is number of nodes in the feature map, U is number of unsupervised iterations, and S is the number of supervised iterations. Since no pre-training feature selection or data reduction is performed, n is going to be very large and increase the training time many-fold. This problem is faced, particularly, when high frequency signals such as AE and vibration are to be input to a Kohonen's feature map.

3.4.3 Application of SCBC (Jammu et al. 1993)

Jammu et al. (1993) proposed an adaptive and unsupervised Single-Category-Based Classifier (SCBC) for detecting on-line tool breakage. The performance

74

of this classifier is compared to that of a Kohonen's feature map and an ART2 network.

Experimentation: Experiments were conducted on a 37 KW NC vertical turret lathe. Diamond shaped tungsten carbide cam lock inserts were used to cut Inconel 718 cylindrical workpieces of approximately 4800 mm inside diameter. The process variables were varied in the following ranges:

Speed 120–173 sfpm (0.61–0.88 m/s),

Feed 0.0071–0.0079 ipr (0.18–0.20 mm/rev),

Depth 0.08–0.15 in. (2.03–3.81 mm).

Signal Processing and Feature Selection: In this work, four measurements were used for detecting tool breakage: processed ultrasonic energy, vibration in radial direction, vibration in cutting direction, and residuals of the ARX model of feed-direction motor velocity. Therefore, the measurement vectors contained four components.

A total of six tests were conducted. Each test contained 50 measurement vectors collected during a 30 second window centered around the time instance of tool breakage. The data was applied to a SCBC, a Kohonen's feature map, and an ART2 network to detect tool breakage. The performances of these neural networks were evaluated by the number of false alarms and undetected tool breakages.

Results: The SCBC detected tool breakage correctly without any false alarms in six out of six tests. The Kohonen's feature map detected tool breakage correctly in five out of six tests. The ART2 network failed to detect tool breakage in three tests and produced three false alarms.

75

Comments: The proposed SCBC seems to be a good alternative classifier to a Kohonen's feature or ART2 network for tool breakage detection. However, the overall approach is too elaborate for the simple task of tool breakage detection. Tool breakage is a phenomena which is easy to distinguish from the normal cutting process. Methods employing a single sensor, such as acoustic emission, are reported to give reliable results (Emel and Kannatey-Asibu 1988, Liang and Dornfeld 1989).

3.5 Conclusions

This chapter reviewed various neural network applications in tool wear monitoring in turning processes. Compared to analytical models or observers, neural networks offer more practical means for tool wear estimation. Neural networks do not depend on analytical models of tool wear. Instead, they learn the relationship between process variables and tool wear from the experimental data. This obviates the problem of understanding the complex physics of the cutting process. However, there will always be scope for improvement of neural network-based methods when the physics of the cutting process is well understood.

From the review of tool wear monitoring methods in this chapter, one can make an observation: when applying neural networks to tool wear monitoring problems, good sensor data representation schemes are as important as effective neural network models. Here the sensor data representation refers to sensor data reduction or feature extraction to create neural network input patterns. Assuming that good sensor data representation schemes are available, the following sections discuss some of the issues related to neural network applications.

76

The tool wear monitoring problems discussed in this chapter can be classified into the following three categories:

Category 1: In this category of tool wear monitoring problems, tools are classified into one of the two decision classes (*good* or *bad*).

Category 2: In this category of problems, tools are classified into relatively finer classes then in category 1 problems. Usually tools are classified into one of three or more classes (*fresh, medium,* or *worn* tools).

Category 3: In this category of problems, the quantitative value of tool wear is estimated at any given time while the cutting is in progress (*continuous* tool wear estimation).

The following sections discuss the application of neural networks for these three categories of tool wear monitoring problems.

3.5.1 Neural Networks for Category 1 Problems

In this category of tool wear monitoring problems, a tool is classified into one of two classes. For example, a tool is classified into fresh or worn based on a tool wear threshold. Or, a tool is classified into normal or broken, depending on the state of the tool. As reviewed earlier in this chapter, Rangwala and Dornfeld (1990), Dornfeld and DeVries (1990), Emel (1991), and Chryssolouris et al. (1992) addressed the tool wear monitoring problem of this category using multilayer neural networks. Burke (1993) used adaptive resonance theory networks to address the same problem. Between these two neural network models, multilayer neural networks are more suitable for addressing this problem. Multilayer neural networks are more stable and will give more accurate and consistent results than

77

adaptive resonance theory neural networks. Though adaptive resonance theory neural networks provide the advantage of shorter training time and unsupervised learning, they are relatively less stable in learning and will give less accurate results in the presence of noise in inputs. The major strength of the adaptive resonance theory networks is their property of adaptive classifications. They are designed for classifying inputs into adaptively created classes (Carpenter and Grossberg 1987, 1991). They are unsuitable for the present application because in this application, inputs are to be classified into one of the two predetermined classes. If the adaptive resonance theory neural networks are used in this application, their main strength is not utilized.

The researchers who used multilayer neural networks for the present application, trained the networks with a standard backpropagation (BP) algorithm (Rumelhart and McClelland 1986). Neural networks in these application could also be trained with the Bayesian backpropagation (BBP) algorithm (Nedeljković 1993) to improve the accuracy of classification. The BP algorithm minimizes the mean-square error criterion. In contrast, the BBP algorithm minimizes the probability of misclassification of inputs. The BP algorithm does not necessarily minimize the probability of classification error while the BBP does. Therefore, the classification accuracy achieved by the BP algorithm is worse or at most equal to that achieved by the BBP algorithm.

Jammu et al. (1993) addressed this problem in a different way. They proposed an unsupervised classifier called a Single-Category-Based Classifier for classifying the tools into normal or broken classes. It is, again, the author's opinion that equally good results can be obtained by more elegant unsupervised Competitive

Learning algorithms proposed by Xu et al. (1993), Kong and Kosko (1991), and Ahalt et al. (1990).

3.5.2 Neural Networks for Category 2 Problems

In this category of tool wear monitoring problems, a tool is classified into more than two classes. For example a tool can be classified as fresh, intermediate, or worn. Kamarthi et al. (1991) and Leem and Dreyfus (1992) investigated the application of Kohonen's feature maps to address this category of tool wear monitoring problems. In this approach, a Kohonen's feature map is trained using the unsupervised feature map algorithm (Kohonen 1990) in the first step. In the next step, the feature map is fine tuned by a supervised training procedure using a set of labeled input patterns. This labeled input set actually determines the numbers classes into which the inputs are classified during the network operational phase. If the labeled input set contains input patterns of N classes, then the output nodes in a Kohonen's feature map are partitioned into M ($\leq N$) classes during the supervised training. Thus, during the network operational phase, the input patterns are classified into M classes. For example, in the investigations of Kamarthi et al. (1991), and Leem and Dreyfus (1992), $N = M = 3$ and the corresponding three input pattern classes were fresh, medium, and worn tools. In this approach, the larger the size of M, the finer the classification of the input patterns.

Alternatively, adaptive resonance theory networks can be used to address this class of tool wear monitoring problems. When an adaptive resonance theory network is trained with input patterns, it will adaptively classify the inputs into their natural classes. The fineness or coarseness of the classification is controlled by the vigilance limit of the network. Once the initial classification is over, the

79

output classes created by the network are labeled using a set of labeled input patterns. The advantage of this network is that it requires much less training time compared with that required by a Kohonen's feature map. However, as discussed in a previous section, this network is less stable (in retaining learned patterns) and likely to provide less accurate classification results than a Kohonen's feature map.

It seems that the Rival Penalized Competitive Learning (RPCL) algorithm (Xu et al. 1993) would be a better procedure to address the tool wear monitoring problems of this class. The unsupervised competitive learning part of the RPCL algorithm is similar to that employed in a Kohonen's feature map. In the RPCL algorithm, the weights of the winning node and its rival node (the second winning node) are modified to adapt to an input pattern. In a Kohonen's feature map, the winning node and all its neighboring nodes, determined by a neighborhood function, are modified to adapt to an input pattern. Therefore, the RPCL algorithm requires considerably less training time compared to a Kohonen's feature map.

In a Kohonen's feature map, the weights of neighbors of the winning node are modified in an attempt to form a topologically ordered feature map. Though the Kohonen's feature map training algorithm is designed to form such topological ordering of its nodes, there is no existing theoretical proof that guarantees such a topological ordering in a feature map whose dimension is smaller than the dimension of input space. In the tool wear monitoring applications where an input pattern dimension is usually greater than three and the dimension of a Kohonen's feature is one or two, the topological ordering of the feature map nodes is most unlikely to happen. The author's experience with Kohonen's feature maps confirms this result. Since the training time spent in creating topological ordering of nodes is not going to give any fruitful results in the present application, it appears

80

that an RPCL algorithm is a better choice than a Kohonen's feature map for this application.

3.5.3 Neural Networks for Category 3 Problems

In this category of tool wear monitoring problems the quantitative value of the tool wear is estimated at any given time. The tool wear monitoring methods investigated by Wang and Dornfeld (1992) and Elanayar et al. (1990) provide continuous tool wear estimates. Whereas the other methods reviewed in this chapter classify tools into two or three classes depending on their wear levels, they do not provide continuous tool wear estimates. To predict tool failure and to implement adaptive control strategies, a method which provides continuous tool wear estimates is more useful than a method which classifies tool wear into two or three classes.

The tool wear monitoring problems of this category are much more difficult to address than the problems of the previous two categories. To address this category of problems, the complex relationship between inputs and tool wear must be approximated. Multilayer neural networks are the most suitable neural networks for this purpose. The results obtained by Wang and Dornfeld (1992) and Elanayar et al. (1990) proves the effectiveness of multilayer neural network for the continuous tool wear estimation. However, implementing multilayer neural networks poses some problems. No elegant way has yet been found to determine the structure of a multilayer neural network for a given problem. Multilayer neural networks need an enormous amount of supervised training time and expensive supervised training data. Though in most of the practical cases these neural networks give good results, there is no theoretical guarantee that the network mean-square-error

81

criteria will converge to a global minima.

Therefore, more emphasis is necessary to investigate neural networks that converge to global minima, require shorter training time and learn primarily from unsupervised training data. Kumara et al. (1993) have investigating the application of Radial Basis Function (RBF) neural networks for on-line flank wear estimation in turning. The preliminary results indicate that RBF neural networks have good potential for continuous flank estimation and at the same time addresses the difficulties posed in implementing multilayer neural networks.

Though this chapter discussed the issues related to the neural network application for tool wear monitoring in turning processes, most of these issues are relevant to milling processes as well. Some of the most important neural network applications in milling were investigated by Noori-Khajavi and Komanduri (1993), Tansel and McLaughlin (1993), and Tansel et al. (1993).

The previous chapter and this chapter together have reviewed the important tool wear monitoring methods in turning processes. The next chapter describes the available sensors for monitoring tool wear in turning processes. This chapter also presents the criteria for the selection of sensors for the on-line tool wear estimation method developed in this thesis.

Chapter 4

SELECTION OF SENSORS

4.1 Introduction

This chapter briefly reviews the potential sensors for monitoring process output variables such as force, vibration, and temperature useful for tool wear estimation in turning processes. This chapter then discusses the selection of sensors based on their applicability, robustness, and reliability in a real world situation.

The methodology in this thesis considers multiple sensors for flank wear estimation because more reliable and consistent tool wear estimation is possible through sensor data fusion. Sensor data fusion refers to the combination of data from multiple, possibly dissimilar, sensors into a single result (Geirsson 1990). The potential advantages of integrating information from multiple sensors (Chryssolouris and Domroese 1988) are as follows:

1. The quality of decisions will be better if the information comes from multiple sensors rather than from a single sensor because using information from multiple sensors means using more information simultaneously. Using information from several sensors is analogous to taking several samples from a random population. This obviously narrows the mean confidence interval and reduces uncertainty in a decision.

2. The reliability of information from different sensors might change relative to one another depending on the process input parameters. If this change can be correlated, a sensor integration scheme can be designed to selectively weight

83

the appropriate sensors at different input conditions to make the decisions more reliable for a wide range of cutting conditions.

3. The decisions based on multiple sensors will be more fault-tolerant because even if some of the sensors fail, the other sensors will compensate for the lack of information from the defective sensors.

Though it can be argued that dissimilar information from multiple sensors may not lead to higher information value, multiple sensor information can provide redundancy which is essential for critical systems.

In the light of these benefits, the combination of force, vibration, and acoustic emission sensor data is proposed in this research for estimating tool wear states. The rational for choosing these sensors is discussed in this chapter.

Section 4.2 sets the criteria for screening the sensors useful for the present application from available sensors. Section 4.3 reviews the potential sensors that can be used in turning processes for monitoring various process outputs. Section 4.4 discusses the selection of sensors.

4.2 Criteria for Sensor Selection

Most of the methods for tool wear estimation described in literature use one or other combination of cutting force, temperature, vibration, and acoustic emission measurements. To a lesser extent, other process variables such as power/motor current, sound, and workpiece surface roughness have also been used for the same purpose. All these possible sensors for the present application are evaluated using the following criteria:

84

1. The process variable being monitored by a sensor should have a good correlation with tool wear.

2. A sensor should be able to give reliable and consistent measurements of the process variable being measured.

3. A sensor should be able to provide a signal of high signal to noise ratio. Even if the signal carries noise, it should be possible to filter out the noise from the signal.

4. A sensor should have a short response delay. This is necessary to implement on-line tool wear estimation and control.

5. It should be possible to physically locate a sensor on the tooling set up without obstructing the cutting operation. The sensor should be robust enough to withstand the impact of chips, spillage of coolants, and temperature fluctuations.

6. A sensor should have a long life and be cost effective.

The following section reviews the potential state-of-the-art sensors for tool wear estimation in turning processes and evaluates them according to the above criteria.

4.3 Potential Sensors for Tool Wear Estimation

Micheletti et al. (1976), Tlusty and Andrews (1983), and Dan and Mathew (1990), reviewed sensors and techniques for tool wear prediction. These reviews discussed the sensors which have potential for tool wear prediction. Some of them are listed below:

1. Cutting Force

2. Vibration

3. Acoustic emission

4. Temperature

5. Roughness of machined surface

6. Workpiece dimensions

7. Distance between tool post and the workpiece

8. Power/Spindle motor current

4.3.1 Cutting Force

Tool wear monitoring using cutting force measurements has been widely investigated. It is commonly agreed that there is a correlation between machining forces and tool wear (Micheletti et al. 1968, Taraman et al. 1974). It has been observed by several researchers that the feed force is more sensitive to wear than force components in cutting and radial directions (Takeyama et al. 1967, Uehara et al. 1979, Tlusty and Andrews 1983, Schaffer 1983, Jetly 1984). However most researchers are of the opinion that measurement of two or more force components is necessary for good correlation with tool wear (Martin et al. 1986).

Lindstrom and Lindberg (1983) developed a force transducer to measure the dynamic forces from the chip formation process. This sensor uses a piezo-electric element which, when dynamically compressed, produces an electrical output proportional to the dynamic forces transmitted through it. Extensive cutting tests showed that dynamic cutting forces can be used to indicate flank wear. These tests also showed that force measurements are more effective than vibration acceleration measurements in detecting tool wear.

86

Dan and Mathew (1990) reported that the power and the auto-correlation coefficient of force signals reflect tool wear. The amplitude of the high frequency component of feed force keeps increasing during initial and intermediate tool wear stages and subsequently starts to decrease during the final wear stage and tool breakage.

Tlusty and Andrews (1983) and Lan and Dornfeld (1984) reported that both cutting and feed forces are sensitive to tool fracture. At the instant of tool breakage these force components drop suddenly and increase after a few milliseconds. This change in force components was used to detect tool breakage.

Although force components are sensitive to tool wear, the mechanisms causing tool wear and failure have complex relationships with the force components. The relationship between force components and tool wear is influenced by the tool geometry, the properties of tool and workpiece materials and by operating conditions (mainly feed and speed). However, the sudden drop in cutting and feed forces are observed to be reliable indicators of tool breakage.

4.3.2 Vibration

When the cutting edge of a tool is sharp, vibrations are generated by the formation of chips as metal is being removed from the surface of the workpiece by the cutting tool. As the cutting edge wears, additional vibrations are generated by the sliding contact between the workpiece and wear-land along the flank of the tool.

Several methods have been developed to relate vibration characteristics with tool wear. Pandit and Kashou (1982) used a data dependent system to isolate a model of vibrations sensitive to tool wear. This model gave the power contribution

87

of vibration frequencies sensitive to flank wear. The band selective power decreased with increasing flank wear until it reached a minimum at certain flank wear level. Afterwards, the power increased until the tool failure occurred. Thus the criterion of minimum power contribution was used as an indicator of critical tool wear.

Weller et al. (1969) developed a worn flank detector. In this method, vibration energy is divided into two frequency components: one in a low-frequency range 0 to 4 kHz, the other in a high-frequency range 4 to 8 kHz. It was found that the low-frequency component band energy for a sharp cutting edge and a worn cutting edge is about the same, while the high-frequency band energy for a worn cutting edge is considerably higher than for a sharp one. The total amount of high-frequency vibration energy increases as the width of the wear-land increases. This phenomenon was observed consistently, even if the cutting parameters such as feed, speed, and depth of cut are varied within wide ranges. Thus, Weller et al. (1969) used the ratio of high-frequency-vibration energy to low-frequency-vibration energy as an indicator to detect worn cutting tools.

Taglia et al. (1976) investigated the relationship between flank wear and power spectrum of a cutting tool vibration. The results showed that the power spectra, cross power spectra and coherence function of signals in the low frequency band (0 to 2.5 kHz) undergo some noticeable and regular variations with wear; such variations are less sensitive and almost uncertain in the wide frequency band (0 to 12.5 kHz).

In the frequency range up to 2.5 kHz, a very small percentage of the total power of acceleration signal was found to be sensitive to the tool wear. It increased up to seven times when the flank wear increased from a very small value to about 1.3–1.5 mm. A further increase of flank wear beyond this point caused the power

88

contained in this frequency range to fall rapidly. The cross power spectra between two successive data records in the frequency range of 0 to 2.5 kHz showed a trend similar to that of the power spectra with increased wear. The mean coherence between two successive data records in the frequency range of 0 to 2.5 kHz decreased steadily with wear and then increased until failure.

Jiang et al. (1987) proposed a frequency band-energy method for in-process flank wear monitoring. It was found that the cutting vibration signal is very sensitive to wear states of cutting tools. The frequency composition of the signal energy varies regularly with the development of tool wear. The proposed strategy involved monitoring the energy of certain frequency bands that are sensitive to level of flank wear over the tool life.

Sokolowski et al. (1992) studied the effect of variation in tool shank-holder configuration and process operating conditions on the spectrum of the cutting tool vibration signals. It was found that tool shank-holder configuration changes result in a shift in the frequency band related to the flank wear. The natural frequencies of the overhang portion of the tool shank are the dominant frequencies sensitive to flank wear. The amplitudes of the dominant frequencies monotonically increase with increasing flank wear. The largest increase in the amplitude of vibration velocity was observed in the feed direction.

Currently, there is a growing interest in using vibration signals for on-line flank wear monitoring. Vibration transducers are easy to mount on the tool shank or tool holder and, thus, easy to measure without interfering with the cutting process. The frequency ranges of the vibration signals (usually between 0 to 20 kHz) makes the instrumentation for data acquisition and analysis cost-effective.

4.3.3 Acoustic Emission

Acoustic emission (AE) is the transient elastic energy spontaneously released in materials undergoing deformation, fracture, or both (Kannatey-Asibu and Dornfeld 1982). AE is related to the grain size, dislocation density, and distribution of second-phase particles in crystalline materials. In turning processes, AE generated at the following sources:

1. Elastic and plastic deformation of the workpiece.

2. Fracture of both the workpiece and the cutting tool.

3. Friction between the chip and the rake face of the tool.

4. Friction between the workpiece and the flank face the tool.

AE signals are usually picked up by a piezoelectric sensor, amplified by an amplifier, and band-pass filtered in the frequency range of 100 kHz to 1 MHz. The normal AE parameters used for flank wear estimation and tool failure detection are count rate, AE-mode, RMS, amplitude of dominant frequencies, or power of dominant frequency bands. The count rate of AE events for a fixed voltage threshold is monitored by a digital electronic counter.

Emel and Kannatey-Asibu (1988) found that progressive flank wear affects the AE power in the frequency range 400–700 kHz and catastrophic tool failure boosts the AE power to the frequency range 100 kHz to 1 MHz. Kannatey-Asibu and Dornfeld (1982) found that skewness and kurtosis of the RMS data, based on an assumed Beta distribution, have good sensitivity to flank wear.

Inasaki and Yonetsu (1981) observed that in the case of carbon steels, the amplitudes of AE signals near frequencies of 120, 170, and 210 kHz. These frequencies increase with increasing flank wear height and cutting speed but are hardly affected by the changes in feed rate and depth of cut. Further, a stepwise

90

increase in the AE signal energy was observed with the increasing contact area after tool fracture.

Iwata and Moriwaki (1977) found that AE measurements are hardly affected by ambient vibrations and noise, and the frequency spectrum of AE signals in the frequency range of 100 to 250 kHz increases with flank wear. However, this increase in the frequency spectrum reaches a saturation at the advanced flank wear stage. They also found that the total count of AE events has a good correlation with flank wear.

Inasaki et al. (1987) found that the AE Mode, defined as the mode of the probability density function of the AE signal, is very sensitive to tool chipping. Lan and Dornfeld (1984) observed that significant AE bursts are generated at the instant of tool breakage and that the RMS level of the bursts depend on the fracture area; the larger the tool fracture area the stronger the AE energy. Diei and Dornfeld (1987) proposed a quantitative model relating the peak value of an RMS AE signal to both the fractured area and the resultant cutting force at tool fracture.

From the study of various experimental results on tool wear and and tool failure detection using AE signals, it appears that AE spectral power, AE Mode, RMS value, count rate, and total count of AE events are consistently correlated to tool wear and tool fracture. This technique has a quick response time and poses no problem in mounting the sensor on the tooling set up without hindering the cutting process.

4.3.4 Temperature

During the cutting operation, heat is generated within the primary and the

91

secondary shear zones due plastic deformation of material. Along the tool-chip interface, heat is generated. If the cutting edge is not perfectly sharp, an additional frictional heat is generated due to ploughing. The methods to monitor temperature at tool-chip interface which may be practical for on-line measurement are grouped into the following categories:

1. Tool-work thermocouple method

2. Embedded or remote thermocouple method

3. Radiation method

4. Thermo-paint method

In the tool-work thermocouple method both the tool and workpiece are electrically isolated from the machine tool structure. This technique was used for tool failure detection (Colwell 1975, Colwell and Mazur 1979, Boothroyd et al. 1967), controlling the cutting process (Billett 1968, Shillam 1971), and tool wear estimation (Turkovich and Kramer 1986). It was found that the reliability of the tool-work thermocouple method is affected by the material properties at the junction and by the noisy thermal voltage signal which is sensitive to cutting conditions.

In embedded or remote thermocouple methods, one or more thermocouples are located on the cutting tool remote from the cutting edge (Groover and Kane 1971). Usually thermocouples are located at the seat of the tool insert. This approach was used for predicting tool wear from temperature measurements (Levy et al. 1976, Lipman et al. 1967, Solaja and Vukelja 1973). Using this approach, it is difficult to measure temperature at the cutting edge because the thermocouples are embedded within the cutting tool a distance away from the cutting edge. The temperature measurements in this method tend to have errors primarily because

92

the temperature gradients near the edge are steep, and the heat conduction characteristics in the cutting tool are altered by the presence of holes in which the thermocouples are embedded. This method can be used to estimate the average interface temperature only after the system has reached steady state conditions (Chow 1984). In a slight variation to this approach, thermocouples are placed touching the back of the cutting inserts. In this approach the temperature profile at the tip of the tool is predicted using numerical methods (Groover et al. 1977, Raman 1989).

Radiation pyrometers and infrared detectors are used for determining temperature distributions on the surfaces of the tool and workpiece and at the tool-chip interface. In this technique, the total radiant energy from the cutting area is measured and is calibrated to give temperatures directly (Boothroyd 1963). Although the radiation method for temperature measurement is independent of tool and workpiece materials and fast in response, the application of this technique to on-line temperature sensing is impaired by the need for direct access to the cutting zone (Schaffer 1983). This method is highly sensitive to the surface roughness and it only detects the temperatures on the surface (Chow 1984).

The thermo-paint techniques utilize thermosensitive paints which change colors with temperature. The thermosensitive paints exhibit different colors at different ranges of temperature. The thermocolors are painted on the tool surface, and the colors of the surfaces are examined after cutting. Knowing the transition temperatures for each color, the isothermals are inferred (Jetly 1984, Koch 1970, Lister and Barrow 1986, Micheletti et al. 1976). This method tends to be inaccurate due to the low sensitivity of the thermocolors. It is difficult to detect small changes in temperature. There is a relatively long time lag for the color to change.

93

It shows temperatures only on the surface of the tool. The applicability of this technique is restricted to accessible surfaces.

4.3.5 Roughness of Machined Surface

The roughness of a machined surface is influenced by the nose radius, by the sharpness of a cutting tool, and by cutting speed. The nose radius creates a regular pattern of waves on the machined surface. The wavelength of this pattern is usually determined by the nose radius, depth of cut, and feed rate. This wave pattern is almost flat if the feed rate is low, the speed is high (cutting speed affects the surface finish through the shear strain rate), and the depth of cut is small. When this regular pattern is distorted, it indicates that the tool is in a poor state.

Spirgeon and Slater (1974) used a fiber-optics transducer to monitor the surface roughness during a finish turning. This transducer works on the principle that the reflectivity of light from a newly turned surface varies inversely proportional to the roughness of the surface. Takeyama et al. (1976) used a pair of reflection systems to monitor surface roughness. This method was found to be effective in detecting changes in cutting edges due to chipping or wear. Whitehouse et al. (1983) investigated an optical diffraction technique to assess the surface finish of diamond turned parts. The surface finish reading obtained from this method was used to predict tool wear.

Although the roughness of a machined surface helps to detect the changes in a cutting tool edge, it does not indicate what mode of tool wear is affecting the roughness of the machined surface.

94

4.3.6 Workpiece Dimensions

Workpiece diameter will change as the flank wear progresses. This is more noticeable in finish-cut turning. If expressions relating tool wear and diameter change are developed, tool wear can be estimated by measuring workpiece diameter. Both contacting and non-contacting sensors have been developed for measuring workpiece diameter on-line. Colding and Novak (1981) used a laser beam for sensing a workpiece diameter. El Gomayel and Bregger (1986) used an electromagnetic sensing probe to monitor the change in workpiece diameter.

Monitoring workpiece diameter is useful in tool wear detection and taking corrective action to maintain the workpiece dimension. However, this method may fail to distinguish between the tool wear modes. This method is affected by workpiece misalignment, deflection, and thermal expansion.

4.3.7 Distance Between Tool Post and the Workpiece

When cutting is in progress, the distance between the tool holder and the workpiece decreases as the tool wear increases. This distance can be measured directly by techniques such as electronic micrometers (Takeyama et al. 1967), reflected ultrasonic waves, pneumatic gauges (Bath and Sharp 1968), and inductive proximity devices (Suzuki and Weinamann 1985). However, the sensitivity of these techniques may possibly be influenced by workpiece surface temperature, surface finish, cutting fluids, and workpiece diameter.

4.3.8 Power/Spindle Motor Current

In turning processes, power/spindle motor current is related to the motor shaft torque or force component in cutting direction. When the tool is sharp, less

95

force is required in cutting, feed, and radial directions. As the tool develops flank wear, forces in all these directions increase to varying degrees. This variation in forces, particularly in cutting direction, increases the power required to drive the shaft. This variation is used for on-line prediction of tool wear (Liao 1986).

The spindle current is usually measured through a current transformer. The measured signal was found to drop instantaneously and soon recover to the pre-drop level when tool breakage occurs. This behavior is used to detect tool breakage. This technique is relatively simple and the current measurement system does not cause any hindrance to the cutting operation. It was found that power/spindle current is less sensitive for tool wear sensing compared to force and vibration measurements (Martin et al. 1986). However, it was also found that power/spindle current is a reliable indicator of tool breakage in medium and heavy cuts (Novak and Ossbahr 1986).

4.4 Selection of Sensors for Tool Wear Estimation

From a careful study, Table 4.1 is created to rank the process variables sensitive to flank wear. The table assigns 0 (poor), 1 (good), or 2 (very good) to each criterion for each sensor. The criterion 1 through 6 (denoted by crt.1, crt.2, $\cdots$ crt.6 in the table) corresponds to the six criteria listed in section 4.2. From this table it can be concluded that cutting force, vibration, and AE sensors are the three best choices for on-line flank wear estimation. The reasons can be further be explained as follows:

1. Cutting force, vibration, and acoustic emission are sensitive to tool wear for a wide range of operating conditions.

2. Cutting force, vibration, and acoustic emission can reliably and accurately

96

Table 4.1: Ranking of sensors for flank wear estimation

Sensor	Crt.1	Crt.2	Crt.3	Crt.4	Crt.5	Crt.6	Total
Force	2	2	1	2	2	2	11
Vibration	2	2	2	2	2	2	12
AE	1	2	1	2	2	2	10
Temperature	2	1	1	0	0	1	5
Roughness	1	1	1	1	0	1	5
Diameter	1	0	1	1	0	1	4
Tool-Work Dist.	1	0	1	1	0	1	4
Motor Current	1	0	1	2	2	2	8

be measured by a dynamometer, an accelerometer, and an acoustic emission sensor respectively.

3. The signal to noise ratio of force, vibration, and acoustic emission signals can be improved by filtering out signal noise outside the frequency band of interest.

4. These sensors are relatively robust and easy to mount or incorporate in a tooling set up without obstructing the cutting process.

5. These sensors have quick response time and are suitable for on-line tool wear estimation.

6. The increased sophistication of these sensors and related instruments make them more practical and cost-effective in industrial applications.

Force and vibration signals can be measured in three machine-tool coordinate directions: cutting, feed, and radial. Since machining force in all three coordinate directions is sensitive to tool wear in varying degrees, force in all three directions are measured for the current research.

Vibration signals in feed direction are more sensitive to flank wear than vibration signals in radial and cutting directions. The studies of Rao (1986) and

97

Sokolowski et al. (1992) indicate that vibration in radial and cutting directions behave similarly with progressing flank wear. Therefore, vibration signals in the cutting and feed directions are measured in this research.

From a multisensor-data-fusion point of view many sensors may be useful. However, it is necessary to verify if the three forces signals and two vibration signals contain independent sets of information. This verification is done using the principal component analysis. The details of this analysis is discussed in chapter 5. Table 4.2 gives the list of process variables used in the tool wear estimation method developed in this research. This tables also gives the sensors used to measure these process variables.

Table 4.2: List of process variables measured.

Process Variable	Sensor
Force in cutting direction	*Dynamometer*
Force in feed direction	*Dynamometer*
Force in radial direction	*Dynamometer*
Vibration in cutting direction	*Accelerometer*
Vibration in feed direction	*Accelerometer*
Acoustic emission	*Wideband AE Sensor*

4.5 Conclusions

This chapter has reviewed various sensors that can potentially be used for measuring process outputs for tool wear estimation in turning processes. After evaluating these sensors according to a set of criteria, three most promising sensors have been selected: force, vibration, and acoustic emission sensors. In the

98

following chapters, sensor data collection, processing, and fusion will be limited to these three types of sensors. For force and vibration sensors, multiples of the same sensor are used in different machine-tool coordinate directions. A three axis dynamometer is used to measure forces in three directions: cutting, feed, and radial. Similarly, two accelerometers—one in cutting direction and the other in feed direction—are employed to measure vibrations.

METHODOLOGY AND EXPERIMENTATION

5.1 Introduction

This chapter describes the methodology for flank wear estimation and the necessary experimentation to verify the methodology.

The detection and measurement of tool wear has attracted much research for more than two decades. A variety of methods have been explored. Most of them can be classified as methods based on analytical models, observers, empirical models, and neural networks. These methods were reviewed in Chapters 2 and 3. The following are some of the limitations with these methods:

1. Metal cutting is inherently a very complex process to model for flank wear estimation. The current understanding of how the different process parameters and material properties are related to tool wear is limited. These factors make it is difficult to develop a correct analytical model structure for flank wear estimation.

2. Even if a reasonably correct analytical model is available, determining the model parameters for different tool and workpiece materials at different ranges of operating conditions is time consuming and often expensive.

3. Currently, the existing observers are based on analytical models. Consequently, they inherit the difficulties associated with the analytical models. The adaptive observers, which are data dependent, need off-line measurements

100

to minimize the observer error.

4. Empirical models consist of regression models that relate flank wear levels to the process parameters considered to have an influence on the flank wear. The empirical models fail to give accurate flank wear estimates over a considerable range of operating conditions because they are too simple to model a complex process such as metal cutting.

5. Recently, several neural network-based methods have been developed for flank wear monitoring but most of them were applied to classify flank wear levels into two or three classes such as fresh, medium, and worn tools. Only a few of them focus on continuous flank wear estimation.

In order to overcome some of the above limitations in the existing flank wear estimation methods, a new approach is taken in the present work. Figure 5.1 gives the important differences between the existing methods and the proposed method.

Figure 5.2 shows the proposed flank wear estimation method as a part of the process monitoring, diagnosis, and control cycle. As shown in the figure, this method has four important steps: sensor data acquisition, sensor data preprocessing, sensor data representation, and flank wear estimation. Figure 5.3 shows the details of the proposed flank wear estimation method. The following sections explain these steps.

5.2 Sensor Data Acquisition (Step 1)

The review of the literature suggests that a more reliable and consistent tool wear monitoring is possible through sensor data fusion. Sensor data fusion refers to the combination of data from multiple, possibly dissimilar, sensors into a single result (Geirsson 1990). The selection of correct sensors is the first step in

101

Existing Methods	Current Method
Most of the existing methods depend on analytical models to relate tool wear to process variables	**Current method uses neural networks to map tool wear to processes variables through training data**
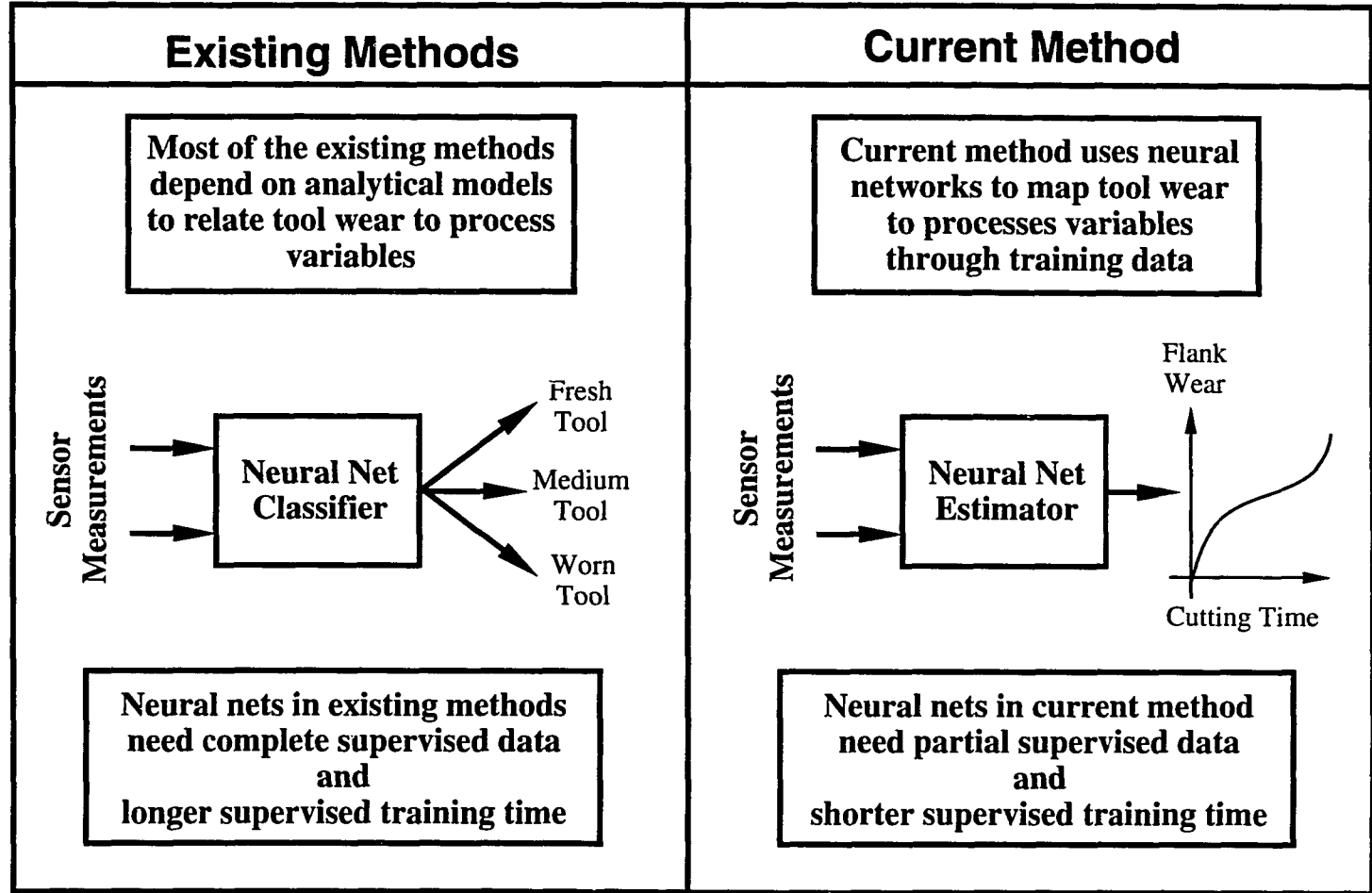
Figure 5.1: Comparison of the existing methods with the current method
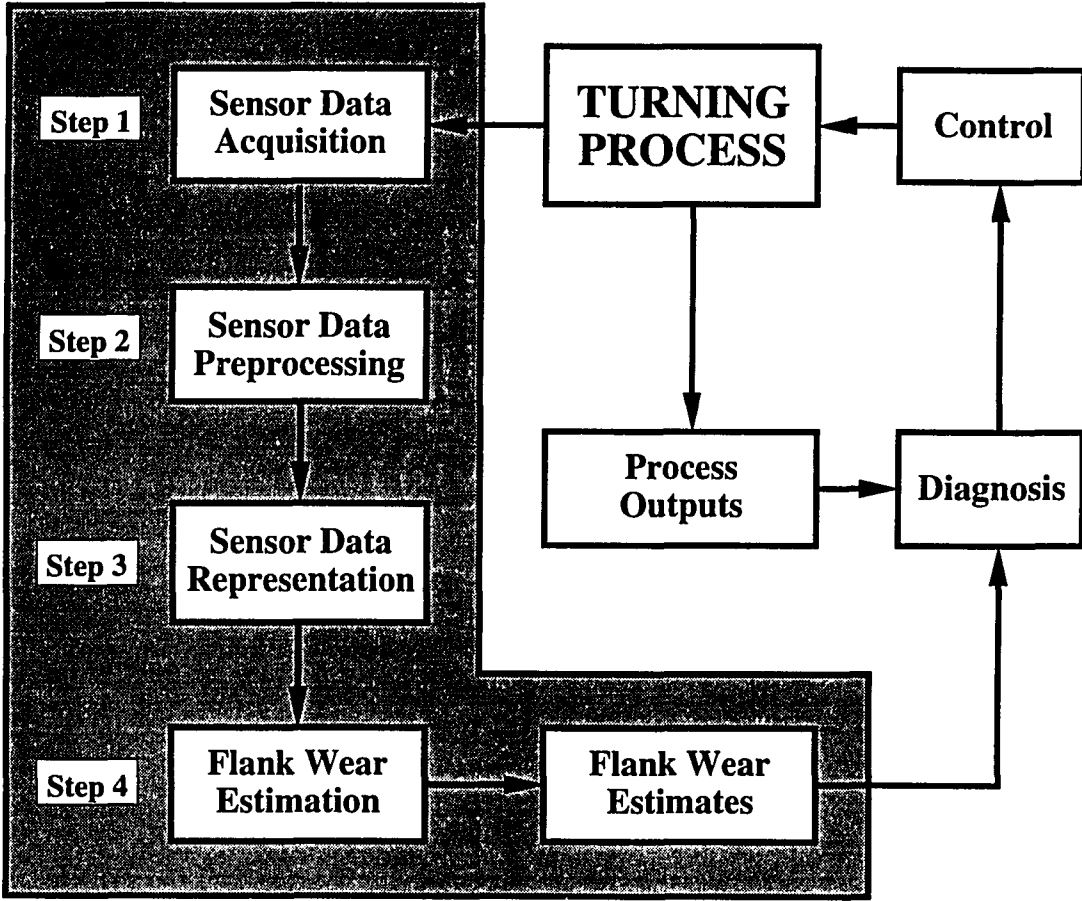
Steps in Flank Wear Estimation

Figure 5.2: Tool wear estimation in monitoring-diagnosis-control cycle

103
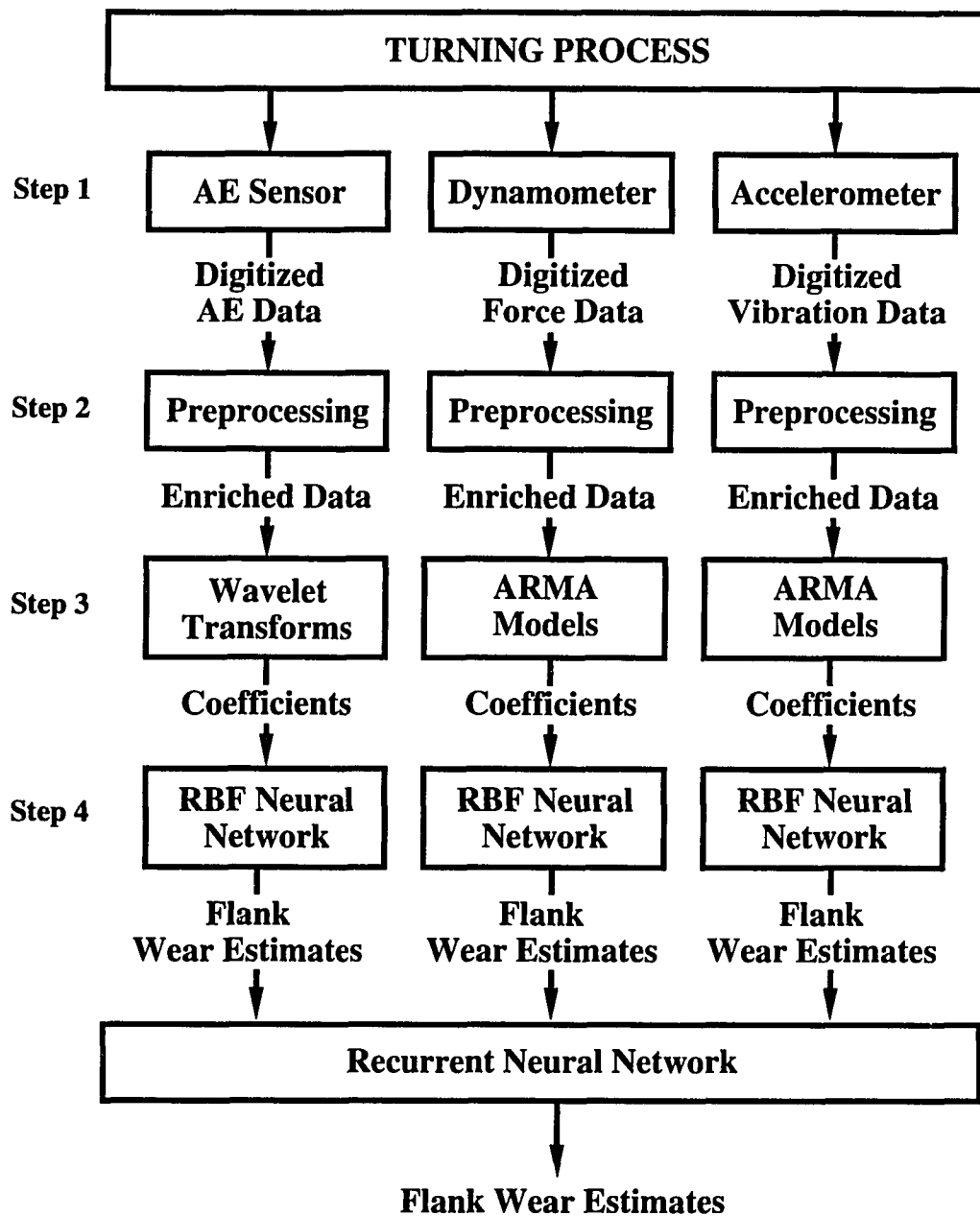
Figure 5.3: Methodology for flank wear estimation

104

implementing a successful sensor integration scheme. Therefore in this work, the following five signals are measured for flank wear estimation:

1. Force in cutting direction.

2. Force in feed direction.

3. Force in radial direction.

4. Vibration in cutting direction.

5. Vibration in feed direction.

6. Acoustic emission.

The sophisticated transducers and instrumentation available for measuring force, vibration, and AE signals make these variables practical and cost-effective for flank wear estimation. These sensors also are relatively robust, easy to mount, and have quick response time which makes them attractive for real-time flank wear estimation. The issues related to the selection of sensors were discussed in Chapter 4.

In the list of six sensors, the three force sensors are of one type and two vibration sensors are of another type. There is a possibility that linear dependencies exist between the same type of sensors. There is little possibility that such linear dependencies exist between different types of sensors because frequency bands of interest are different for each type of sensor. Typically, the frequency band of interest in AE signals are in the range of 100 kHz to 1 MHz, while the frequency band of interest in force signals is in the range of 5 Hz to 1kHz, and in vibration signals is in the range of 1 kHz to 15 kHz.

It is necessary to investigate any linear dependencies among force and vibration sensors and use only those sensors which are linearly independent. In this research, principal component analysis is conducted on the three time series

105

of force signals and on the two time series of vibration signals to determine the possible linear dependencies among them.

5.2.1 Detection of Linearly Dependent Sensors

Let x be the multivariate time series of either three force signals or two vibration signals. Let k be the dimension of the multivariate time series. In the case of force signals $k = 3$, and in the case of vibration signals $k = 2$. The components of the multivariate time series of force signals are:

x_1 time series of force in cutting direction,

x_2 time series of force in feed direction, and

x_3 time series of force in radial direction.

The components of the multivariate time series of vibration signals are:

x_1 time series of vibration signals in cutting direction, and

x_2 time series of vibration signals in feed direction.

The eigenvalue-eigenvector analysis (Box and Tiao 1977, Tiao and Box 1981) is performed on the multivariate time series of force and vibration signals to determine the linear dependency among the components of multivariate time series.

The theory behind this method comes from the principal component analysis. The central idea of the principal component analysis is to reduce the dimensionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set (Jolliffe 1986). This reduction is achieved by transforming the original variables to a new set of uncorrelated variables called the principal components. These principal components are ordered so that the first few retain most of the variation present in all of the original variables. More generally, if a set of k variables has substantial

106

correlations among them, then the first few principal components will account for most of the variation in the original variables. Conversely, the last few principal components identify directions in which there is very little variation, revealing near constant linear relationships between the original variables. Computation of the principal components reduces to the solution of an eigenvalue-eigenvector problem for a positive-semidefinite symmetric matrix (Jolliffe 1986, Dunteman 1989).

The eigenvalue-eigenvector analysis, for the present problem, consists of the following steps:

1. Estimate $\Gamma_{\mathbf{x}}(0)$, the lag zero cross covariance matrix of the multivariate time series $\mathbf{x}$.

2. Compute the eigenvalues of the estimated covariance matrix $\Gamma_{\mathbf{x}}(0)$

3. Fit a multivariate AR model of a relatively large order p to $\mathbf{x}$.

4. Compute the vector residual series $\mathbf{e}$ of the AR model determined in step 2.

5. Estimate $\sum$, the covariance matrix of the vector residual series $\mathbf{e}$.

6. Compute the eigenvalues of the estimated covariance matrix $\sum$.

If the estimated covariance matrix $\Gamma_{\mathbf{x}}(0)$ has m eigenvalues equal or close to zero, then the contemporaneous components of $\mathbf{x}$ have m linear relationships among them. These relationships are of the form

$$\mathbf{c}_j^T \mathbf{x} = 0, \qquad \text{for } j = 1, 2, \ldots, m, \tag{5.1}$$

where $\mathbf{c}_j^T = [c_{j1}, \ldots, c_{jk}]$ is a vector of constants. Such relationships occur when one or more time series are computable from contemporaneous values of the others.

If the estimated covariance matrix $\sum$ has m eigenvalues equal or close to zero, then m component of $\mathbf{x}$ form linear relationships among them. These relationships

107

are of the form

$$\mathbf{h}_1^T \mathbf{x}_t + \mathbf{h}_2^T \mathbf{x}_{t-1} + \ldots + \mathbf{h}_r^T \mathbf{x}_{t-r+1} = 0, \tag{5.2}$$

where $\mathbf{h}_i = [h_{i1}, \ldots, h_{ik}]^T$ are vectors of constants and r $(r \leq p)$ is a lag constant.

By identifying these linear relationships among the time series of multiple sensors, it is possible to select a linearly independent subset.

In the present work, eigenvalues and eigenvectors of the $\Gamma_{\mathbf{x}}(0)$ and $\sum$ are computed using a Jacobi method for real symmetric matrices (Press et al. 1988) and multivariate AR models are fit to $\mathbf{x}$ by least square estimation method. The eigenvalues of the covariance matrices $\Gamma_{\mathbf{x}}(0)$ and $\sum$ of the experimental data are presented in section 5.8.

5.3 Sensor Data Preprocessing (Step 2)

In this work, the digitized data is preprocessed by passing it through appropriate band-pass filters. Section 5.9 describes the band widths of the filters that are used for the experimental data. The preprocessing of the data servers two important purposes:

- The noise accumulated due to non-homogeneities in the workpiece material and electric interferences during the signal transmission through cables and instrumentation is eliminated.

- The frequencies that are outside the frequency band sensitive to flank wear are filtered out.

This preprocessing increases the signal to noise ratio and improves the accuracy of flank wear estimates.

108

5.4 Sensor Data Representation (Step 3)

The main objective of the sensor data representation is to provide the most compact signal representation possible while preserving the structural features of interest in the signal or signals being represented. In the present problem, the structural features of interest are the signal features that are sensitive to flank wear. It is important to select the representation scheme that matches the structural features of interest in the sensor signal and that highlights these structural features from the features of no interest.

Many investigations have shown that certain frequencies of force, vibration, and AE signals are sensitive to flank wear. These frequencies mainly depend on the dynamics of the cutting system, and tool and workpiece materials. If these spectral features that are sensitive to flank wear are extracted from force, vibration, and AE signals, they can be used for flank wear estimation. The success of flank wear estimation depends on how accurately the spectral features can be extracted.

In present methodology, the spectral features are extracted by representing force and vibration signals using autoregressive moving average models, and AE signals by using discrete wavelet transforms.

5.4.1 Representation of Force and Vibration Signals

In the present work, force and vibration signals are modeled by either univariate or multivariate autoregressive moving average (ARMA) models.

A parametric spectral estimation technique such as ARMA modeling gives higher resolution spectral estimates than a nonparametric spectral estimation technique such as FFT, particularly when data records are short.

109

When the spectrum of a sensor signal changes due to progressing flank wear, the coefficients of the ARMA model of the signal which parameterize the spectrum also change. Therefore, the model parameters themselves can be used as wear sensitive spectral features and they can be input a neural network to estimate flank wear. The details of the representation of force and vibration signals using ARMA models are discussed in Chapter 6.

5.4.2 Representation of AE signals

In the proposed method, AE signals are modeled by discrete wavelet transforms (DWT). The DWT possesses low computational complexity and simple computational structure. Algorithmically, the coefficients of DWT are calculated from a quadratic mirror filter implementation (Mallat 1989, Nason and Silverman 1993).

A wavelet transform of a function $f(\cdot)$ is obtained through a combination of dilations and translations of a set of orthogonal wavelet bases. A wavelet basis may be obtained from a self-similar scaling function $\phi(\cdot)$ and a wavelet generator $\psi(\cdot)$. Wavelet transform performs a time-frequency domain analysis of a signal. The standard wavelet representation of a signal $f(t) \in \mathbf{L}^2(\mathbf{R})$ is as follows:

$$f(t) = \sum_{k \in \mathbf{Z}} c_{0,k} \phi_{0,k}(t) + \sum_{m=0}^{j-1} \sum_{k \in \mathbf{Z}} d_{m,k} \psi_{m,k}(t)). \tag{1}$$

This transform may be considered a filtering operation in which the band width, center frequency, and the sampling rate of ψ_{2^j} are twice those of $\psi_{2^{j-1}}$.

The nonlinear nature of the functions $\phi(\cdot)$ and $\psi(\cdot)$ lends wavelets to capture signals of arbitrary shapes. Wavelets have the ability to zoom in on discontinuities, edges, and singularities. The superior spatial localization and scaling properties allow wavelets to represent both the continuous and the transient parts of signals

110

precisely. Since AE signals are time-dependent and consist of both continuous and transient parts, wavelet transforms are better suited to represent AE signals than are ARMA models. The discrete wavelet transform coefficients computed from AE signals are treated as the compressed representation or features of AE signals. These signal features are presented to a neural network to obtain flank wear estimates. The details of the representation of AE signals using DWT are discussed in Chapter 7.

5.5 Tool Wear Estimation (Step 4)

In this research, a neural network is used for relating sensor-signal features to flank wear. This neural network learns the relationship between sensor-signal features and flank wear from a set of representative training examples. Once the neural network learns the mapping between sensor measurements and flank wear, it can compute flank wear estimates for any given inputs from the sensors.

The architecture of the neural network for flank wear estimation is designed to satisfy two important design considerations: it reduces the demand on supervised training data and it remains fault tolerant to fluctuations in metal cutting. This neural network is constructed by integrating Kohonen's feature maps, radial basis functions, and recurrent neural networks. This neural network architecture also guarantees the network convergence, reduces the network training time, and eliminates empiricism in determining the neural network topology. The theoretical and empirical details of this neural network are presented in Chapter 8.

111

5.6 Experimentation

The flank wear estimation method proposed in this work is verified by conducting a set of experiments. In these experiments, sensor signals and flank wear are measured intermittently during a cutting process. From these measurements, a set of supervised data and unsupervised data are generated. A supervised data point is one which has both sensor measurements and the corresponding flank wear value. An unsupervised data point is one which has only sensor measurements but not the corresponding flank wear value. A set of data collected in these experiments is used for training a neural network, and another set of data is used for testing the performance of the neural network. The details of the workpiece, cutting inserts, sensors, and instrumentation are discussed in the following subsections.

5.6.1 Workpiece

AISI 6150, a chromium-vanadium steel alloy, is chosen as the workpiece material for the present experimentation. This material is chosen because it has uniform material properties and high hardness and toughness. Presence of vanadium inhibits grain growth and produces a sound, uniform, fine-grain casting because vanadium is a good deoxidizer and a strong carbide former. When dissolved, vanadium has a marked effect on hardenability and yields high mechanical properties on air cooling. The medium-carbon chromium-vanadium steels have high toughness and strength and are used for axles and springs. The high-carbon chromium-vanadium steels have high hardness and wear resistance and are used for bearings and tools. Table 5.1 gives the material composition of the actual workpiece used in the present experiments and Table 5.2 gives the workpiece heat treatment given to the workpiece.

112

Table 5.1: Composition of AISI 6150 workpiece material

Element	Percent
Fe	*97.14*
C	*0.50*
Mn	*0.77*
P	*0.010*
S	*0.007*
S	*0.25*
Ni	*0.11*
Cr	*0.92*
Mo	*0.03*
V	*0.16*
Cu	*0.10*

Table 5.2: Workpiece heat treatment

AISI 6150
1. *Quenched and Tempered*
2. *Heated to 1550° F*
3. *Oil Quenched*
4. *Tempered at 600° F*
5. *Air cooled*

The cutting stock is in the form of cylinders of 36 inches long and 7 inches diameter. The hardness of the workpiece varies from the surface layers of the cylinder to the center of the cylinder. The hardness of the workpiece at a radius of 3.25 inches is 425 Bhn and at the center of the cylinder is 360 Bhn. To maintain the uniformity of surface roughness in each experiment, the surface of the workpiece is prepared with a finish cut before every run of the experiment.

113

5.6.2 Cutting Tools

The workpiece is cut with uncoated carbide grade K68 inserts. K68 is a tough WC/Co unalloyed grade carbide. These inserts have excellent edge resistance for machining stainless steels, cast iron, and most high-temperature alloys. The geometric specifications of the insert and the insert holder are given in Table 5.3 and Table 5.4 respectively.

Table 5.3: Insert specifications

Insert Specifications	
Specification	*SPG-422*
Material	*Carbide*
Insert Shape	*Square*
Size	1/2″
Thickness	1/8″
Side Relief Angle	11°
Side Rake Angle	5°
Back Rake Angle	0°
Cutting Point Radius	1/32″
Nose Angle	90°

5.6.3 Sensors and Instrumentation

Experiments are conducted on a 20 HP LeBlond 1610 heavy duty lathe. In these experiments a three-axis dynamometer, two accelerometers, and an acoustic emission sensor are used. Figure 5.4 shows the details of instrumentation and Figure 5.5 shows the tooling setup and location of the sensors. The cutting forces are measured in cutting, feed, and radial directions.

114

Table 5.4: Insert holder specifications

Insert Holder Specifications	
Specification	*KSBR-164C*
Insert Holding Method	*Clamp locking*
Insert Shape	*Square*
Holding Style	*Straight Shank*
Side Cutting Edge Angle	15°
End Cutting Edge Angle	15°
Lead Angle	15°
Width of Holder	1″
Height of Holder	1″
Length of Holder	6″

A three-axis Kistler Z3392/b piezo-electric force dynamometer is used to measure machining forces. This dynamometer is located underneath the tool post. The vibration signals are measured in cutting and feed directions. Two PCB accelerometers are used for measuring these vibration signals. The accelerometers are located on the tail of the toolholder in order to protect sensors from the high cutting temperatures and the impact of chips.

It would be better if these accelerometers are located closer to the cutting end of the toolholder instead of at the tail of the toolholder. The vibration signals contain richer flank wear information when the accelerometer is placed nearer to the point of tool-workpiece contact (Martin et al. 1986, Sokolowski et al. 1992). However, in order to isolate the accelerometer from temperature effects and to protect the cables from being damaged by the chips the accelerometers are placed at the tails of the toolholder.
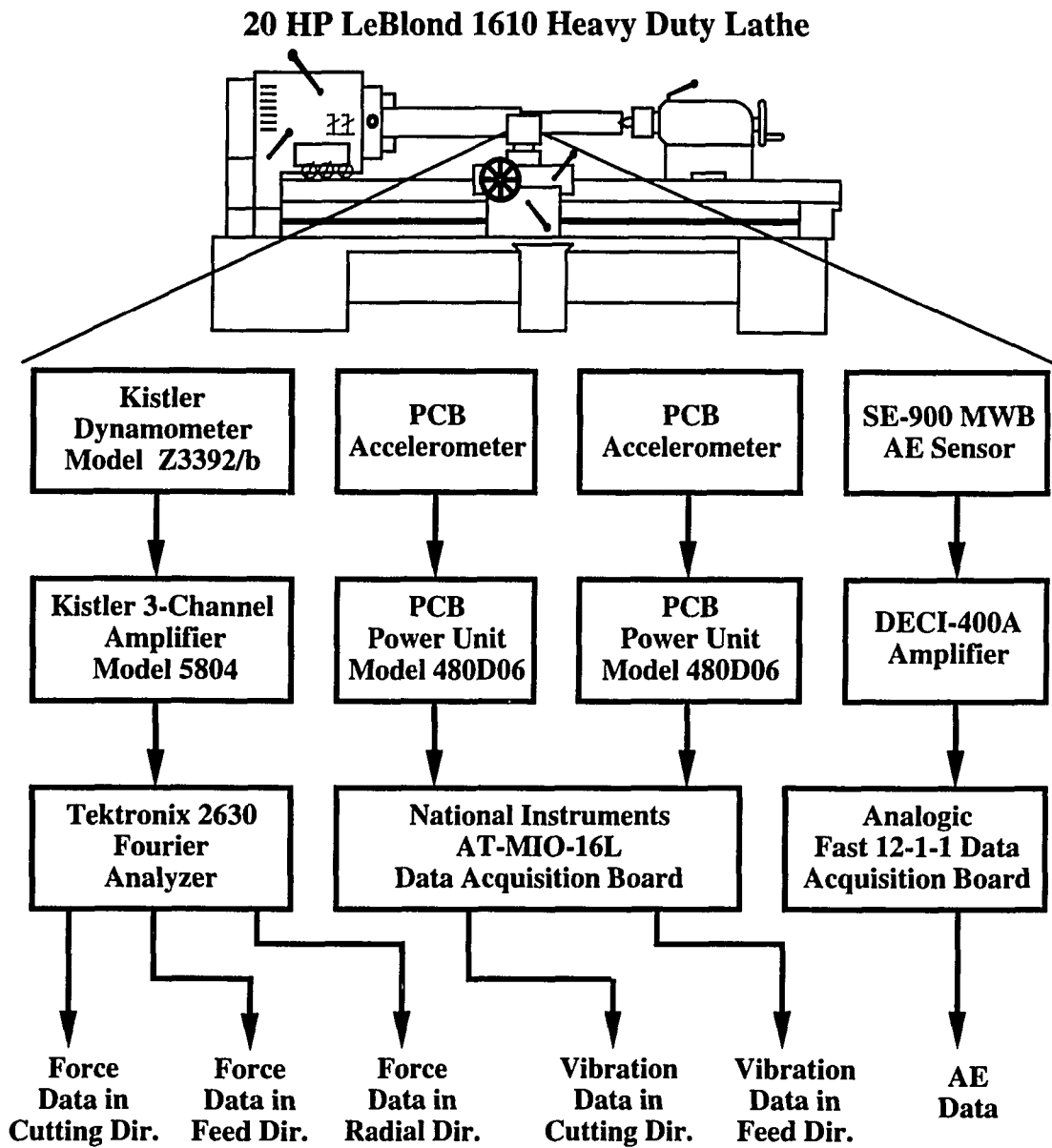
115

20 HP LeBlond 1610 Heavy Duty Lathe

```
┌─────────────────┐  ┌──────────────┐  ┌──────────────┐  ┌──────────────┐
│     Kistler     │  │     PCB      │  │     PCB      │  │  SE-900 MWB  │
│  Dynamometer    │  │ Accelerometer│  │ Accelerometer│  │   AE Sensor  │
│  Model  Z3392/b │  │              │  │              │  │              │
└─────────────────┘  └──────────────┘  └──────────────┘  └──────────────┘
         │                  │                  │                  │
         ▼                  ▼                  ▼                  ▼
┌─────────────────┐  ┌──────────────┐  ┌──────────────┐  ┌──────────────┐
│ Kistler 3-Channel│ │     PCB      │  │     PCB      │  │              │
│   Amplifier     │  │  Power Unit  │  │  Power Unit  │  │  DECI-400A   │
│   Model 5804    │  │ Model 480D06 │  │ Model 480D06 │  │  Amplifier   │
└─────────────────┘  └──────────────┘  └──────────────┘  └──────────────┘
         │                  │                  │                  │
         ▼                  ▼                  ▼                  ▼
┌─────────────────┐  ┌──────────────────────────────┐  ┌──────────────┐
│ Tektronix 2630  │  │    National Instruments      │  │   Analogic   │
│    Fourier      │  │        AT-MIO-16L            │  │ Fast 12-1-1 Data│
│    Analyzer     │  │   Data Acquisition Board     │  │Acquisition Board│
└─────────────────┘  └──────────────────────────────┘  └──────────────┘
```

| Force Data in Cutting Dir. | Force Data in Feed Dir. | Force Data in Radial Dir. | Vibration Data in Cutting Dir. | Vibration Data in Feed Dir. | AE Data |

Figure 5.4: Lathe, sensors, and instrumentation used for data acquisition
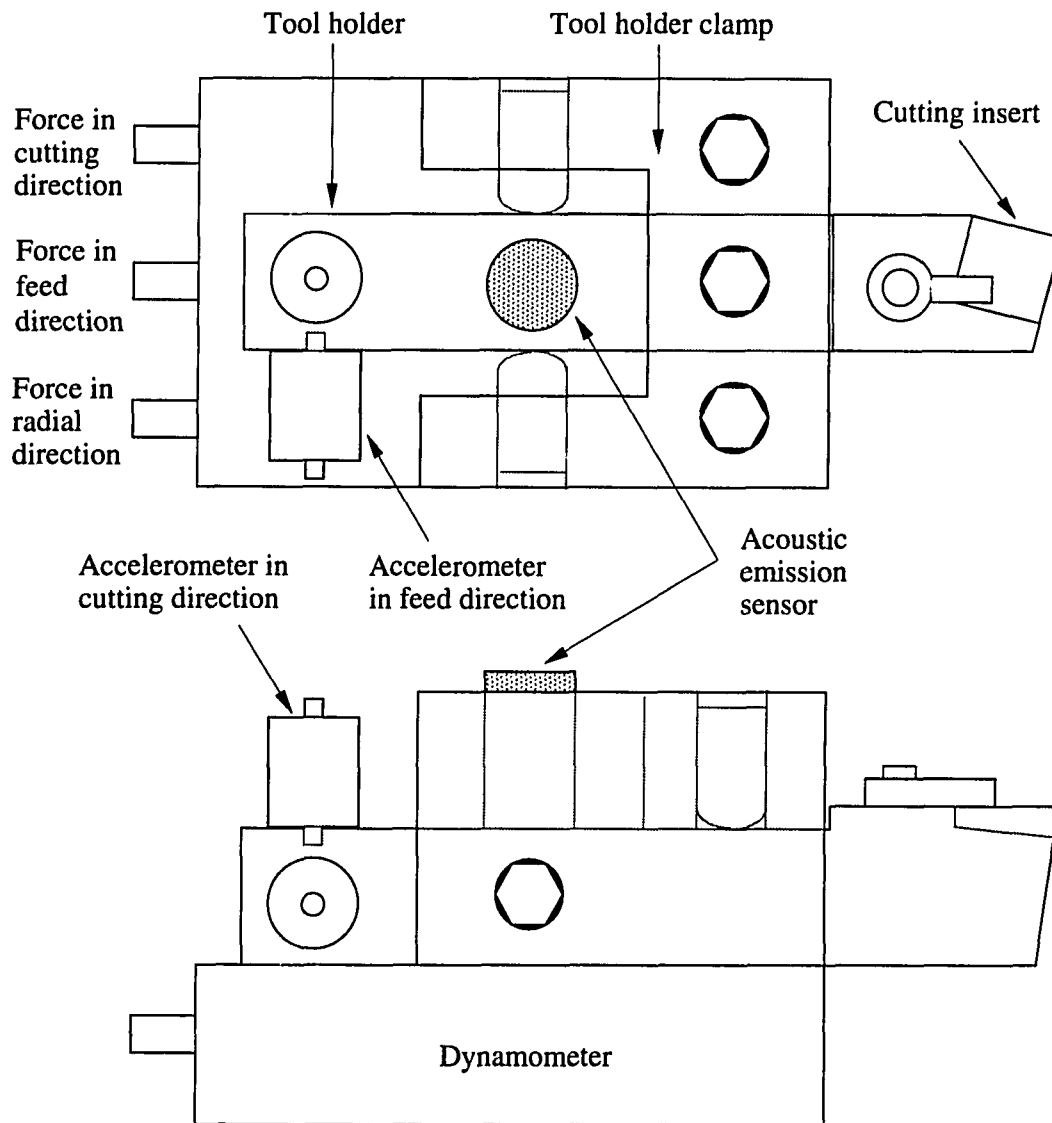
116

Figure 5.5: Tooling setup and sensor
location

117

The AE signals are measured using a SE-900 MWB wide bandwidth AE sensor with an output response of 100 kHz to 900 kHz. An AE sensor is mounted in the middle of the toolholder. A magnetic clamp is used to secure the AE sensor to the toolholder.

The force signals from the dynamometer are passed through a 5804 series Kistler three channel charge amplifier. The vibration signals from the two accelerometers are passed through two separate 480D06 type PCB amplifiers. The signals from the AE sensor are passed through a DECI-400A preamplifier. It has a fixed gain of 40 dB and a passband range of 100 kHz – 2 MHz.

The force signals are digitized using a Tektronix 2630 Fourier Analyzer. The vibration signals are digitized using a National Instrumentation data acquisition board installed in a IBM compatible PC. The AE signals are digitized using an Analogic Fast 12-1-1-1 high-frequency data acquisition board installed in a IBM compatible PC.

The sampling rates for digitization of sensor signals are determined by observing the spectra of the signals. Figures 5.6, 5.7, and 5.8 show the typical spectra of force, vibration, and AE signals respectively when flank wear is 0.0035 inches and 0.0180 inches. By observing these spectra, the force signals are digitized with a sampling frequency of 3 kHz, vibration signals with 26 kHz, and AE signals with 1 MHz. All these signals are sampled for a length of 4096 points.

5.6.4 Data Collection

Once the cutting is started with a fresh insert edge, force, vibration, and AE signals are sampled once in every 60 seconds until the cutting edge is considered worn out. The cutting process is also interrupted once in every 60 seconds
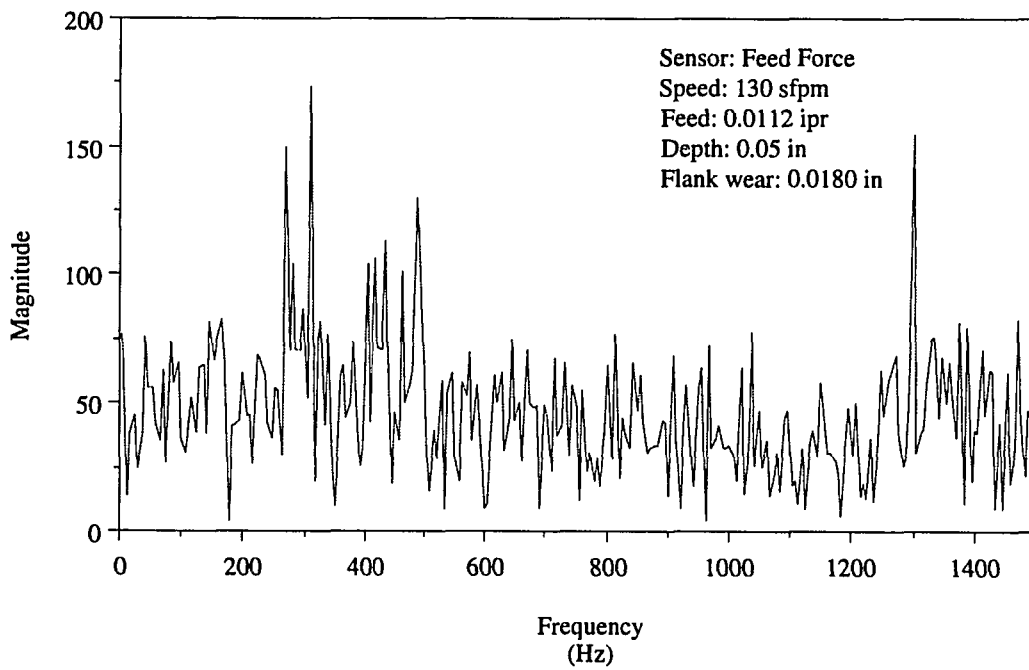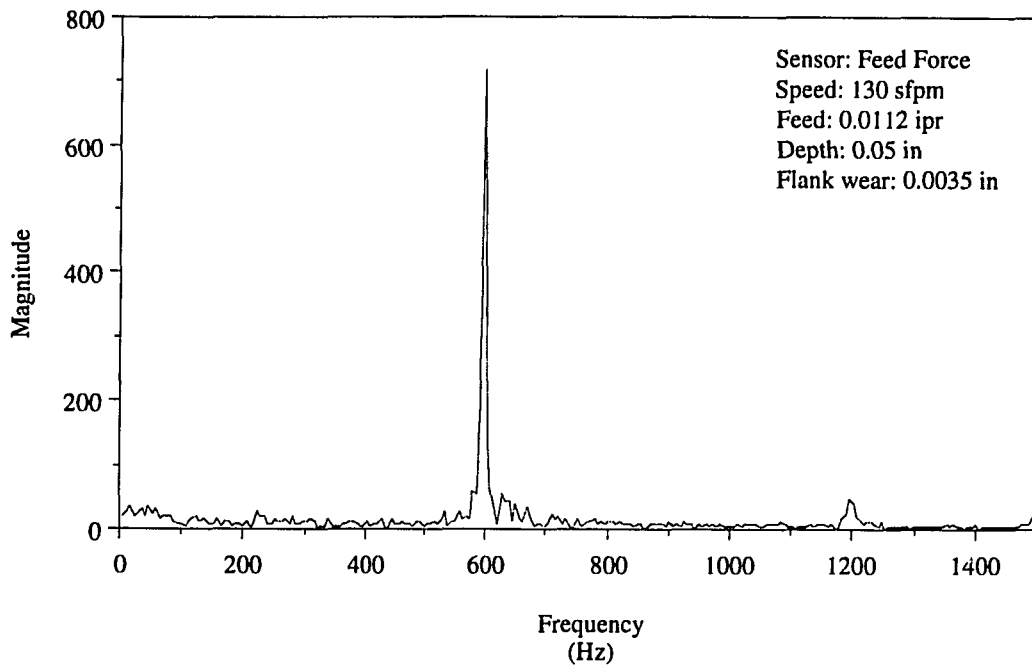
118

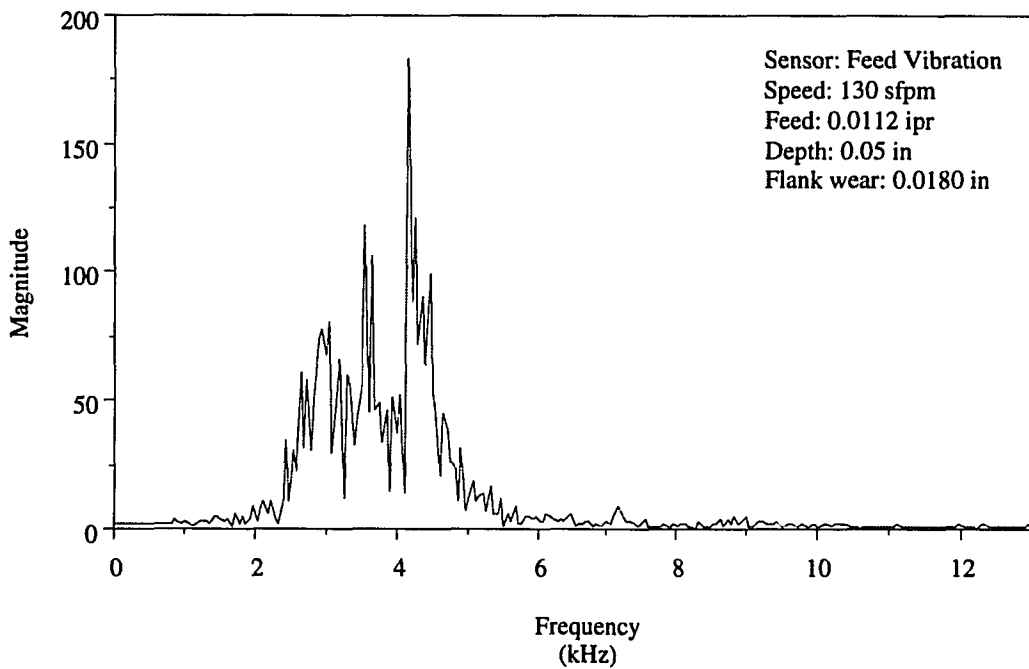**Figure 5.6:** Frequency spectrum of feed force signals
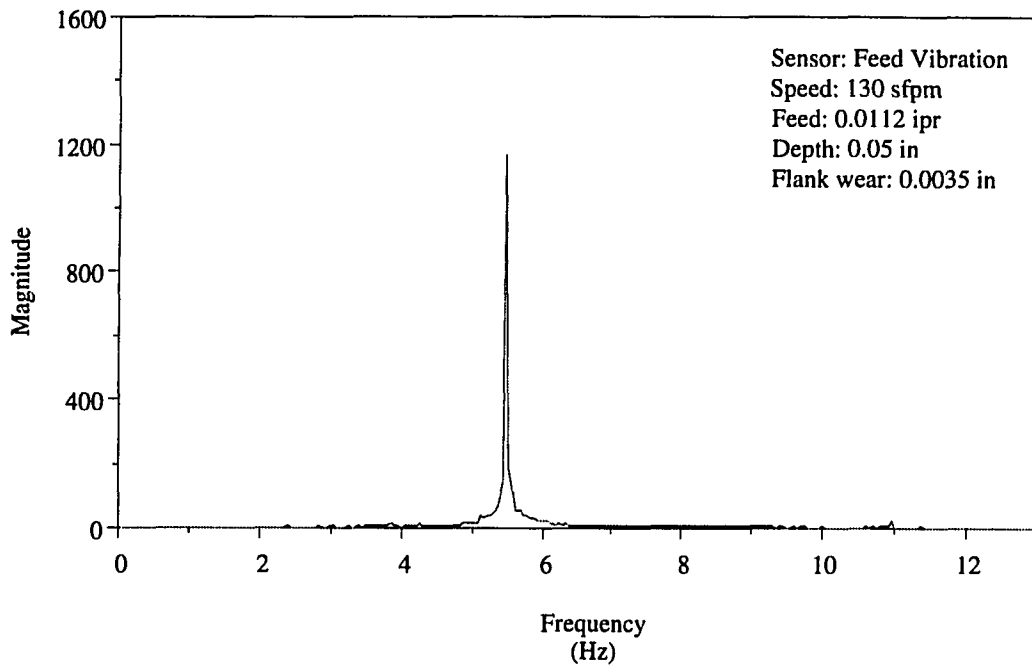
119

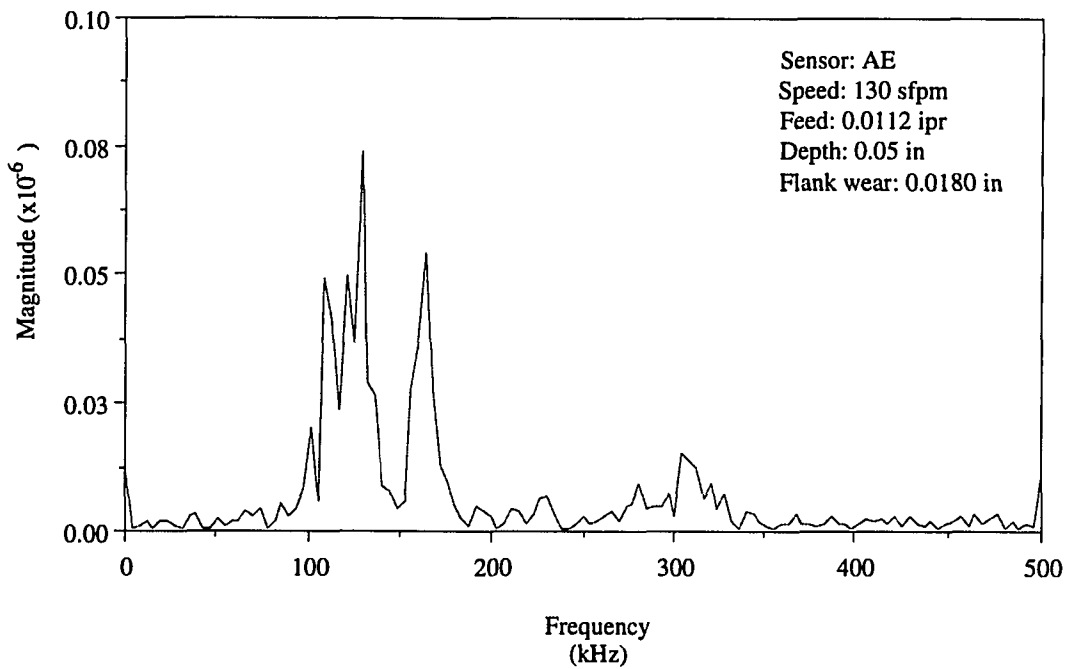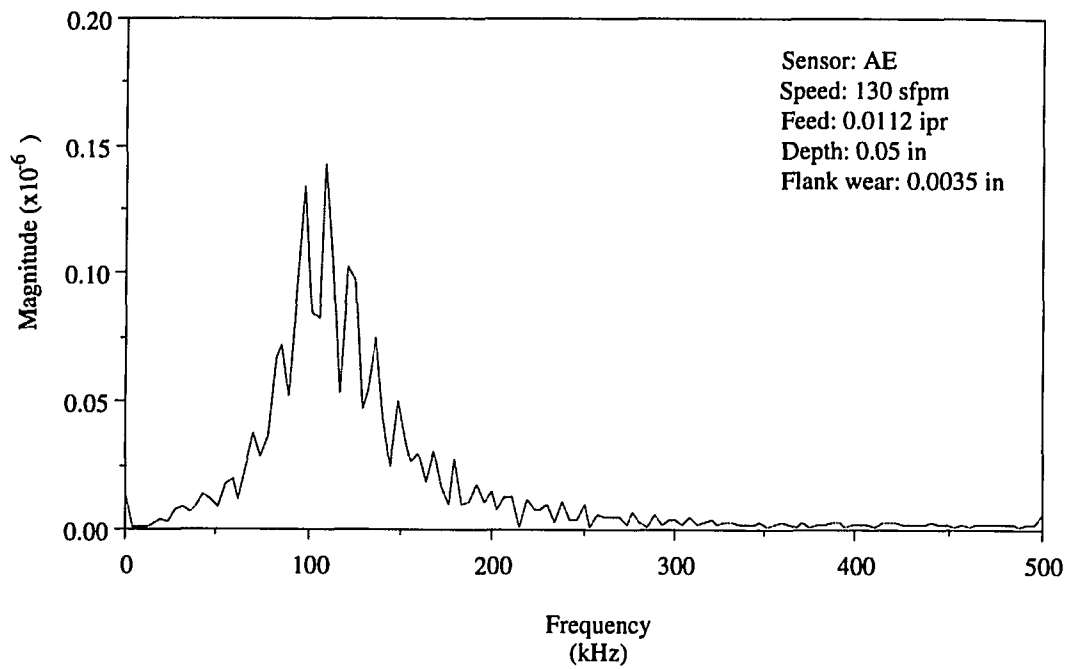**Figure 5.7:** Frequency spectrum of vibration signals in feed direction

120

Figure 5.8: Frequency spectrum of AE signals

121

for off-line flank wear measurement. The flank wear is measured using a tool-maker's microscope. When the cutting edge develops an average flank height of 0.018 inches, it is considered to be a worn out edge. This limit is chosen according to the criteria recommended by the ISO to define the effective tool life for carbide tools (ISO 1972). This criteria recommends a flank wear height between 0.0118 and 0.0236 inches (0.3 and 0.6 mm) to consider a tool to be worn out. In the present experiments, 0.0180 inches (the average of 0.0118 and 0.0236), is taken as the criteria to consider a tool to be worn out. The experiments have shown that an insert edge usually runs for about 10 to 14 minutes before it is considered worn out. The following steps are used to record sensor data and flank wear:

1. Start cutting a workpiece with a fresh insert edge.

2. Continue cutting for about 55 seconds without sampling any data.

3. At the 55th second press the trigger switch. This switch will simultaneously trigger force, vibration, and AE data acquisition systems. The data acquisition is complete on all three data acquisition systems in less than 2 seconds.

4. Interrupt the cutting process at the 60th second.

5. Measure flank wear on the cutting tool edge using a toolmaker's microscope.

6. Resume cutting the workpiece with the same insert edge.

7. Repeat steps 2 to 6 until the average flank wear height on the major flank face exceeds 0.018 inches.

8. Repeat steps 1 to 7 for every combination of speed and feed specified in the experimental design.

5.7 Experimental Design

In this work, three sets of experiments are conducted. In the rest of the thesis, these three sets of experiments are referred to as Set 1, Set 2 and Set 3. In all three sets of experiments the depth of cutting is kept constant at 0.05 inches and only the cutting speed and feed are varied. For easy referencing, each cutting tool used in these three experiments is assigned a unique number.

The Set 1 experiments are conducted according to a full-factorial design with three levels of speed and five levels of feed. Table 5.5 lists the speed and feed combinations and the corresponding cutting tool numbers in Set 1 experiments. The operating conditions for this set of experiments are chosen such that the cutting tools develop predominantly flank wear with a negligible amount of crater wear.

Table 5.5 Grid of speed and feed in Set 1 experiments

Set 1 Experiments		Speed (sfpm)		
		100	130	160
Feed (ipr)	0.0064	Tool-1	Tool-6	Tool-11
	0.0088	Tool-2	Tool-7	Tool-12
	0.0112	Tool-3	Tool-8	Tool-13
	0.0136	Tool-4	Tool-9	Tool-14
	0.0154	Tool-5	Tool-10	Tool-15

123

The Set 2 experiments are conducted according to a full-factorial design with three levels of speed and three levels of feed. Table 5.6 lists the speed and feed combinations and the corresponding cutting tool numbers in Set 2 experiments. The operating conditions in these experiments, as in Set 1 experiments, cause mostly flank wear on the cutting tools.

Table 5.6: Grid of speed and feed in Set 2 experiments

Set 2 Experiments		Speed (sfpm)		
		100	130	160
Feed (ipr)	0.0088	Tool-16	Tool-19	Tool-22
	0.0112	Tool-17	Tool-20	Tool-23
	0.0136	Tool-18	Tool-21	Tool-24

The Set 3 experiments are conducted according to a full-factorial design with two levels of speed and three levels of feed. Table 5.7 lists the speed and feed combinations and the corresponding cutting tool numbers in Set 3 experiments. The cutting speeds in these experiments are higher than those used in Set 1 and Set 2 experiments. These cutting speeds develop not only flank wear but also crater wear on the cutting tools.

Table 5.8, gives a picture of how the operating conditions in Set 1, Set 2, and Set 3 experiments are related. The cutting tools numbered from 1 to 15 belong to the Set 1 experiments. The cutting tools numbered from 16 to 24 belong to the Set 2 experiments, and those numbered from 25 to 30 belong to the Set 3 experiments. This table shows that the operating conditions used in Set 2 experiments

124

Table 5.7: Grid of speed and feed in Set 3 experiments

Set 3 Experiments		Speed (sfpm)	
		190	220
Feed (ipr)	0.0088	Tool-25	Tool-28
	0.0112	Tool-26	Tool-29
	0.0136	Tool-27	Tool-30

are the subset of the operating conditions used in Set 1 experiments. The cutting speeds used in Set 3 experiments are higher than the cutting speeds used in Set 1 experiments.

Table 5.8: Grid of speed and feed in Set 1,2, 3 experiments

Set 1, 2, 3 Experiments		Speed (sfpm)			Speed (sfpm)	
		100	130	160	100	130
Feed (ipr)	0.0064	Tool-1	Tool-6	Tool-11		
	0.0088	Tool-2 / Tool-16	Tool-7 / Tool-19	Tool-12 / Tool-22	Tool-25	Tool-28
	0.0112	Tool-3 / Tool-17	Tool-8 / Tool-20	Tool-13 / Tool-23	Tool-26	Tool-29
	0.0136	Tool-4 / Tool-18	Tool-9 / Tool-21	Tool-14 / Tool-24	Tool-27	Tool-30
	0.0154	Tool-5	Tool-10	Tool-15		

125

The data from Set 1 experiments is used for training a neural network for flank wear estimation. The data from Set 2 and Set 3 experiments are used to study the performance of the neural network. The data in Set 2 experiments test the performance of the neural network at the same operating conditions that were used to train the neural network. The data in Set 3 experiments test the performance of the neural network at higher cutting speeds than those used to train the neural network. The data in Set 3 experiments is also useful in studying the performance of the neural network when both flank wear and crater wear are present on a cutting tool.

5.8 Linear Independence of Sensor Signals

As described in section 5.2.1, the linear independence of three force signals and two vibration signals are verified by principal component analysis. This analysis is conducted on the data collected on tool#20 at cutting speed 130 sfpm and feed 0.0112 ipr. The choice of tool#20 is arbitrary. On this cutting tool, 10 data points were collected between the initial flank wear level and the final flank wear level with a time interval of one minute. Each data point consists of signal records of length 4096 from three force sensors (in cutting, feed, and radial direction), two vibration sensors (in cutting and feed direction), and an AE sensor. Out of 4096 points in each signal record, only the first 512 points are used for the present analysis. That means the time series of 512 points from force and vibration sensors are considered for the principal component analysis. Note that the time series from the AE sensor is not used in this analysis.

The principal component analysis is conducted separately on 10 data points that were collected on the cutting tool#20. The results for force sensors are

126

presented in Tables 5.9, 5.10, and 5.11. The results for vibration sensors are presented in Tables 5.12, 5.13, and 5.14. The results in Tables 5.9 and 5.12 present the principal component analysis results on the original time series of force and vibration signals respectively. The results in Tables 5.10 and 5.13 present the principal component analysis results on the residual of the multivariate AR(10) model of force and vibration signals respectively. The results in Tables 5.11 and 5.14 present the principal component analysis results on the residual of the multivariate AR(20) model of force and vibration signals respectively. The main source code for principal component analysis is presented in Appendix A.

The eigenvalues computed in most of the above cases are far from close to zero. This indicates that the time series of both force and vibration signals are linearly independent. In other words, force sensors in cutting, feed, and radial direction and vibration sensors in cutting and feed direction carry independent information about the cutting process. Therefore, in the present application force sensors in cutting, feed, and radial direction and vibration sensors in cutting and feed direction are used for flank wear estimation. Note that in addition to these five sensors, an AE sensor is also used for flank wear estimation.

5.9 Preprocessing Sensor Data

As explained in section 5.3, the preprocessing of the digitized sensor data is done by passing the sensor data through appropriate Bessel band-pass filters. The lower and upper cut-off frequencies for the band-pass filter are chosen by studying a representative set of power spectra of sensor signals. Figure 5.6, 5.7 and 5.8 show the typical spectra of force, vibration, and AE sensors. Table 5.15 gives lower and upper cut-off frequencies of force, vibration, and AE signals. However a more

127

Table 5.9: Eigenvalues of time series of three cutting forces. These time series are collected from a single cutting edge at different flank wear levels. Speed: 130 sfpm, Feed: 0.0112 ipr, Depth of cut: 0.05 inches.

Point	Flank Wear	Eigenvalue 1	Eigenvalue 2	Eigenvalue 3
1	0.0035	0.106937	0.861926	0.004305
2	0.0056	0.102880	1.141307	0.016752
3	0.0071	0.096087	1.292137	0.005974
4	0.0090	0.103092	1.234015	0.002709
5	0.0103	0.027462	0.008514	0.032711
6	0.0126	0.036603	0.011542	0.031279
7	0.0143	0.116835	0.045250	0.025200
8	0.0156	0.092655	0.269875	0.017949
9	0.0174	0.040265	0.009510	0.016572
10	0.0027	0.018807	0.008631	0.011154

Table 5.10: Eigenvalues of residual of the multivariate AR(10) model of three cutting forces. These time series are collected from a single cutting edge at different flank wear levels. Speed: 130 sfpm, Feed: 0.0112 ipr, Depth of cut: 0.05 inches.

Point	Flank Wear	Eigenvalue 1	Eigenvalue 2	Eigenvalue 3
1	0.0035	0.000725	0.000872	0.000258
2	0.0056	0.001771	0.003146	0.001029
3	0.0071	0.001766	0.002581	0.000669
4	0.0090	0.001509	0.000870	0.000465
5	0.0103	0.004175	0.002102	0.008950
6	0.0126	0.009991	0.002920	0.006335
7	0.0143	0.013624	0.003675	0.005888
8	0.0156	0.009522	0.002756	0.004654
9	0.0174	0.008415	0.002425	0.004642
10	0.0027	0.005452	0.001593	0.002992

Table 5.11: Eigenvalues of residual of the multivariate AR(20) model of three cutting forces. These time series are collected from a single cutting edge at different flank wear levels. Speed: 130 sfpm, Feed: 0.0112 ipr, Depth of cut: 0.05 inches.

Point	Flank Wear	Eigenvalue 1	Eigenvalue 2	Eigenvalue 3
1	0.0035	0.000581	0.000792	0.000262
2	0.0056	0.001980	0.003010	0.001026
3	0.0071	0.001459	0.001943	0.000628
4	0.0090	0.001643	0.000930	0.000519
5	0.0103	0.002486	0.001087	0.005428
6	0.0126	0.005667	0.001510	0.003859
7	0.0143	0.008193	0.002445	0.003805
8	0.0156	0.007443	0.001964	0.002676
9	0.0174	0.004784	0.001155	0.002951
10	0.0027	0.003057	0.000585	0.001860

Table 5.12: Eigenvalues of time series of two vibration signals. These time series are collected from a single cutting edge at different flank wear levels. Speed: 130 sfpm, Feed: 0.0112 ipr, Depth of cut: 0.05 inches.

Point	Flank Wear	Eigenvalue 1	Eigenvalue 2
1	0.0035	0.027614	20.269867
2	0.0056	0.081825	40.006184
3	0.0071	0.000001	0.000002
4	0.0090	0.359844	35.127215
5	0.0103	0.000365	0.006025
6	0.0126	0.298323	2.417005
7	0.0143	0.281171	43.073801
8	0.0156	0.429933	3.249731
9	0.0174	0.258988	2.252180
10	0.0027	0.009245	0.022541

129

Table 5.13: Eigenvalues of residual of the multivariate AR(10) model of two vibration signals. These time series are collected from a single cutting edge at different flank wear levels. Speed: 130 sfpm, Feed: 0.0112 ipr, Depth of cut: 0.05 inches.

Point	Flank Wear	Eigenvalue 1	Eigenvalue 2
1	0.0035	0.000862	0.008893
2	0.0056	0.001950	0.021213
3	0.0071	0.000000	0.000001
4	0.0090	0.002578	0.020710
5	0.0103	0.000012	0.000027
6	0.0126	0.004299	0.011764
7	0.0143	0.003354	0.025956
8	0.0156	0.005398	0.012694
9	0.0174	0.005434	0.013784
10	0.0027	0.000163	0.000419

Table 5.14: Eigenvalues of the residual of the multivariate AR(20) model of two vibration signals. These time series are collected from a single cutting edge at different flank wear levels. Speed: 130 sfpm, Feed: 0.0112 ipr, Depth of cut: 0.05 inches.

Point	Flank Wear	Eigenvalue 1	Eigenvalue 2
1	0.0035	0.000782	0.004410
2	0.0056	0.001332	0.009889
3	0.0071	0.000000	0.000001
4	0.0090	0.002781	0.009336
5	0.0103	0.000008	0.000017
6	0.0126	0.005235	0.002304
7	0.0143	0.002792	0.015160
8	0.0156	0.007477	0.003560
9	0.0174	0.005777	0.002977
10	0.0027	0.000152	0.000059

130

rigorous data preprocessing is possible by employing an appropriate combination of low-pass, high-pass, and band-pass filters.

In the present work, the band-pass filtering of the data is performed using the signal processing routines provided by SIG package. Sample source codes for band-pass filtering of feed force data, feed vibration data, and AE data are given in Appendices B, C, and D respectively.

Table 5.15: Cut-off frequencies of band-pass filters

Band-Pass Cut-Off Frequencies		
Signal	Lower Cut-Off	Upper Cut-Off
Force in Cutting Direction	3 Hz	1 kHz
Force in Feed Direction	3 Hz	1 kHz
Force in Radial Direction	3 Hz	1 kHz
Vibration in Cutting Direction	1 kHz	10 kHz
Vibration in Feed Direction	1 kHz	10 kHz
Acoustic Emission	50 kHz	300 kHz

5.10 Conclusions

This chapter described a methodology for flank wear estimation and the experiments for verifying the methodology. The methodology involves four important steps: sensor selection, sensor data preprocessing, sensor data representation, and flank wear estimation using a neural network. The issues related to sensor selection were discussed in Chapter 4. The sensor data preprocessing and the experimentation details are described in this chapter. Chapters 6 will discuss the representation of force and vibration signals using ARMA models. Chapter 7 will discuss the representation of AE signals using the DWT. Chapter 8 will describe the theoretical and implementation details of the neural network for flank wear estimation.

TOOL WEAR ESTIMATION USING
FORCE AND VIBRATION MEASUREMENTS

6.1 Introduction

This chapter primarily discusses the issues related to the representation of force and vibration signals using time series models. In this problem, sensor data representation refers to the identification and extraction of features sensitive to flank wear from sensor data.

It has been reported by many researchers that spectral characteristics of force and vibrations signals are affected by tool wear (Weller et al. 1969, Martin et al. 1974, Taglia et al. 1976, Pandit and Kashou 1982, Lindstrom and Lindberg 1983, Martine et al. 1986, Rao 1986, Jiang et al. 1987, Rangawala and Dornfeld 1990, Sokolowski et al. 1992). This indicates that spectral features of force and vibration signals can be used for flank wear estimation. Therefore, this chapter addresses the problem of representing spectral features of force and vibration signals.

Section 6.2 reviews the spectral characteristics of force and vibration signals in a turning processes and identifies the requirements of a method to estimate the spectra of these signals. Section 6.3 briefly describes the basics of nonparametric and parametric spectral estimation methods and discusses their application to the current work. Section 6.4 discusses time series models to represent force and

132

vibration signals. Section 6.5 reviews the concepts of univariate and multivariate ARMA models and explains how these models can be used for sensor data representation.

6.2 Spectral Properties of Sensor Signals

This section reviews the results of various studies on spectral characteristics of force and vibration signals in turning processes. Based on this review, the requirements of a spectral estimation method to characterize force and vibration spectra are identified.

6.2.1 Spectral Characteristics of Force Signals

In the literature the spectral characteristics of vibration signals are studied more extensively than those of force signals. Most of the research on cutting forces has been focused on relating DC cutting forces (average cutting forces) to tool wear. However, Eman and Wu (1987) showed that dynamic cutting forces contain richer information about tool-workpiece interaction during machining than DC forces do.

Lindstrom and Lindberg (1983) showed that dynamic changes in cutting forces can be used for detecting flank wear. They have further shown that force signals are more effective than acceleration signals in detecting tool wear due to the transfer function of the tool holder and the tool holder components.

Dan and Mathew (1990) reported that the power and the autocorrelation coefficient of force signals reflect tool wear. The high frequency component and amplitude of feed force keep increasing during normal wear stage and subsequently start decreasing during advanced wear stage and tool breakage.

133

Rao (1986) studied the frequency properties of both force and vibration signals in cutting, feed, and radial directions for the possible existence of a parameter correlated to flank wear. He found that force and vibration signals in feed direction carry useful information related to flank wear, and signals in the other two coordinate directions do not reveal any useful information.

Even under stable turning conditions, oscillations of small magnitudes exist. These oscillations are generally around the natural frequencies of the toolholder dynamometer systems. In general, lower frequencies dominate when a tool is new. However, the weakest mode with the lowest damping ratio, which is at the first natural frequency of the cantilever portion of the toolholder, tends to dominate as the tool wears. These shifts in dominant frequencies as opposed to pure increase or decrease in vibration or force amplitude appears to be a more common occurrence as the tool wear progresses.

Rangawala and Dornfeld (1990) observed that vibrations in cutting force direction are induced by flank wear both in high frequency (>5 kHz) and low frequency (<300 Hz) regions. The high frequency vibrations are attributed to vibrations of the tool holder and the low frequency vibrations to workpiece vibrations. The force spectrum is dependent on process variables such as cutting velocity, feed rate, and oscillations in shear angle during chip formation.

Sata and Matsushima (1974) found that the force spectrum is affected by chip formation, chatter, and build-up edge.

6.2.2 Spectral Characteristics of Vibration Signals

Martine et al. (1974) found that vertical vibrations (in the cutting direction) of a cutting tool in the course of stable machining are almost sinusoidal, with a

134

frequency equal to the natural frequency of the tool. For XC 38 steel at cutting speeds of 75 to 150 m/min, the power of the acceleration signal obtained by spectral analysis is a linear function of cutting speed and tool wear, varying in the ratio of 1:100 between a new tool and a worn tool.

Martine et al. (1986) found that the amplitude of vibration signals in the frequency range of 0 to 2.5 kHz increases as a cutting tool approaches the worn out stage. The signal amplitude increases up to 30% of the original value and the average increase is about 20%.

Jiang et al. (1987) found that vibration signals from a cutting tool contain a large amount of information about the cutting process and are very sensitive to tool wear, and the frequency composition of the signal energy varies regularly with the gradually progressing tool wear. The distribution of the signal energy in the frequency domain is influenced by various factors such as spindle speed and length of the work piece (Jiang et al. 1987). The vibration signal is very sensitive to the tool wear condition and the frequency composition of the energy of the signal varies regularly with the development of tool wear.

Some higher peaks appear near 117 Hz and 510 Hz which are close to the first and the second order resonance frequency of the cutting system. The energy of these spectral peaks increases step by step during normal wear stages, but their frequency positions do not change throughout the wear progress. During the micro-breakage stage between normal and rapid tool wear stages, the signal amplitude in the frequency range below 20 Hz rises significantly with an unstable pattern of strong rises and falls. However, the energy of the resonant frequencies, 117 Hz and 510 Hz, decreases slowly.

135

Pandit and Kashou (1982) observed that the actual power contribution in the frequency range of 4–5 kHz decreases with increasing flank wear, passes through a minimum, and then increases at the later stage of the wear. The minimum value occurs when flank wear is between 0.254 mm and 0.38 mm. It was also found that the frequency band that is sensitive to tool wear is the natural frequency of the tool holder.

Sokolowski et al. (1992) observed that the amplitude of vibration velocity at a certain dominant frequency changes monotonically with increasing flank wear. This dominant frequency shifts if the configuration of tool holder and shank is changed. This dominant frequency is considered to be the natural frequency of the overhang portion of the tool shank. In each direction of the machine tool system, the same dominant frequency is observed. The largest increase in the amplitude of vibration velocity is recorded in the feed direction. However, the largest amount of stochastic behavior of the amplitude is observed in the feed direction as well. The increase of normalized amplitude is more consistent in the cutting and radial directions rather than in the feed direction.

The increase in cutting speed causes an increase in the amplitude of the dominant frequency of the vibration velocity. This increase in amplitude is more prominent in the radial direction and less prominent in the feed direction. The increase in feed rate results in a substantial decrease in the amplitude in the cutting direction at low depths of cut. Otherwise, the change in feed rate has little influence on the spectral characteristics of vibration signals.

The dominant frequency shifts with the progress of flank wear. This shift in the dominant frequency is considered to be connected with the changes in transfer function of the machine tool-cutting system.

136

Taglia et al. (1976) noticed that power spectra, cross power spectra, and coherence function of vibration signals in the low frequency range of 0 to 2.5 kHz undergo some noticeable and regular variation with wear. This variation is less sensitive and almost uncertain in the range of 0 to 12.5 kHz. Though the signal power contained under 2.5 kHz is much smaller than that contained in the upper frequencies, it is magnified when either vibration velocity or vibration displacement signal is regarded instead of acceleration signal.

Weller et al. (1969) found that vibrations generated in a cutting action contain considerably more high-frequency energy when the cutting edge is worn than when the edge is sharp. The vibration energy is divided into two frequency components: high-frequency range and the low-frequency range. The low-frequency component is typically in the range of 0 to 4 kHz, while the high-frequency component is typically in the range of 4 to 8 kHz. The low-frequency component band energy for a sharp cutting edge and a worn cutting edge is about the same, but the high-frequency band energy for a worn cutting edge is considerably higher than for a sharp one. It was found that the total amount of high-frequency vibration energy increases as the flank wear height increases. This phenomenon is observed consistently, even if the cutting parameters such as feed, speed, and depth of cut are varied within wide ranges.

Further analysis of vibration signals showed that the additional high frequency vibration energy generated by a worn cutting edge is concentrated in a few narrow frequency bands. These vibrations comprise a fundamental frequency plus several overtones, including at least the first and the second harmonics. The fundamental frequency is at the resonant frequency of the system because it is excited by the friction between the cutting edge and the workpiece. The fundamental frequency

137

and the overtones are influenced by a number of factors such as workpiece material, geometry of insert and toolholder, and size of tool block and workpiece.

6.2.3 Requirement of Spectral Estimation Method

The previous sections reviewed various properties of the spectra of force and vibration signals in turning processes. The following are the main properties that are to be considered in selecting a spectral estimation method:

1. The dominant frequencies in either force or vibration spectra are usually the natural frequencies of the cantilever portion of the toolholder, toolholder-dynamometer system, or machine tool system.

2. The amplitudes of the dominant frequencies are usually sensitive to the flank wear.

3. These dominant frequencies, which are sensitive to tool wear, continue to shift with increasing flank wear. Sokolowski et al. (1992) have given an explanation to this shift in dominant frequencies. The shift in frequencies is caused by changes in the transfer function of the machine tool-cutting process system. The increasing flank wear, inturn, increases the cutting force and shifts the operating point of the force-displacement characteristics. The shift in the operating point can reduce the machine tool stiffness and consequently decrease the natural frequency. The increasing flank wear also alters the stress in the cutting edge region and the friction between the workpiece and the tool flank face. These changes affect the stiffness and damping of the cutting process transfer function which directly influences the natural frequencies of the machine tool-cutting process system.

4. The spectral characteristics of the force and vibration signals are not only af-

138

fected by tool wear but also by various other factors such as cutting conditions, workpiece material properties, and cutting tool and toolholder geometry.

5. Though the influence of tool wear on the spectral characteristics of force and vibration signals is obvious, the change in the pattern of frequencies and their amplitudes is neither linear nor monotonic with respect to tool wear progress.

Most of the methods reported in the literature (Rangwala and Dornfeld 1990, Emel 1991, Leem and Dreyfus 1992, Burke 1993) that use spectral features for tool wear estimation consist of the following steps:

1. Determine the spectrum of a sensor signal via an FFT operation on the sample data.

2. Consider the power spectrum as a feature vector whose components are the signal power at discrete frequencies.

3. Reduce the size of the feature vector by retaining only selected discrete frequencies that are more sensitive to the tool wear and less sensitive to the noise. This selection is typically done based on a representative set of power spectra of the sampled data. Once these discrete frequencies are selected, they are considered to be valid at all operating conditions and tool wear levels.

4. Use these reduced feature vectors as inputs to a neural network or a statistical procedure to compute tool wear estimates.

Although this approach to tool wear estimation through FFT-based spectral estimation has shown promising results, it has some drawbacks:

1. It assumes that the dominant frequencies in sensor-signal spectra undergo little shift during the course of tool life, and their amplitude variation is the indicator of the tool wear. Since the dominant frequencies do change with increasing flank wear, changes in both frequency position and amplitude must

139

be considered in the signal representation.

2. Since only the amplitude changes of certain frequencies are used for tool wear estimation, the useful information indicated by the changes in the position of all dominant frequencies and their amplitudes is not considered.

3. In this approach, the spectral estimation is based on an FFT operation on the sample data of finite duration. In the present research, the resolution and the fidelity of the spectra may not be satisfactory because as these spectra are computed from short data records using an FFT operation.

To overcome these limitations, spectral features chosen for tool wear estimation should characterize the entire spectrum of a sensor signal with good resolution. The next section will briefly discuss the spectral estimation methods and identify the choice of a spectral estimation technique for the present application.

6.3 Spectral Estimation Methods

The power spectral density (PSD), or spectrum, of a discretely sampled deterministic or stochastic process refers to the distribution of the energy of the discrete time series over a frequency domain. Many spectral estimation techniques have been developed, but they can broadly be classified into two main groups: nonparametric and parametric. While nonparametric methods obtain the spectrum directly from the time series data, parametric methods compute the spectrum from a closed-form mathematical model of the time series data. Both nonparametric and parametric methods may take a variety of forms. Traditional spectral estimation techniques based on the Fourier transform operation are some of the well known nonparametric methods. The commonly used parametric methods are autoregressive, moving average, and autoregressive moving average models.

140

The following sections briefly discuss the nonparametric and parametric spectral estimation methods.

6.3.1 Nonparametric Methods of Spectral Estimation

The spectral estimation through Fourier transform is possible in two ways. The first method was introduced by Blackman-Tukey (1959). This method estimates the autocorrelation lags from the measured data, windows the autocorrelation estimates in an appropriate manner, and then Fourier transforms the windowed lag estimates to obtain the PSD estimate. The second method is the periodogram approach based on the fast Fourier transform (FFT) introduced by Cooley and Tukey (1965). The periodogram spectral estimate is obtained as the squared magnitude of the output values from an FFT performed directly on the data set. Currently, the periodogram is the most popular spectral estimation method because of its computational superiority over the Blackman-Tukey method.

Let the measured data sequence consist of N data samples $x_0, x_1, \ldots, x_{N-1}$ observed from a wide-sense stationary real or complex valued ergodic and zero mean time series. According to the Blackman-Tukey (1959) method, the PSD of the discrete time series is estimated by

$$\hat{\mathcal{P}}_{BT}(f) = \Delta t \sum_{m=-M}^{M} \hat{R}_{xx}(m)e^{(-j2\pi f m \Delta t)}, \qquad (6.1)$$

where f is the frequency in the range $-1/(2\Delta t) \leq f \leq 1/(2\Delta t)$. The hat denotes variables estimated. The autocorrelation lag estimates $\hat{R}_{xx}(m)$ are computed using the unbiased estimator

141

$$\hat{R}_{xx}(m) = \frac{1}{N-m} \sum_{n=0}^{N-m-1} x_{(n+m)} x_n^* , \qquad \text{and} \qquad (6.2)$$

$$\hat{R}_{xx}(-m) = \hat{R}_{xx}^*(m) , \qquad (6.3)$$

for $m = 0, \ldots, M$, $M \leq N - 1$. The negative estimates are determined in accordance with the conjugate symmetric property of the autocorrelation function of a stationary process. The superscript $*$ denotes complex conjugate.

Assume that the data are available only in a finite time window from $n = 0$ to $n = N - 1$, and the transform is discretized for N values by taking samples at the frequencies $f = m\Delta f$, $m = 0, 1, \ldots, N - 1$ and $\Delta f = 1/N\Delta t$. Then the PSD of the discrete time series is estimated using the periodogram approach:

$$\hat{\mathcal{P}}_m = \frac{1}{N\Delta t} |X_m|^2 , \qquad (6.4)$$

where X_m is the Discrete Fourier Transform (DFT) given by

$$X_m = \Delta t \sum_{n=0}^{N-1} x_n e^{-j2\pi m \Delta f n \Delta t} , \qquad (6.5)$$

$$= \Delta t \sum_{n=0}^{N-1} x_n e^{-j2\pi mn/N} \qquad \text{for } m = 0, \ldots, N - 1 . \qquad (6.6)$$

The Blackman-Tukey is a computationally efficient approach to spectral estimation if autocorrelation lags are windowed. However, this method may result in negative PSD values when some autocorrelation sequence estimates are used. Compared to Blackman-Tukey, the periodogram based on an FFT operation is

142

computationally a more efficient approach and is thus the most widely used spectral estimation method.

Though the Fourier transform-based methods are computationally efficient and produce reasonable results for a large class of signal processes, these techniques suffer from certain inherent limitations, particularly when the finite data sequence is short. Most of these problems are caused by the assumptions made about the data sequence outside the measurement interval (Kay and Marple 1981).

The finite data sequence used in a FFT computation is considered to be obtained by windowing an infinite length sample sequence with a boxcar function. The use of the finite data sequence implicitly assumes that the data beyond this finite sequence is zero, which is usually not the case. The multiplication of the actual time series by a window function means that the overall transform is the convolution of the desired transform with the transform of the window function. If the true power of a signal is concentrated in a narrow bandwidth, this convolution operation will spread that power into adjacent frequency regions. This leads to distortion in the spectrum due to side-lobe leakage and suppression of weak signal main-lobe responses by strong signal side-lobes.

Data windowing is also a fundamental factor that determines the frequency resolution of the periodogram. Frequency resolution in a PSD is always limited by the length of the data sequence regardless of the characteristics of the data or its signal to noise ratio. The convolution of the window transform with that of the actual signal limits the most narrow spectral response of the resultant transform to the main-lobe width of the window transform independent of the data. The resolution of the frequency will approximately be the inverse of the total duration of the finite data sequence $N \Delta t$. Leakage effects due to data windowing can be

143

reduced by the selection of windows with nonuniform weights, such as Hanning or Hamming windows. This reduction in the side-lobes will always broaden the main-lobe of the window transform and consequently reduce the resolution of the spectral estimate.

Even zero-padding the data sequence before transforming will not improve the resolution of the periodogram. Transforming a data set with zeros only serves to interpolate additional PSD values between those that would be obtained with a non-zero-padded transform, within the frequency interval $-1/(2\Delta t) \leq f \leq 1/(2\Delta t)$ (Kay and Marple 1981).

6.3.2 Parametric Methods of Spectral Estimation

In Fourier transform-based approaches, the windowing of data or lag values makes the implicit assumption that the unobserved data or lag values outside the window are zero, which is normally an unrealistic assumption that results in a smeared spectral estimate. In a parametric (also known as model-based approach) to spectral estimation, the unrealistic assumption about the unobserved data or lag values is replaced by a model of the process underlying the data sequence. If an exact or approximate model of the process from which the data sequence is drawn is available, it is possible to obtain a better spectral estimate based on the model by determining the parameters of the model from observations. Thus, a modeling approach to spectral estimation has three main steps:

1. Select a time series model that best describes the process underlying the data sequence.

2. Estimate the model parameters from either the available data samples or autocorrelation lags using computationally efficient and stable algorithms.

144

3. Determine the spectral estimates by substituting the estimated model parameters into the theoretical power spectral density function (PSDF) implied by the model.

The modeling approach to spectral estimation provides more realistic assumptions about the nature of the measured process outside the measured interval, instead of assuming it to be zero or cyclic. This approach eliminates the need for window functions along with their distorting impacts. Thus, this approach may significantly improve a spectral estimate, especially for short data records.

A modeling technique usually provides spectral estimates of higher frequency resolution than those achievable with the traditional Fourier transform based techniques. The degree of improvement in the resolution and the spectral fidelity, if any, are determined by the ability to fit an assumed model to the measured data with a few parameters (Kay and Marple 1981). The following sections describe the three possible time series models and their corresponding theoretical PSDFs.

Most of the deterministic and stochastic discrete-time processes encountered in practice can be approximated by one of three time series models: Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA). In these rational transfer function models, an input sequence $\{v_n\}$ and the output sequence $\{x_n\}$ that models the time series data are related by the following linear difference equations:

$$x_n = \sum_{k=0}^{p} a_k x_{n-k} + v_n \qquad \text{(AR process)}, \qquad (6.7)$$

$$x_n = \sum_{l=0}^{q} b_l v_{n-l} \qquad \text{(MA process)}, \qquad (6.8)$$

145

$$x_n = \sum_{k=1}^{p} a_k x_{n-k} + \sum_{l=0}^{q} b_l v_{n-l} \qquad \text{(ARMA process)}. \qquad (6.9)$$

The system transfer function $H(z)$ between the input v_n and output x_n for AR, MA, and ARMA models are given by

$$H(z) = A(z) \qquad \text{(all-pole model)}, \qquad (6.10)$$

$$H(z) = B(z) \qquad \text{(all-zero model)}, \qquad (6.11)$$

$$H(z) = \frac{B(z)}{A(z)} \qquad \text{(pole-zero model)}. \qquad (6.12)$$

where $A(z)$ and $B(z)$ are the z-transforms of AR and MA processes respectively and are expressed as follows:

$$A(z) = \sum_{m=0}^{p} a_m z^{-m}, \qquad (6.13)$$

$$B(z) = \sum_{m=0}^{q} b_m z^{-m}. \qquad (6.14)$$

$P_x(z)$, the power spectrum of a linear filter's output process is related to $P_v(z)$ the power spectrum of the filter's input stochastic process by (Kay and Marple 1981)

$$P_x(z) = H(z)H^* \left(\frac{1}{z^*} \right) P_v(z), \qquad (6.15)$$

$$= A(z)A^* \left(\frac{1}{z^*}\right) P_v(z) \qquad \text{for all-pole model}, \qquad (6.16)$$

$$= B(z)B^* \left(\frac{1}{z^*}\right) P_v(z) \qquad \text{for all-zero model}, \qquad (6.17)$$

$$= \frac{B(z)B^* (1/z^*)}{A(z)A^* (1/z^*)} P_n(z) \qquad \text{for pole-zero model}. \qquad (6.18)$$

The Eq. (6.18) is generally evaluated along the unit circle, $z = e^{j2\pi f \Delta t}$ for $-1/(2\Delta t) \leq f \leq 1/(2\Delta t)$. If the driving process is assumed to be white-noise with zero mean and variance σ^2, then the PSD of the noise is $\sigma^2 \Delta t$. The Δt factor is included in the expression for power spectral density of the noise so that $P_x(e^{j2\pi f \Delta t})$, when integrated over $-1/(2\Delta t) \leq f \leq 1/(2\Delta t)$, yields the true power of an analog signal. The PSDFs of the AR, MA, and ARMA output processes are given by

$$\mathcal{P}_{AR}(f) = \frac{\sigma^2 \Delta t}{\left| A \left(e^{(j2\pi f \Delta t)} \right) \right|^2} \qquad \text{for the AR model}, \qquad (6.19)$$

$$\mathcal{P}_{MA}(f) = \frac{\sigma^2 \Delta t}{\left| B \left(e^{(j2\pi f \Delta t)} \right) \right|^2} \qquad \text{for the MA model}, \qquad (6.20)$$

$$\mathcal{P}_{ARMA}(f) = \frac{\sigma^2 \Delta t}{\left| B \left(e^{(j2\pi f \Delta t)} \right) / A \left(e^{(j2\pi f \Delta t)} \right) \right|^2} \qquad \text{for the ARMA model}.$$

$$(6.21)$$

From Eqs. (6.19), (6.20), and (6.21) it is clear that specification of the autoregression coefficients $\{a_k\}$, or moving average coefficients $\{b_k\}$, or both $\{a_k\}$ and $\{b_k\}$, and the variance of the driving process σ^2 is equivalent to specifying the spectrum of the process $\{x_n\}$ (Kay and Marple 1981).

147

6.3.3 Selection of Spectral Estimation Method

From the description of spectral estimation methods in the previous section, two main points can be concluded. Nonparametric methods are desirable when computational speed of spectral estimation is critical, and parametric methods are desirable when the resolution of the spectrum is important. Parametric methods will give more accurate spectral estimates than nonparametric methods when data sequences are short.

In a FFT-based method $N/2$ spectral features are necessary to cover the entire spectrum of a signal, where N is the number of points in the data sequence used for computing the spectrum. However, in a parametric method, the number of spectral features required to describe the spectrum of the signal is much smaller than $N/2$. This indicates that parametric methods will provide better spectral resolution than nonparametric methods.

In the present application, sensor signals are sampled for a finite period and the spectral features are to be extracted from this finite data sequence. The spectral features should represent the entire spectrum of a sensor signal with a sharp resolution. In this situation an obvious choice seems to be a parametric method rather than a nonparametric method. With the advent of fast computers, the computational speed required in a parametric method may not be a limitation.

In the present application, implementation of a parametric spectral estimation method for tool wear estimation involves the following steps:

1. Select a time series model that best describes the cutting process as observed by force and vibration signals.

2. Identify the model order and estimate the model parameters from the data sampled from sensor signals.

148

3. Use the vector of model parameters as an input to a neural network to estimate flank wear.

6.4 Time Series Model of Cutting Process

In a parametric approach, the resolution and accuracy of the spectral esti-mates depend on how accurately and precisely the model describes the actual process.. In most of the applications, the process is assumed to be an AR process because it results in a system of linear equations and offers computational ad-vantages. However, the performance of the AR spectral estimator deteriorates as the signal to noise ratio of the process decreases. MA spectral estimators are not commonly used because narrow-band processes usually require too many model coefficients and hence produces poor quality spectral estimates.

The ARMA model generalizes the AR and the MA models. It is the most suitable model for estimating the power spectrum of a process that has both spectral zeros and peaks. The use of either an AR or an MA model to represent such a process often requires a large number of coefficients, resulting in a spectral estimate of poor quality. An ARMA model is also very useful in the estimation of an AR process with additive white noise. Though an ARMA spectral estimator is better than either an AR or an MA estimator, it is not always preferred because of the difficulty in estimating the model parameters. The optimal procedures for estimating ARMA parameters involve iterative optimization techniques which may pose convergence problems and are, therefore, not computationally efficient. To minimize the computational burden, suboptimal procedures of estimating ARMA parameters are introduced. The suboptimal ARMA techniques have been found to provide high resolution estimates, especially for short data records (Salami 1985).

149

The Wold decomposition theorem (Wold 1954, Koopmans 1974) relates the ARMA, MA, and AR models. The theorem asserts that any stationary ARMA or MA process of finite variance can be represented as a unique AR model of possibly infinite order; likewise, any ARMA or AR process can be represented as a MA process of possibly infinite order. This theorem assures that even if a wrong model among these three is chosen, a reasonable approximation may still be obtained by using a high model order (Kay and Marple 1981).

Several methods have been reported which use either residuals or coefficients of AR models of various signals from cutting processes to classify tool wear levels. Liang and Dornfeld (1989) used coefficients of AR models of AE signals to classify tool wear levels. Takata and Sata (1986) used residuals of AR models of cutting-torque signals to detect tool breakage. Yao and Fang (1992) used trivariate AR models of three orthogonal components of the machining force for dispersion analysis to extract features sensitive to the rate of various types of wear in a finish-cut turning process. Dornfeld and DeVaries (1990), and Choi et al. (1990) used coefficients of multivariate AR models of AE, force, and spindle-motor-current signals to classify tool wear levels.

In the present research, coefficients of ARMA models instead of AR models are used for tool wear estimation. The choice of ARMA models is motivated by the stochastic nature of a cutting process. The existence of inevitable nonhomogeneous material properties and other uncertainties add stochastic variations to a cutting process. In addition, sensor signals pick up noise during their transmission through electric cables and instrumentation. These additive stochastic variations and noise can badly affect the performance of an AR based estimator. The reason for this degradation in performance is that the all-pole model assumed in AR spectral

150

analysis is not valid when noise is present. The spectral peaks are broadened and displaced from their true positions. The effects of noise on an AR process is demonstrated in the following analysis. Let y_n denote a noise corrupted AR process, then

$$y_n = x_n + w_n, \tag{6.22}$$

and

$$x_n = \sum_{l=1}^{p} a_l x_{n-l} + v_n, \tag{6.23}$$

where x_n is a pure AR process, and w_n and v_n are uncorrelated zero mean white noise processes of variance σ_w^2 and σ_v^2 respectively. The PSDF of y_n is given by

$$P_y(z) = \frac{\sigma_v^2 \Delta t}{A(z)A * (1/z^*)} + \sigma_w^2 \Delta t, \tag{6.24}$$

$$= \frac{\left[\sigma_v^2 + \sigma_w^2 A(z)A * (1/z^*)\right]}{A(z)A * (1/z^*)}. \tag{6.25}$$

The above equation shows that the PSD of y_n is described better by a pole-zero model rather than by a all-pole model. That is, even though x_n is an AR(p) process, its noise corrupted version y_n is best modeled as an ARMA(p, p) process rather than as an AR(p) process (Pagano 1974). For a cutting process which consists of both deterministic and nondeterministic components, neither an AR nor an MA model will be able to give an appropriate representation of the process from a parameter parsimony point of view. Instead, such a stochastic process can be efficiently represented by an ARMA model.

151

6.5 ARMA Models of Cutting Process

A univariate ARMA model relates a stochastic input sequence to the output observations using a linear difference equation of past values of the signal and the input. A general form of a univariate ARMA model is given by

$$\phi_p(B)x_t = \theta_q(B)v_t, \qquad (6.26)$$

where $\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta_q(B) = 1 - \theta_1(B) - \cdots - \theta_q B^q$. The v_t's are independently, identically, and normally distributed random shocks (or white noise) with zero mean and variance σ_v^2; B is the back-shift operator such that $Bx_t = x_{(t-1)}$; and $x_t = y_t - \eta$ is the deviation of the observation y_t from some convenient location η. Usually η is taken as the mean of the series if the series is stationary (Tiao and Box 1981).

Multivariate ARMA models are useful extension of univariate ARMA models and have the general form of:

$$\Phi_p(B)\mathbf{x}_t = \Theta_q(B)\mathbf{v}_t, \qquad (6.27)$$

where $\Phi_p(B)\mathbf{x}_t = I - \Phi_1 B - \ldots - \Phi_p B^p$ and $\Theta_q(B)\mathbf{v}_t = I - \Theta_1 B - \ldots - \Theta_q B^q$ are matrix polynomials in B, the Φ's and Θ's are $k \times k$ matrices (k is the number of series), and $\mathbf{x}_t = \mathbf{y}_t - \eta$ is the vector of deviations from some origin η. If the series is stationary, η it taken to be the mean of the series. $\{\mathbf{v}_t\}$ with $\mathbf{v}_t = (v_{1t}, \ldots, v_{kt})^T$ is a sequence of random shock vectors identically, independently, and normally distributed with zero mean and covariance matrix Ψ (Tiao and Box 1981).

Development of a univariate or a multivariate ARMA model involves two steps: model identification and parameter estimation. In the first step, model orders p and q are determined from the observed data. Once the model orders p and

152

q are fixed, the model coefficients are estimated from the observed data. Several methods have been developed for both identification and parameter estimation of ARMA models. An extended autocorrelation analysis method (Tsay and Tiao 1984) and a smallest canonical correlation analysis method (Tsay and Tiao 1985) are among the widely applicable ARMA model identification methods. Once the order of an ARMA model is selected, the model parameters are determined by maximizing the likelihood function or an approximate conditional likelihood function. Extensive literature exists on properties of the likelihood function and the various simplifying approximations to this function. Some of these methods are discussed in (Tiao and Box 1981).

In the present work, ARMA models are identified using the smallest canonical correlation analysis method proposed by Tsay and Tiao (1985), and the parameters are estimated using conditional and exact likelihood methods proposed by Box and Jenkins (1970), Wilson (1973), and Hillmer and Tiao (1979). The statistical package SCA (Scientific Computing Associates) which is used in this work provides the computational implementation of these methods.

In this application, it is appropriate to regard sensor signals as stationary processes at different stages of wear development because it takes only a fraction of a second for a set of a few hundred data points to be sampled each time. The coefficients of ARMA models are estimated based on these stationary time series data. Once the coefficients of ARMA models are estimated, the elements of the coefficient vectors ϕ and θ or coefficient matrices Φ and Θ are considered as the representations of sensor measurements. These coefficients describe the instantaneous dynamics of the cutting process as observed by sensor signals.

153

For example, consider a tool wear monitoring scheme in which force and vibration signals are measured as shown in Figure 6.1. Consider the time series data of N points collected at two time instances t_1 and t_2. Let these vectors be represented by $x_1^1 \cdots x_N^1$ and $x_1^2 \cdots x_N^2$ respectively. Let these time series be modeled by multivariate ARMA models $\Phi_p^1(B)\mathbf{x}_t^1 = \Theta_q^1(B)\mathbf{v}_t$ and $\Phi_p^2(B)\mathbf{x}_t^2 = \Theta_q^2(B)\mathbf{v}_t$ respectively. For a given input sequence $\mathbf{v}_t$, the change in output sequences $\mathbf{x}_t^1$ and $\mathbf{x}_t^2$ is reflected by the change in the model coefficients $[\Phi_p^1(B), \Theta_q^1(B)]$ and $(\Phi_p^2(B), \Theta_q^2(B))$. If there is any change in the signal spectral characteristics from time instance t_1 to time instance t_2 due to the progressing tool wear, that change will be reflected in the coefficients of the ARMA models determined at time instances t_1 and t_2.

In this work, individual sensor signals could be modeled using univariate ARMA models or all sensor signals together could be modeled using a multivariate ARMA model. For individual sensors, coefficients of the individual univariate ARMA models are input to separate neural networks. The tool wear estimates given by individual neural networks are integrated using another neural network. Here, sensor data fusion is expected to take place at the neural network stage. In the later case, coefficients of the multivariate ARMA model form an input vector to the neural network. In this case, sensor data fusion takes place at the modeling stage itself. Theoretically, it is difficult to say which method will produce better tool wear estimates. Only the experimental results will give an answer to this question.

6.6 Implementation

As described in section 6.3.2, the process of representing force and vibration

154
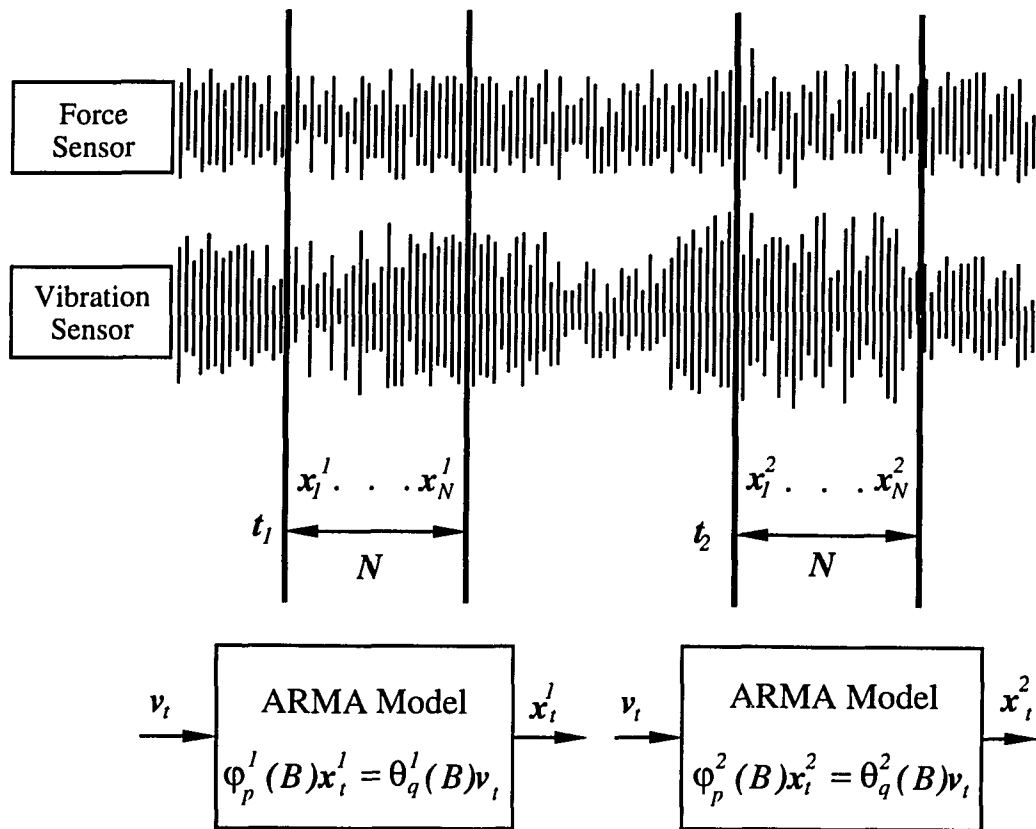
Figure 6.1: Sensor data representation using ARMA models

signals starts with the selection of time series models that best describes those signals. In the present work, this decision is made based on the orders of AR and MA parts of ARMA models. The orders of AR and MA parts of ARMA models for either force or vibration data are determined using smallest canonical correlation analysis method (Tsay and Tiao 1985). The statistical package SCA (Scientific Computing Associates) which implements this method is used to find model orders of sensor data.

155

If the order of the MA part of an ARMA model is zero, then the sensor data is considered to be best described by an AR model. Similarly, if the order of the AR part of an ARMA model is zero, then the sensor data is considered to be best described by an MA model. If both the AR and the MA parts of an ARMA model are non-zero, then the sensor data is considered to be best described by an ARMA model.

Appendix E gives a sample of SCA code for finding the orders of univariate ARMA models of cutting, feed, and radial force signals. Appendix F gives a sample of SCA code for finding the order of a multivariate ARMA model of cutting, feed, and radial force signals. Appendices G and H give samples of SCA code for finding the orders of univariate and multivariate ARMA models of cutting and feed vibration signals.

In the case of univariate ARMA modeling, all three force signals are best described by AR(3) and MA(3). Even in the case of multivariate ARMA modeling, the force signals are best described by AR(3) and MA(3).

In the case of univariate ARMA modeling, vibration signals are best described by AR(8) and MA(2). In the case of multivariate ARMA modeling, vibration signals are best described by AR(6) and MA(4).

Once the order of an ARMA model is selected, the model parameters are determined by maximizing the likelihood function or an approximate conditional likelihood function (Tiao and Box 1981). The SCA package is used to find the ARMA model coefficients both in the univariate and the multivariate cases. Appendices I and J give the sample SCA codes for determining the univariate and the multivariate ARMA model coefficients of force signals. Appendices K and L gives the sample SCA codes for determining the univariate and the multivariate

156

ARMA model coefficients of vibration signals.

Once the ARMA model coefficients are determined, they are input to neural networks to get flank wear estimates. The dimensionality of the input patterns to neural networks is determined by the number of coefficients defining the ARMA models which in turn are determined by the order of the models.

6.7 Conclusions

This chapter discussed the theory and implementation of representing force and vibration signals using univariate and multivariate ARMA models. Once the sensor data are modeled by ARMA models, the coefficients of the models are used as the features of sensor data.

The next step is to relate these ARMA model coefficients to flank wear values. As discussed in chapter 5, a specially designed neural network is used for this purpose. The theory and implementation of this neural network architecture is discussed in Chapter 8. The results of flank wear estimation using force and vibration sensors are presented in Chapter 9.

157

Chapter 7

TOOL WEAR ESTIMATION USING
ACOUSTIC EMISSION MEASUREMENTS

7.1 Introduction

This chapter deals with the representational and analysis issues of Acoustic Emission (AE) signals for flank wear estimation in turning processes. The nature of AE signals and their properties in regard to flank wear monitoring were described in Chapter 4. Chapter 5 dealt with the details of data acquisition and preprocessing of AE signals. This chapter introduces the discrete wavelet transform and discusses how it can be used for representing AE signals.

Spectral analysis of AE signals is more effective than other methods based on AE parameters such as RMS, event count, or count rate for predicting flank wear and assessing tool condition. In the existing methods, the power spectral density of AE signals is computed using Fourier transform-based methods. Fourier transform expands the original function in terms of orthonormal basis functions of sine and cosine waves of infinite duration. Thus, Fourier transform-based methods are good for analyzing long time periodic signals but not for analyzing signals such as AE which have transient components.

Recently, new families of orthonormal basis functions called *wavelets* have been discovered which overcome the problems of the Fourier transform. Unlike the sine and cosine waves of the Fourier transform, wavelets need not have infinite support. They can be nonzero for only a short duration of time. This compact

158

support allows the wavelet transform to translate a time-domain signal into a representation that is localized both in time and frequency domains.

The performance of wavelet transform surpasses that of Fourier transform in analyzing signals that change rapidly, in which high frequency components are studied with sharper time resolution than low frequency components. It analyzes all the signal frequencies and the signal phases. The wavelet coefficients represent the correlations between the wavelet and the signal (McAulay and Li 1992).

Section 7.2 describes the properties of AE signals in turning processes. Section 7.3 discusses why time-frequency analysis of AE signals is more appropriate than pure frequency-domain analysis. Section 7.3 also compares time-frequency analysis through short-time Fourier transform and discrete wavelet transform. Section 7.4 briefly reviews the construction of discrete wavelet transform through multiresolution analysis. Section 7.5 describes the representation of AE signals using discrete wavelet transform coefficients.

7.2 Properties of AE Signals in Turning

Several methods have been investigated using AE signals for flank wear estimation, tool fracture detection, and chip formation detection (continuous or discontinuous). The reviews of these methods are available in Dornfeld (1984), Dan and Mathew (1990), and Chittayil (1994). The primary interest of this section is to highlight some of the features of AE signals to show that time-frequency analysis is more appropriate than a pure frequency domain analysis for extracting tool-wear sensitive features from AE signals.

159

7.2.1 What are AE Signals?

AE signals are the elastic stress waves generated as a result of the rapid release of strain energy within a material due to the rearrangement of its internal structure. These stress waves produce displacement on the surface of the material which can be detected by converting them into electrical signals using piezoelectric transducers (Kannatey-Asibu and Dornfeld 1981).

In metal cutting, AE is the transient elastic energy spontaneously released in materials undergoing deformation or fracture or both. At a microscopic level, AE is related to the grain size, dislocation density, and distribution of second-phase particles in crystalline form (Kannatey-Asibu and Dornfeld 1982).

7.2.2 Source of AE in Turning

In turning processes, AE is generated at different locations in chip-tool work-piece contact region. Several researchers in this area agree on the following AE sources (Kannatey-Asibu and Dornfeld 1981, Moriwaki 1983, Lan and Dornfeld 1986):

1. Deformation in the primary shear zone.

2. Deformation in the secondary shear zone (chip-tool interface zone).

3. Fracture of both the workpiece and the cutting tool.

4. Friction between the chip and the rake face of the tool.

5. Friction between the workpiece and the flank face the tool.

6. The breakage of chips and their impact on the cutting tool or the workpiece.

160

7.2.3 Nature of AE Signals in Turning

The AE generated in turning processes, has two forms: continuous and transient. The continuous form of AE is primarily generated by dislocations in the primary shear zone, dislocations in the secondary shear zone, and rubbing friction at flank wear and crater wear regions. The transient form of AE is generated by discrete events such as chip breaking, chip impact on workpiece or tool, cutting tool fracture, and micro-cracks in either the workpiece or the tool material (Kannatey-Asibu and Dornfeld 1981, Lan and Dornfeld 1986).

In cutting processes, acoustic emission from the friction between the tool and the workpiece can be distinguished from the AE signals generated by the material deformation process at the shear zones. The AE signals due to flank wear can also be distinguished from AE signals due to chip-breaking because the latter ones are typically large amplitude burst signals compared to the former ones (Dornfeld and Lan 1983, Liang and Dornfeld 1989). This distinction between AE signals from different sources makes it possible to use AE signals for tool condition monitoring.

7.2.4 Factors that Influence AE Signals

AE generated from a machining process fundamentally depends on the properties of workpiece and tool materials, applied stress, strain rate, and volume involved in the deformation process (Lan and Naerheim 1985, Lan and Dornfeld 1986, Moriwaki and Tobito 1990).

Lan and Dornfeld (1986) found that the mean RMS energy of AE signals is linearly related to the Brinell hardness of the workpiece material. Kannatey-Asibu and Dornfeld (1981) found that higher AE energy is generated when machining SAE 1018 steel than when machining 6061-T6 Al for the same cutting conditions

161

due to higher strength of the steel. Lan and Naerheim (1985) found that the AE signal amplitude generated by machining Ti-6Al-4V is substantially larger than when machining SAE 4340 steel (Lan and Naerheim 1985).

Kannatey-Asibu and Dornfeld (1981) observed that AE is very sensitive to cutting speed and rake angle in orthogonal cutting. Lan and Naerheim (1985) observed similar results in a three-dimensional turning test. These results indicate that AE amplitude is very sensitive to the strain rate which is affected by both the cutting speed and rake angle. Lan and Dornfeld (1986) noticed that the energy of AE signals decreases with increasing tool rake angle due to decreasing energy consumption at both primary and secondary deformation zones. The mean RMS energy of AE signals generated during normal machining is not sensitive to the variation in feed rate and depth of cut (Inasaki and Yonetsu 1981, Lan and Dornfeld 1986). Lan and Naerheim (1985) found that depth of cut does affect AE signals in a finish cut. The depth of cut mainly affects AE through change in volume of material involved in the deformation process.

Increasing flank wear land increases the portion of AE that is generated due to friction between flank face and the workpiece. The AE generated from primary shearing exhibits different signal characteristics compared to the AE generated from friction because of the difference in the mechanisms by which the AE is produced. As a result, the characteristics of AE signals change as the flank wear land progresses (Liang and Dornfeld 1989).

Crater wear influences AE by causing a change in the effective rake angle. As the crater wear progresses, the effective rake angle increases. The energy of AE signals decreases with increasing rake angle (Liang and Dornfeld 1989).

162

It is found that AE from chip fracture contributes significantly to the signal generated during machining. A periodic burst of activity is superimposed on the continuous AE of chips during machining. Each burst event due to each chip fracture appears as a spike in the RMS measurement (Lan and Dornfeld 1986).

A significant level of AE is generated by tool fracture caused by excessive metal removal rates, inclusions in the workpiece, and deterioration of the tool geometry. An AE burst occurs during the formation and propagation of a crack in the cutting tool. The amplitude of the RMS signal energy at the tool fracture is dependent upon the area of the fracture (Lan and Dornfeld 1984, and Dornfeld 1984).

7.2.5 AE Parameters Sensitive to Tool Wear

AE signals are usually picked up by a piezoelectric sensor, amplified by an amplifier, and band-pass filtered in the frequency range of 100 kHz to 1 MHz. The normal AE parameters that are used for flank wear estimation and tool failure detection are AE RMS, event count, count rate, AE-mode, amplitude of dominant-frequencies, and power of dominant-frequency bands.

AE RMS: The Root Mean Square (RMS) value of an AC voltage is equal to the DC voltage that, if applied to the same electric circuit for the same period of time, produces the same energy dispersion as the AC voltage. This RMS value is given by the following equation (Teti and Dornfeld 1989).

$$V_{RMS} = \left[\frac{1}{\Delta T} \int_0^{\Delta T} V^2(t)dt \right]^{\frac{1}{2}} . \tag{7.1}$$

A constant and relatively uniform level of AE RMS corresponds to continuous chip formation. A constant level AE RMS with rather regular spikes corresponds to discontinuous chip cutting. A relatively noisy AE RMS with a sudden increase in level corresponds to chip congestion or other undesirable chip variations during cutting (Lan and Dornfeld 1986).

Dornfeld (1984) observed that AE RMS increases with increasing flank wear up to 0.015 in. (0.381 mm). For flank wear between 0.015 in. (0.381 mm) and 0.022 in. (0.559) AE RMS values do not increase with increasing flank wear due to the counter effect of crater wear on AE. Kakino (1980), Inasaki and Yonetsu (1981), and Lan and Dornfeld (1982) also found that AE RMS increases in proportion to the amount of flank wear.

Kannatey-Asibu and Dornfeld (1982) found that skewness and kurtosis of an assumed Beta distribution for AE RMS signals are sensitive to progressive flank wear. The skewness increases with increasing flank wear while the kurtosis decreases with the cutting time. Dornfeld (1984) observed that the skewness and the kurtosis of AE RMS signals are also affected by the crater wear but only at the severe levels of crater wear.

Liang and Dornfeld (1989) found that the parameters of autoregressive models of AE RMS signals are sensitive to flank wear.

Lan and Dornfeld (1984) observed that significant AE bursts are generated at the instance of tool breakage and the RMS level of the bursts depends on the fracture area: the larger the tool fracture area the stronger the AE energy.

AE-Mode: Inasaki et al. (1987) found that the AE Mode, defined as the mode of the probability density function of the AE signal, is very sensitive to tool chipping.

164

Event Count and Ring Down Count: The AE event count is the number of AE bursts whose envelope amplitude exceeds a certain threshold. Ring down count is the number of times the AE signal crosses the threshold. Figure 7.1 schematically shows these definitions (Teti and Dornfeld 1989).

Dornfeld and Lan (1983) found that the event count of AE signals shows excellent correlation to the chip-breaking frequency and offers a reliable measure of chip-breaking activity.

Iwata and Moriwaki (1977) and Moriwaki (1983) observed a correlation between the amount of flank wear and the total number of AE event counts.

Teti and Micheletti (1989) reported that both the total AE event count and the total ring down count increase with increasing flank wear but this trend is negatively influenced by the development of crater wear.

Frequency Analysis: Several researchers observed that the amplitude of certain frequencies or the power in certain frequency bands increases with increasing flank wear. These frequencies and frequency bands depend on tool and workpiece materials, AE sensor location, and tooling set up.

Inasaki and Yonetsu (1981) observed that during the cutting of carbon steels, the amplitudes of AE signals near 120, 170, and 210 kHz frequencies increase with increasing flank wear height and cutting speed but are hardly affected by the changes in feed rate and depth of cut. Further, a stepwise increase in the AE signal energy was observed with the increasing contact area after tool fracture.

Emel and Kannatey-Asibu (1988) found that progressive tool wear affects the AE power in the frequency range of 400–700 kHz and catastrophic tool failure boosts the AE power in the frequency range of 100 kHz to 1 MHz.

165
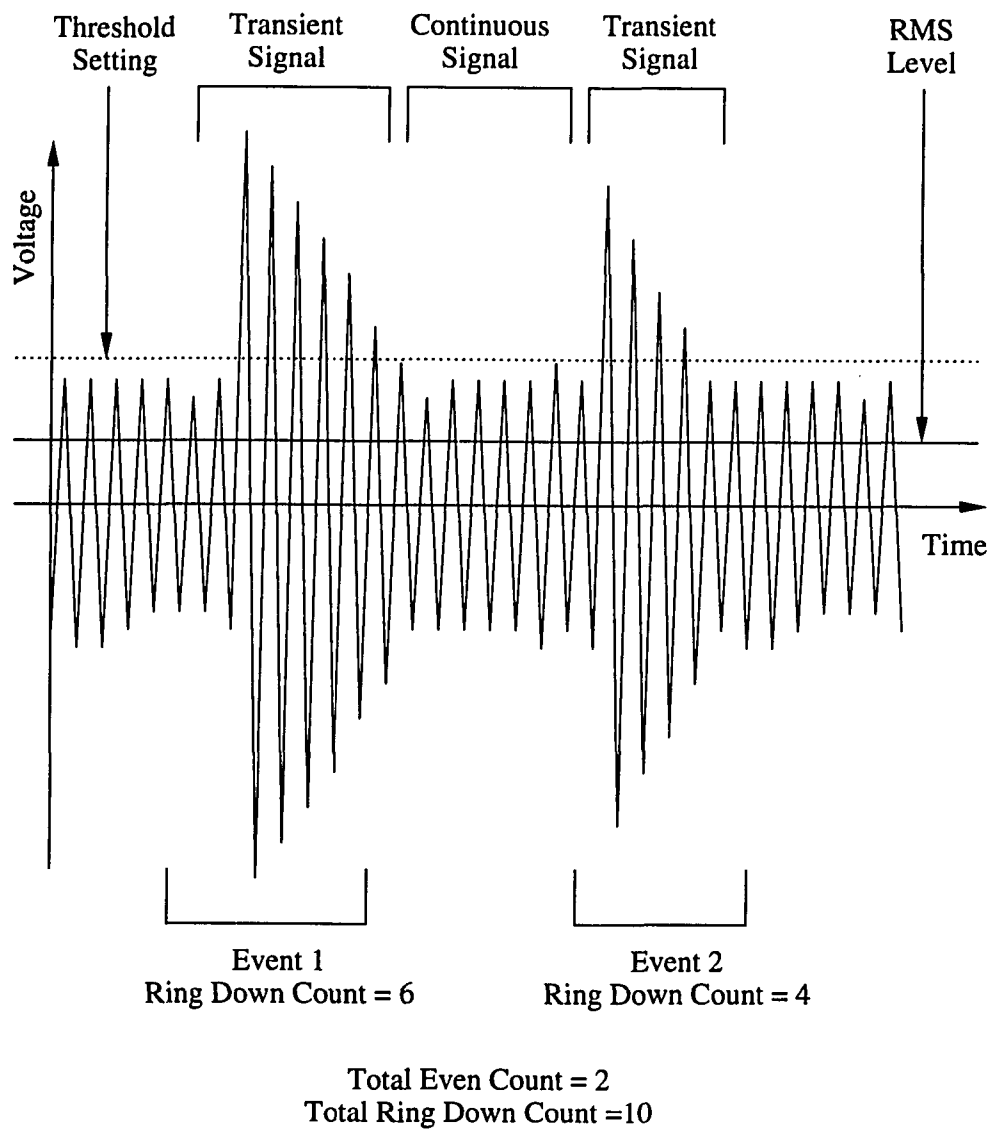
Figure 7.1: Schematic diagram of an AE signal

166

Du and Yan (1991) found that a peak around 100 kHz is always present in the Time-Frequency-Distribution. While cutting with a worn tool or using large feed and large depth of cut, a peak around 300 kHz occurs. They inferred that the frequency around 100 kHz is related to AE release from shear zones, and the frequency around 300 kHz is related to AE release from friction between tool and workpiece.

Iwata and Moriwaki (1977) found that the spectrum of AE signals in the frequency range of 100 to 250 kHz increases with flank wear and AE measurements are hardly affected by ambient vibrations and noise. However, this increase in the frequency spectrum reaches a saturation point at the advanced tool wear stage.

7.2.6 AE Signals for Tool Wear Estimation: A Critical View

AE provides three major advantages for flank wear estimation:

1. The frequency range of an AE signal is much higher than that of the machine vibrations and environmental noise. This makes it possible to obtain a relatively uncontaminated signal by using a high-pass filter.

2. AE is very quick in reflecting the changes in the tool condition.

3. AE can be measured by simply mounting a piezoelectric transducer on the toolholder. This helps to monitor the tool condition without hindering the cutting operation.

On the other hand, it is difficult to extract flank wear-sensitive features from the AE signals for the following reasons:

1. Most of the tool wear detection indices such as AE RMS, event count, count rate, and AE-mode depend on the rate of energy released during cutting. The energy rate depends not only on the level of tool wear but also on the cutting

167

conditions such as cutting speed, feed, and depth of cut. This reduces the effectiveness of flank wear detection indices to predict flank wear.

2. The spectral features of AE signals are sensitive to the tool wear condition, but the frequency analysis of AE signals is difficult due to signal coloring effects caused by the propagation media, sensor frequency response, and instrumentation system function. In their original form, AE waves are basically sinusoidal in nature. During the course of their propagation, they often undergo considerable changes due to scattering by structural defects, multiple reflections at interfaces, and refraction where there is a medium change along the travel path. All these factors considerably modify the waveform by causing changes in phase, amplitude attenuation, and wave repetitions through reflections (Kannatey-Asibu and Dornfeld 1981, Liang and Dornfeld 1989).

3. In a cutting process, the AE characteristics are not only affected by tool condition but also by other factors such as temperature, built-up edge, and effective rake angle. For this reason, crater wear, due to its influence on the tool effective rake angle, tends to reduce the sensitivity of AE energy content to progressive flank wear.

Despite these difficulties, among various approaches taken to analyze AE signals, the power spectral density of AE signals has been found to be the most informative for monitoring progressive tool wear and tool breakage in turning processes (Emel and Kannatey-Asibu 1986, Kannatey-Asibu and Emel 1987). In earlier methods, the power spectral density of AE signals is computed based on the FFT technique (Du and Yan 1991). However, the FFT technique has some drawbacks in estimating power spectral density of AE signals, primarily because AE signals are non-stationary. The following section discusses these drawbacks and

168

also describes how those problems can be overcome using time-frequency analysis through wavelet transforms.

7.3 Time-Frequency Analysis of AE Signals

As shown in Figure 7.2, an AE signal has two parts: a continuous part and a transient part. This figure is generated from the data sampled at 1 MHz from the actual turning operation on AISI 6150 steel workpiece with K68 (Kennametal) grade carbide insert at 130 sfpm (0.6604 m/s) speed, 0.0112 ipr (0.2845 mm/rev) feed, and 0.05 in. (1.27 mm) depth of cut. A continuous part of the signal is contributed by deformation of workpiece material in the shear zones and friction in the tool-workpiece and chip-tool contact regions. Superimposed on the continuous AE signal, the transient part of the signal is generated by micro-cracks of the crystal structure of the workpiece material, non-homogeneity of the workpiece material, and chip breakage. The AE pulses generated during the cutting operation are non-stationary and may possess a different magnitude, damping, frequency, and phase (Ono et al. 1971, Du and Yan 1991).

A stationary random process is one whose statistical properties are time invariant. This implies that the first probability density function for the process, $f(x_1, t_1)$, is independent of the time of observation t_1. Then all the moments of this distribution, such as $E[x(t_1)]$ and $E[x^2(t_1)]$, are also independent of time. The second probability density function is not dependent on the absolute observation time instances t_1 and t_2, but it is dependent on τ the difference between t_1 and t_2. The autocorrelation function of the stationary process has the following properties (Gelb 1974):

169
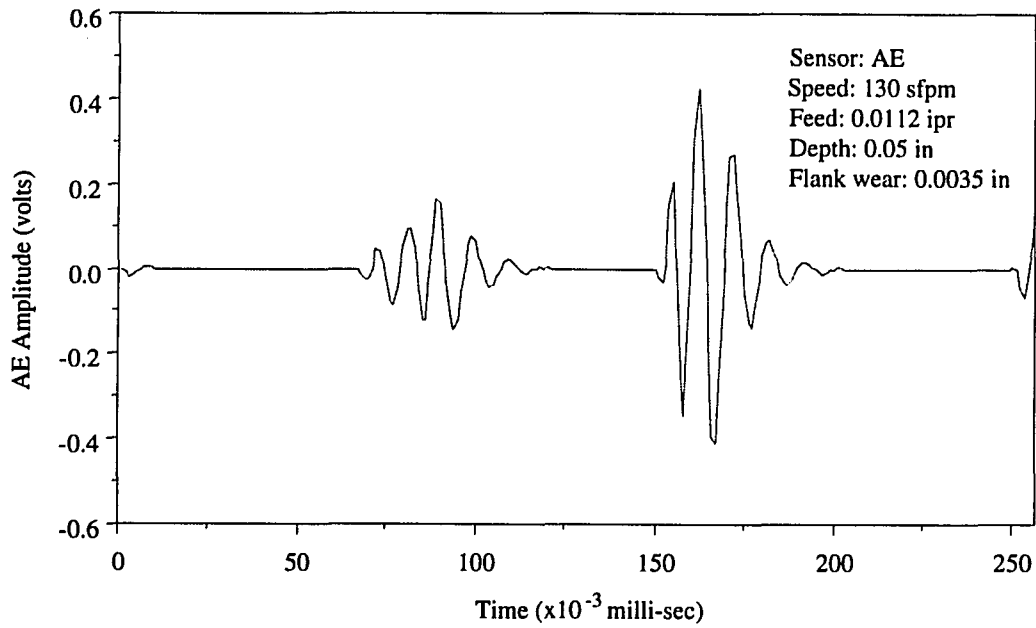
Figure 7.2: A record of AE signal

$$\varphi_{xx}(\tau) = E[x(t_1)x(t_1 + \tau)], \qquad (7.2)$$

$$\varphi_{xx}(0) = E[x^2], \qquad (7.3)$$

$$\varphi_{xx}(-\tau) = \varphi_{xx}(\tau), \qquad (7.4)$$

$$\varphi_{xx}(0) \geq |\varphi_{xx}(\tau)|. \qquad (7.5)$$

For a stationary signal $x(t)$, the frequency information can be obtained by Fourier transform

$$\mathcal{F}(f) = \int_{-\infty}^{+\infty} x(t)e^{-2j\pi ft}dt. \qquad (7.6)$$

170

The analysis coefficients $\mathcal{F}(f)$ define the notion of global frequency f in a signal. They are computed as inner products of the signal with sine wave basis functions of infinite duration. As a result, Fourier analysis works well if $x(t)$ is composed of a few stationary components, but any abrupt change in time in the signal is spread out over the whole frequency axis in $\mathcal{F}(f)$ (Rioul and Vetterli 1991).

Since AE signals are non-stationary and contain both continuous and transient parts, the conventional FFT based methods have two major difficulties in analyzing AE signals:

1. The Fourier transform works under the assumption that the original time-domain signal is periodic in nature. Since AE signals are nonperiodic and have abrupt bursts, the power spectral density estimates computed using FFT techniques will not be accurate.

2. The event count and count rate of AE signals carry some information related to tool condition. FFT methods measure the power content of the transient events in the signal but fails to retain the information related to their occurrence in time, because the Fourier transform of a signal does not convey any information pertaining to translation of transient events over the time scale.

A global representation like the Fourier transform fails to represent information that changes its character unpredictably during the course of the signal. For example, the Fourier representation of a piece of music will tell what notes (or frequencies) were played somewhere within that piece of music but it will fail to tell the timing of those notes. A better representation, like the musical score, tells what notes were played when.

171

To represent signals which have both continuous and transient parts, a time-frequency representation method is necessary. The time-frequency representation of a signal can be achieved through the short-time Fourier transform or the wavelet transform. The next section discusses the properties of these two transforms.

7.3.1 Short-Time Fourier Transform

As mentioned in the previous section, the Fourier transform analyzes a signal in terms of frequency components over the entire time domain. In this global representation, time localization is lost. In order to achieve time localization, conventionally the Short-Time Fourier Transform (STFT) (also known as Gabor transform) is used to compute a local spectrum. The STFT uses a window function to restrict the transform length to achieve better time localization. The coefficients of the STFT measure the amplitude of the sinusoidal wave components of the frequencies locally, around a specified time. If a signal $f(t)$ is considered stationary over the width of a window $g(t)$ centered at time location τ, the STFT of the windowed signal $f(t)\overline{g(t-\tau)}$ is given by

$$STFT(\omega,\tau) = \int f(t)\overline{g(t-\tau)}e^{-i\omega t}dt. \qquad (7.7)$$

As the window moves to different positions in the signal, the STFT gives the spectra of the signal at those positions. The STFT maps the signal $f(t)$ onto a two-dimensional function in the time-frequency plane (ω,τ).

The performance of the STFT critically depends on the selection of the window $g(t)$. Multiplying the signal $f(t)$ by the window function $g(t)$ results in the signal spectrum convoluted with the spectrum of the window. The performance of the STFT method is limited by time and the frequency resolutions of the window

172

$g(t)$. Let Δt and Δf be the time and the frequency resolutions of the window $g(t)$ respectively: any two sinusoids in the frequency domain will be discriminated only if they are more than Δf apart, and any two pulses in the time domain will be discriminated only if they are more than Δt apart. According to the uncertainty principle or Heisenberg inequality, time and frequency resolutions are bounded by (Rioul and Vetterli 1991)

$$\Delta t \, \Delta f \geq \frac{1}{4\pi}. \tag{7.8}$$

Typically Δt and Δf should be as small as possible, but the Heisenberg inequality sets the fundamental trade-off between time and frequency resolutions. In STFT, a Gaussian function is usually used as the window because it meets the bound in Eq. (7.8) with equality, thus minimizing the uncertainty. As the window gets narrower in the time domain, its spectrum gets wider in the frequency domain and *vice versa*. If the window size is determined based on the lowest frequency, the localization of high frequency components will be lost. If the window size is selected to localize high frequencies, the accuracy of low frequency estimates will be lost.

Once a window is selected, both Δt and Δf remain the same over the entire time-frequency plane. The coverage of the time-frequency plane by a STFT is shown in Figure 7.3. The application of constant time and frequency resolutions poses a limitation in analyzing signals composed of small bursts associated with long quasi-stationary components. Each type of signal component can be analyzed with either good time resolution or good frequency resolution, but not both (Rioul and Vetterli 1991).

173
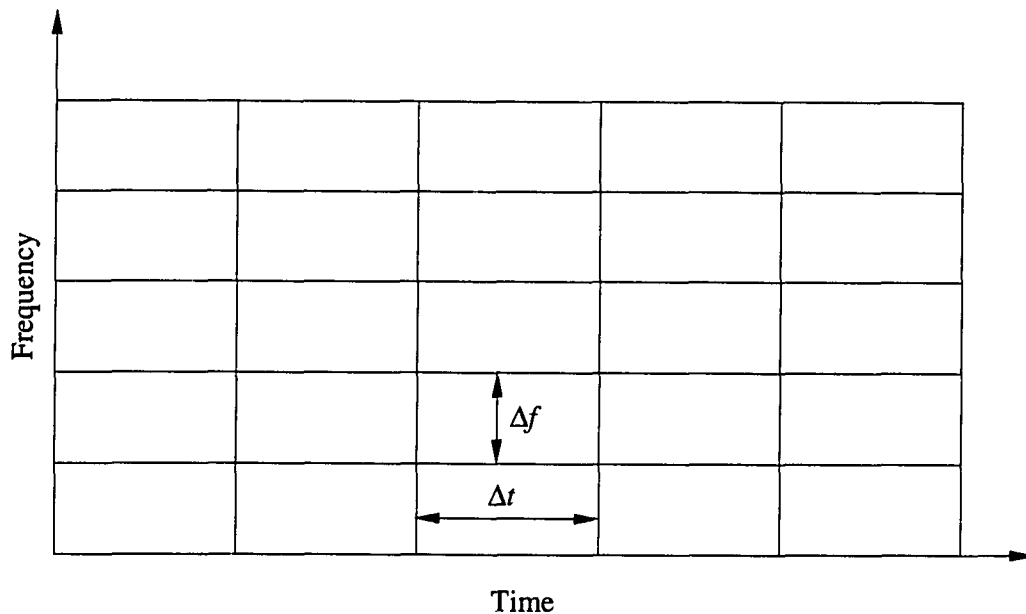
Figure 7.3: The time-frequency plane covered by STFT

The frequency is the measure of cycles per unit signal length. High frequency components take shorter signal length, and low frequency components take longer signal length. The inverse relationship between the frequency level and the window length necessitates a short window for high frequency components and a wider window for low frequency components.

This time-frequency resolution problem is resolved to some extent in a wavelet transform through multiresolution analysis. In multiresolution analysis, time and frequency resolutions are varied according to the frequency of signal components. As shown in Figure 7.4, the time-frequency plane of a wavelet transform adapts the window size according to the frequency level. Without violating the Heisenberg inequality, the time resolution becomes better at high frequencies, and the frequency resolution becomes better at low frequencies.

174
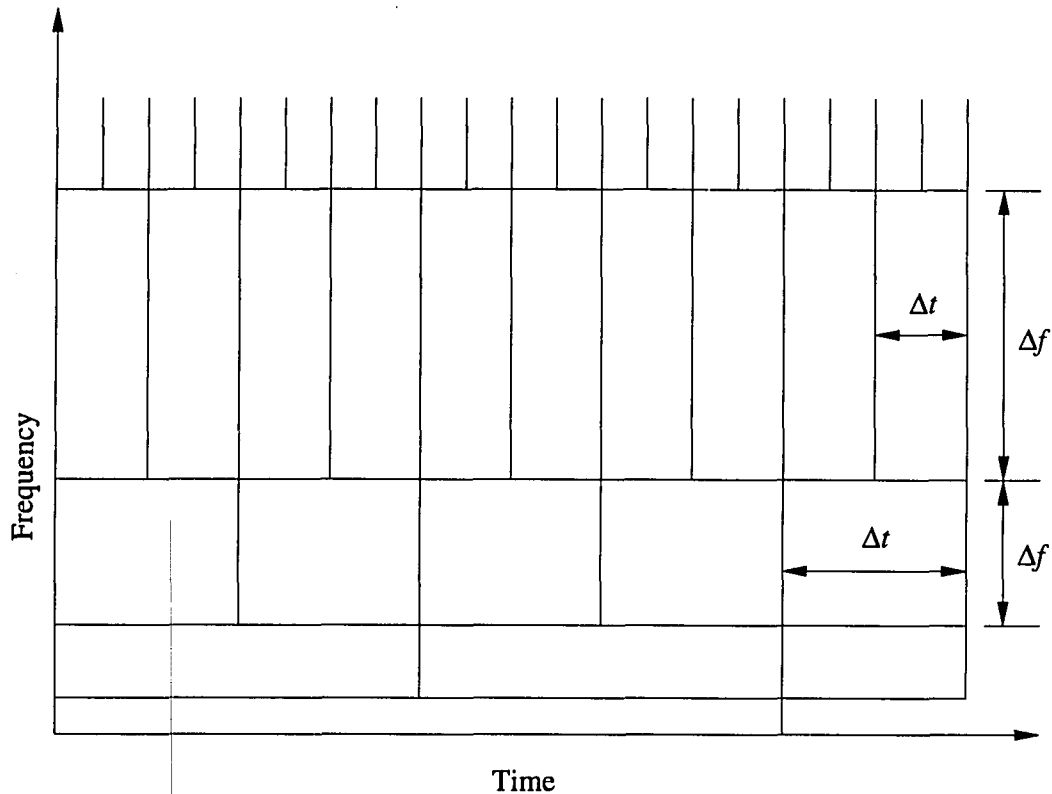
Figure 7.4: The time-frequency plane covered by DWT

This multiresolution analysis offers advantages in representing signals, such as AE, which contain high frequency bursts of short duration and low frequency components of longer duration. The closely located high frequency bursts are analyzed with a sharper time resolution, while low-frequency long-duration signals are analyzed with a better frequency resolution. These are the attractive time-frequency properties of a wavelet transform for analyzing AE signals.

7.3.2 Wavelet Transform

A wavelet transform provides time-frequency analysis similar to that of a

175

STFT. A wavelet transform and the STFT primarily differ in the decomposition functions they use. Wavelet transform decomposes the signal into a family of functions which are obtained by dilation and translation of a unique function called the mother wavelet $\psi(t)$. The family of functions derived by scaling and shifting the fundamental wavelet ψ has following form:

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi \left(\frac{t-b}{a} \right) , \qquad (7.9)$$

where a and b are the dilation and translation parameters of the mother wavelet function. The continuous wavelet transform is defined by the equation:

$$W f(a,b) = |a|^{-\frac{1}{2}} \int f(t) \psi \left(\frac{t-b}{a} \right) dt , \qquad (7.10)$$

$$= \langle \psi_{a,b}, f \rangle , \qquad (7.11)$$

assuming that ψ satisfies $\int \psi dt = 0$.

The dilation and translation parameters (a and b) may vary over either a continuous or a discrete set. As the dilation parameter a changes, the $\psi_{a,\bullet}(t)$ covers different frequency ranges. Larger values of a correspond to smaller frequencies, and smaller values of a corresponds to higher frequencies. That means that when it is dilated, ψ accesses lower frequency information; and when it is contracted, it accesses higher frequency information. On the other hand, the translation parameter b controls the time localization center of ψ. Each $\psi_{\bullet,b}$ is localized around $t = b$. The translation parameter b is analogous to window center τ in a STFT. The following are the similarities and differences between the STFT and the wavelet transform:

176

1. Both the STFT and the wavelet transform take the inner products of $f(t)$ with a family of functions indexed by

$$g_{\omega,\tau}(t) = e^{i\omega t}, \qquad \text{and} \tag{7.12}$$

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi \left(\frac{t-b}{a} \right). \tag{7.13}$$

2. The functions $g_{\omega,\tau}$ all consist of the same envelope function g, translated to the proper time location τ, and filled in with high frequency oscillations. All the $g_{\omega,\tau}$, regardless of the value of ω, have the same width. In contrast, the $\psi_{a,b}$ have time-widths adapted to their frequency: high frequency $\psi_{a,b}$ are narrow while low frequency $\psi_{a,b}$ are much broader. As a result, the wavelet transform is better able than the STFT to zoom in on very short-lived high-frequency phenomena, such as transients in a signal.

The wavelet transform in Eq. (7.10) is continuous if the parameters a and b vary continuously on their range $\mathbf{R}^* \times \mathbf{R}$ (where $\mathbf{R}^* = \mathbf{R}\backslash\{0\}$). Discrete type wavelet transforms can be formulated by restricting the values of the parameters a and b to discrete values: $a = a_0^m$ and $b = nb_0 a_0^m$, $a_0 > 0$, $b_0 > 0$, $m \in \mathbf{Z}$ and $n \in \mathbf{Z}$. Then the family of discrete wavelets and the corresponding discrete type transform are

$$\psi_{m,n}(t) = a_0^{-m/2} \psi \left(a_0^{-m} t - nb_0 \right), \tag{7.14}$$

$$Wf(m,n) = a_0^{-m/2} \int f(t) \psi \left(a_0^{-m} t - nb_0 \right), \tag{7.15}$$

177

$$= \langle \psi_{m,n}, f \rangle \, , \tag{7.16}$$

assuming that ψ satisfies the admissibility condition $\int \psi dt = 0$.

The selection $b = nb_0 a_0^m$ helps to adapt the translation proportional to the dilation rate $a = a_0^m$. For a large positive value of m, $\psi_{m,n}$ is much broader, and accordingly larger translation steps $b_0 a_0^m$ are taken. For a large negative value of m, $\psi_{m,n}$ is narrow, and accordingly shorter translation steps are taken.

The discrete type wavelet transforms can be constructed from either redundant discrete systems (or wavelet frames) or orthonormal bases of wavelets. In the present application the discrete wavelet transform is formulated based on the orthonormal wavelet bases. In rest of this chapter these wavelets are referred to by Discrete Wavelet Transform (DWT). The next section reviews the mathematics behind the DWT. The treatment of DWT in the next section is very brief and is mostly restricted to the concepts that are important to the present problem. The details of the DWT can be found in (Daubechies 1988, 1990, 1992, Mallat 1989).

7.4 Discrete Wavelet Transform

The $\psi_{m,n}$ constitute an orthonormal basis for $L^2(\mathbf{R})$ for some very special choices of ψ and a_0, b_0. In particular, if $a_0 = 2$, $b_0 = 1$, there exists ψ, with good time-frequency localization properties, such that the family of functions

$$\psi_{m,n}(t) = 2^{-m/2} \psi \left(2^{-m} x - n \right) \tag{7.17}$$

constitutes an orthonormal basis for $L^2(\mathbf{R})$ (Daubechies 1992). A systematic way called multiresolution analysis is used to construct the orthonormal basis $\psi_{m,n}$

178

and to show that a function $f \in L^2(\mathbf{R})$ can be approximated to an arbitrarily small precision by a finite linear combination of $\psi_{m,n}$.

7.4.1 Multiresolution Analysis

A multiresolution analysis of $L^2(\mathbf{R})$ is a sequence of closed subspaces $\cdots, V_{-1}, V_0, V_1, V_2, \cdots$ with the following properties:

1. $V_j \subset V_{j+1}, \ j \in \mathbf{Z}$.

2. $\bigcup_{j=-\infty}^{\infty} V_j$ is dense in $L^2(\mathbf{R})$ and $\bigcap_{j=-\infty}^{\infty} V_j = \{0\}$.

3. $f(t) \in V_j \iff f(2t) \in V_{j+1}$.

4. $f(t) \in V_0 \iff f(t-k) \in V_0$ for all $k \in \mathbf{Z}$.

5. There exists a function $\phi \in V_0$ such that $\phi_{0,k \in \mathbf{Z}}(t) = \phi(t-k)$ constitutes an orthonormal basis for V_0.

Properties 1 and 2 make a multiresolution analysis of $L^2(\mathbf{R})$ a sequence of nested subspaces, $\{V_j\}_{j \in \mathbf{Z}}$, whose limit is dense in $L^2(\mathbf{R})$. Let W_j be the orthogonal complement of V_j in V_{j+1}, then for all $j \in \mathbf{Z}$:

$$V_{j+1} = V_j \bigoplus W_j, \tag{7.18}$$

$$L^2(\mathbf{R}) = \bigoplus_{j=-\infty}^{+\infty} W_j, \tag{7.19}$$

$$L^2(\mathbf{R}) = V_0 \bigotimes_{j=0}^{+\infty} W_j. \tag{7.20}$$

The relationship between subspaces in multiresolution analysis of $L^2(\mathbf{R})$ is shown in Figure 7.5.

179
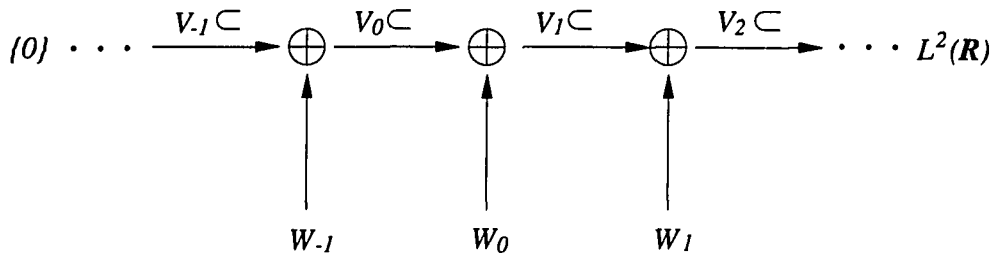
Figure 7.5: Subspaces in multiresolution analysis of L^2 *(R)*

Property 3 defines that a function $f(t)$ is a member of V_j if and only if $f(2t)$, the contraction $f(t)$, is a member of V_{j+1}.

Property 4 defines that a function $f(t)$ is a member of V_j if and only if all translates of $f(t)$ by an integer $k \in \mathbf{Z}$ are also members of V_j.

Property 5 indicates that there exists a function ϕ whose translations by integers $k \in \mathbf{Z}$ form an orthonormal basis for V_0. Let

$$\phi_{j,k} \stackrel{\triangle}{=} 2^{j/2}\phi(2^j x - k), \qquad j, k \in \mathbf{Z}. \tag{7.21}$$

From the properties 1 through 5 it follows that $\{\phi_{j,k}(t)\}_{k \in \mathbf{Z}}$ is an orthonormal basis V_j. $\phi(t)$ is called *scaling function* corresponding to the multiresolution analysis. Since $V_j \subset V_{j+1}$ for all j, any basis element of V_j for $j < 0$ can be expressed as a linear combination of basis elements in V_0:

$$\phi\left(\frac{t}{2}\right) = \sum_{k \in \mathbf{Z}} h_k \phi(t - k), \qquad h_k \stackrel{\triangle}{=} \langle \phi(t/2), \phi(t) \rangle. \tag{7.22}$$

When $\psi_{j,k}$ is set to

$$\psi_{j,k} \stackrel{\triangle}{=} 2^{j/2}\psi\left(2^j t - k\right), \qquad j, k \in \mathbf{Z}, \tag{7.23}$$

180

and ψ is defined as

$$\psi\left(\frac{x}{2}\right) = \sum_k g_k \phi(t-k), \qquad g_k = (-1)^k \bar{h}_{1-k}, \tag{7.24}$$

it follows that $\{\psi_{jk}\}_{k\in\mathbf{Z}}$ is an orthonormal basis for W_j. The function $\psi(t)$ is called the *mother wavelet* associated with the multiresolution analysis.

From Eq. (7.20), a function $f \in L^2(\mathbf{R})$ can be represented by its component at *resolution level* zero and its projection onto W_j, $j \geq 0$, as follows:

$$f(t) = \sum_{k\in\mathbf{Z}} c_{0,k}\phi_{0,k}(t) + \sum_{j=0}^{\infty} \sum_{k\in\mathbf{Z}} d_{j,k}\psi_{j,k}(t), \tag{7.25}$$

where

$$c_{0,k} \stackrel{\triangle}{=} \langle f, \phi_{0,k}(t) \rangle, \qquad d_{j,k} \stackrel{\triangle}{=} \langle f, \psi_{j,k}(t) \rangle. \tag{7.26}$$

Let the projection of a function f onto V_j be denoted by f_j. By recursively applying Eq. (7.18), f_j can be computed as follows:

$$f_j(t) = \sum_{k\in\mathbf{Z}} c_{0,k}\phi_{0,k}(t) + \sum_{m=0}^{j-1} \sum_{k\in\mathbf{Z}} d_{m,k}\phi_{m,k}(t), \qquad j > 0. \tag{7.27}$$

Usually f_j is referred to by *approximation of f at level j*. The projection of f onto W_j is the complementary part to f_j; it is referred to by *detail signal of f at level j*. Let the detail signal of f at level j be denoted by $\hat{f}_j$. Then

$$\hat{f}_j(t) = \sum_{k\in\mathbf{Z}} b_{j,k}\psi_{j,k}(t), \tag{7.28}$$

$$= f_{j+1} - f_j. \tag{7.29}$$

181

In this multiresolution analysis, every function $f \in L^2(\mathbf{R})$ can be approximated as closely as is desired by an $f_j \in V_j$, for some $j \in \mathbf{Z}$. Since $V_j = V_{j-1} + W_{j-1}$ for any $j \in \mathbf{Z}$, f has a unique decomposition:

$$f_j = f_{j-1} + \hat{f}_{j-1}, \tag{7.30}$$

where $f_{j-1} \in V_{j-1}$ and $g_{j-1} \in W_{j-1}$. By repeating this process

$$f_j = \hat{f}_{j-1} + \hat{f}_{j-2} + \cdots + \hat{f}_{j-M} + f_{j-M}. \tag{7.31}$$

The decomposition of a function $f \in L^2(\mathbf{R})$ as expressed above is called *wavelet decomposition*. In this decomposition, M is so chosen that f_{j-M} is sufficiently blurred. The blur is measured in terms of frequency or number of cycles per unit length of f_j (Chui 1992).

7.4.2 Implementation of Discrete Wavelet Transform

The implementation of the DWT begins with the construction of scaling function ϕ and the corresponding mother wavelet function ψ in Eq. (7.21) and Eq. (7.24) respectively. Depending on the construction of the pair of functions ϕ and ψ, different multiresolution analyses are possible. In the present work the DWT algorithm is based on the orthonormal bases of compactly supported wavelets introduced by Daubechies (1988). In the compactly-supported wavelets, the scaling function ϕ and the mother wavelet ψ are defined as:

$$\phi(t) = \sqrt{2} \sum_{k=0}^{2N-1} h_k \phi(2x - k), \qquad N > 1, \tag{7.32}$$

182

$$\psi(t) = \sqrt{2} \sum_{k=0}^{2N-1} g_k \phi(2x - k), \qquad g_k = (-1)^k \bar{h}_{-k+1} \text{ and } N > 1. \quad (7.33)$$

The coefficients h_k satisfy the following admissibility conditions:

1. $\sum_{k=0}^{2N-1} |h_k||k|^\epsilon < \infty$ for some $\epsilon > 0$.

2. $\sum_{k=0}^{2N-1} h_{k-2m} h_{k-2n} = \delta_{mn}$, where δ_{mn} is Kronecker delta.

3. $\sum_{k=0}^{2N-1} = \sqrt{2}$.

Once a wavelet system is created, it can be used to expand a function $f(t)$ in terms of the basis functions shown in Eq. (7.27).

The implementation of the DWT begins with the selection of record length (2^M) of the signal. Let the discrete points in the signal be denoted by y_k. These y_k are viewed as the inner products of $\phi_{0,p}$ and the signal $f(t)$. That is, the sample points are considered an approximation to $c_{0,k}$ coefficients in Eq. (7.27). This allows $c_{j,k}$ and $d_{j,k}$ terms to be calculated by direct convolution of $f(t)$ samples with the coefficients h_k and g_k respectively.

The implementation of a pyramid algorithm for the DWT is shown in Figure 7.6. It is implemented with a pair of Finite Impulse Response (FIR) filters called a *Quadrature Mirror Filter* (QMF) pair. A FIR filter performs the dot product between the filter coefficients and the discrete samples in the tapped delay line of the filter. Passing a set of discrete samples, representing a signal, through a FIR filter causes discrete convolution of the signal with filter's coefficients.

The outputs of the QMF filter pair are decimated (or desampled) by a factor of two; that is, every other output sample of the filter is kept and the others are discarded. The low-frequency filter output is fed into another identical QMF filter pair.

183
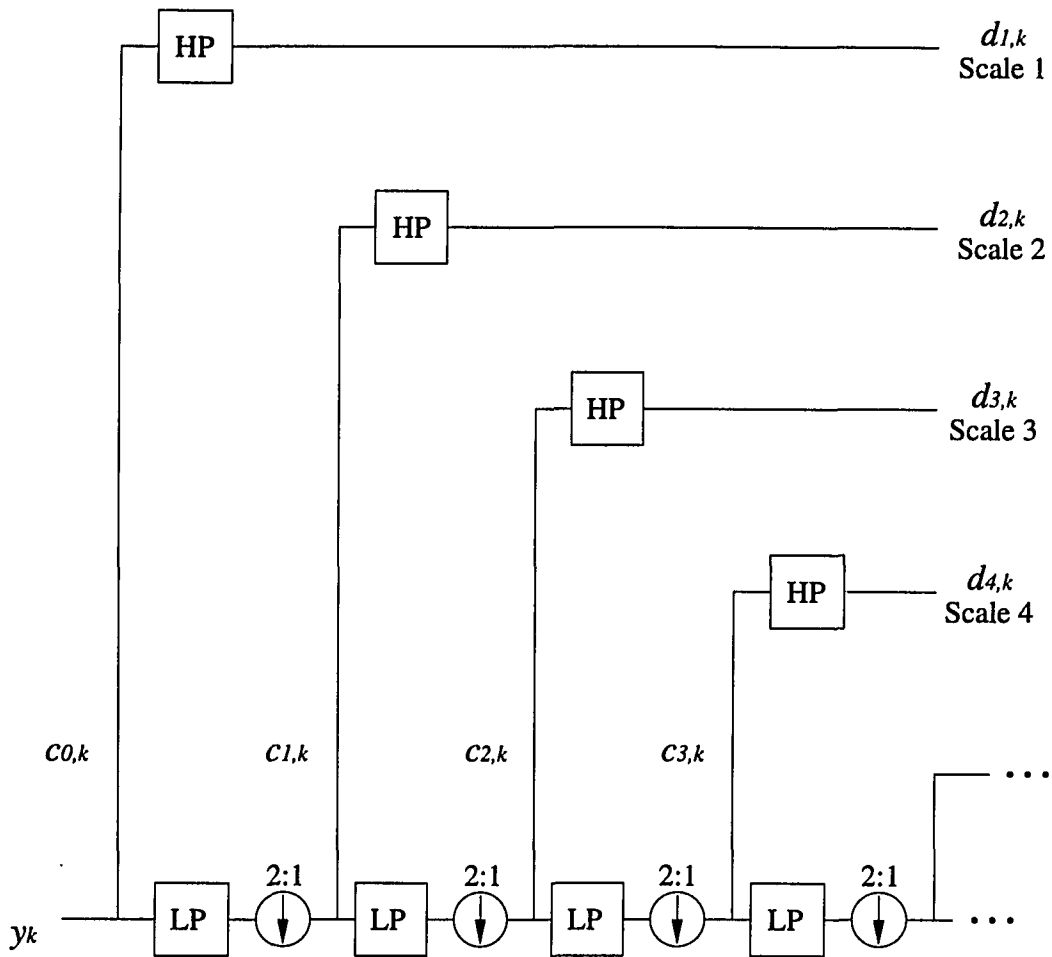
Figure 7.6: Pyramid algorithm for forward discrete wavelet transform

184

In this pyramid algorithm, the wavelet coefficients h_k and g_k in Eq. (7.32) and Eq. (7.33) are used as the filter coefficients of the QMF filter pairs. The wavelet coefficients h_k are used in low-pass (LP) filters and g_k are used in high-pass (HP) filters. The output of each low-pass filter is the approximation components $(c_{j,k})$ of the original signal at resolution of 2^j (at scale j in the tree). The output of each high-pass filter is the detail components $(d_{j,k})$ of the original signal at resolution 2^j. The $c_{j,k}$ of the previous scale are used to generate the new $c_{j,k}$ and $d_{j,k}$ for the next level of the tree.

The DWT is actually computationally more efficient than the FFT algorithm. An FFT of length N takes on the order of $N \log_2 N$ operations. A DWT of length N requires approximately N operations—the best possible. The sample code for computing the DWT of a data record is given in Appendix M.

7.5 Representation of AE signals by Wavelet Decomposition

When the DWT algorithm is applied to an AE signal, the signal is represented in a time-scale plane where scale is related to but not identical to frequency. Scales are implemented by dilating a mother wavelet ψ in the time domain. Each dilation is a doubling of the wavelet length in the time domain which results in a halving of the bandwidth in the frequency domain.

This operation is repeated recursively in the pyramid algorithm, yielding a group of signals that divide the spectrum of the original AE signal into octave bands or *scales* with successively coarser measurements in time as the width of each spectral band narrows and decreases in frequency. The division of frequency domain at each scale is shown in Figure 7.7.

185
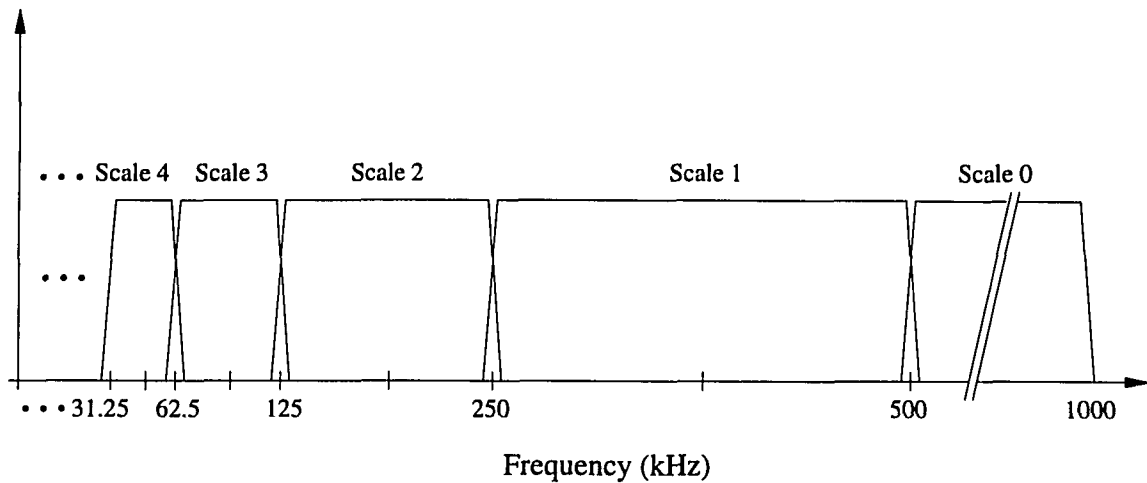
Figure 7.7: Division of frequency domain when sampling rate is 1000 kHz

A scale of the transform corresponds to one octave of AE signal bandwidth. Scale 0 corresponds to the highest frequencies in the AE signal; it contains frequencies starting from the signal sampling rate (in present problem 1000 kHz) to the Nyquist frequency (500 kHz, half the sampling rate). Scale 1 contains frequencies ranging from the Nyquist frequency (500 kHz) to one half the Nyquist frequency (250 kHz). As the scale increases, the bandwidth decreases by a factor of two. Since the bandwidth decreases by a factor of two with each scale increase, the sampling rate or number of coefficients are also halved with each scale increase. This decimation process leads to an economical but complete representation of AE signals in the time-scale plane.

In the present problem AE signal records of length 256 (2^8) are considered for flank wear estimation. The AE signal is sampled at a frequency of 1000 kHz. When the DWT algorithm is applied to an AE record, the original signal is contained in scale 0. Next, scale 1 represents frequencies in the range from 500 kHz to 250 kHz with 128 $d_{1,k}$ coefficients. Similarly, scale 2 represents frequencies of 250 kHz

186

to 125 kHz with 64 $d_{2,k}$ coefficients. Successively larger scales have the following band widths and number of coefficients: 125–62.5 kHz/32, 31.25–15.625 kHz/16, 7.812–3.406 kHz/8, 1.703–0.805 kHz/4, 402-201 Hz/2, and 100–50 Hz/1. The structure of coefficients generated by the DWT algorithm are shown in Figure 7.8.

As mentioned in Chapter 5, the flank wear-sensitive AE frequencies in the present problem are spread in the range of 50 to 300 kHz. From Figure 7.7 it is clear that the frequency band 50–300 kHz is covered by scales 2 and 3 in the time-scale plane. Thus, an AE signal is represented by the detailed coefficients belonging to scale 2 and scale 3. As highlighted in Figure 7.8, the coefficients $d_{2,0}$ to $d_{2,63}$ and $d_{3,0}$ to $d_{2,31}$ represent the AE signals in the present problem. These 96 coefficients are used as the AE signal features for flank wear estimation. The vectors of these coefficients are input to a neural network to get flank wear estimates.

7.6 Conclusions

This chapter studied the properties of AE signals generated in turning processes. The AE signals generated in turning are time-varying and consist of both continuous and transient parts. This chapter discussed the usefulness of discrete wavelet transforms to represent non-stationary AE signals.

The discrete wavelet transform coefficients computed from AE signals are treated as the compressed representation or features of the AE signals. In the present application, AE records of length 256 were used. The detailed coefficients of scale 2 (2^6) and scale 3 (2^5) are used as the AE signal features because they correspond to the AE frequency bands sensitive to flank wear. These AE-signal features are used to train a neural network, and subsequently to obtain the flank

187

Scale 0	Scale 1	Scale 2	Scale 3	Scale 4	··· Scale 7
$c_{0,0}$	$c_{1,0}$	$c_{2,0}$	$c_{3,0}$	$c_{4,0}$	
$c_{0,1}$	$c_{1,1}$	$c_{2,1}$	$c_{3,1}$	$c_{4,1}$	
		...	...	...	
		...	...	...	
	$c_{1,126}$	$c_{2,62}$	$c_{3,30}$	$c_{4,14}$	
	$c_{1,127}$	$c_{2,63}$	$c_{3,31}$	$c_{4,15}$	
	$d_{1,0}$	$d_{2,0}$	$d_{3,0}$	$d_{4,0}$	
	$d_{1,1}$	$d_{2,1}$	$d_{3,1}$	$d_{4,1}$	
		...	...	...	... ...
		...	...	...	
		$d_{2,62}$	$d_{3,30}$	$d_{4,14}$	
		$d_{2,63}$	$d_{3,31}$	$d_{4,15}$	
		- - - -	- - - -	- - - -	
		$d_{1,0}$	$d_{2,0}$	$d_{3,0}$	
...		$d_{1,1}$	$d_{2,1}$	$d_{3,1}$	
...			...	...	
	...		$d_{2,62}$	$d_{3,30}$	
	...		$d_{2,63}$	$d_{3,31}$	
			- - - -	- - - -	... ...
			$d_{1,0}$	$d_{2,0}$	
	...		$d_{1,1}$	$d_{2,1}$	
	...			...	
				...	
				$d_{2,62}$	
			...	$d_{2,63}$	
			...	- - - -	
				$d_{1,0}$	
				$d_{1,1}$	
				...	
				...	
$c_{0,254}$	$d_{1,126}$	$d_{1,126}$	$d_{1,126}$	$d_{1,126}$	... ...
$c_{0,255}$	$d_{1,127}$	$d_{1,127}$	$d_{1,127}$	$d_{1,127}$	

Figure 7.8: Wavelet coefficient structure

188

wear estimates.

The details of the design and training aspects of the neural network for flank wear estimation are discussed in Chapter 8. The flank wear estimation results using force, vibration and AE measurements are analyzed in Chapter 9.

Chapter 8

NEURAL NETWORKS FOR FLANK WEAR ESTIMATION

8.1 Introduction

The objective of this chapter is to discuss a suitable neural network model and its learning algorithm to relate sensor data information with flank wear. The role of a neural network in this research is to approximate a nonlinear function $f : \mathbf{R}^n \to \mathbf{R}$ that maps sensor measurements to flank wear values:

$$w = f(\mathbf{x}), \tag{8.1}$$

where w is flank wear and $\mathbf{x} \in \mathbf{R}^n$ is the corresponding vector of sensor measurements.

The following are the two most desirable properties in a neural network model for flank wear estimation:

1. The network should be able to learn partly in unsupervised mode so that the demands for supervised training data and training time are reduced. The collection of supervised training data involves measuring of both the sensor readings as well as the corresponding flank wear levels. On the other hand, the collection of unsupervised training data involves the acquisition of sensor data only. It is important that a neural network in this application learns at least partly from unsupervised training data because this data can be generated on-line without interrupting the cutting process. Generating supervised

190

training data requires intermittent interruption of the cutting process for flank wear measurements.

2. The network should be fault-tolerant to the fluctuations in sensors signals caused by non-uniformities in workpiece material composition and hardness and phenomenon such as built-up edge and chip breaking.

To satisfy these two criteria, in this research, a novel fault-tolerant neural network is created by combining Kohonen's feature maps, radial basis function networks, and recurrent neural networks. This network learns partly in unsupervised mode and partly in supervised mode. The overall architecture of the proposed neural network is shown in Figure 8.1. This architecture has three stages. In the first stage, a Kohonen's feature map finds the centers in the input space defined by input patterns created from the sensor signals. In the second stage, a radial basis function network utilizes the information about the centers in input space and maps the input patterns to flank wear estimates. In the third stage, a recurrent neural network integrates the first level flank wear estimates, provided by a set of radial basis function networks, to generate a second level flank wear estimate which is more accurate and reliable. The following sections describe the role of the radial basis function networks, Kohonen's feature maps, and the recurrent neural networks in the combined neural network architecture.

8.1.1 Hybrid-Learning Neural Network Architecture

A neural network architecture which uses error-driven top-down supervised learning and data-driven bottom-up unsupervised learning offers several improvements over a feedforward neural network which uses strictly supervised learning. A typical hybrid-learning neural network architecture has a layer of processing

191
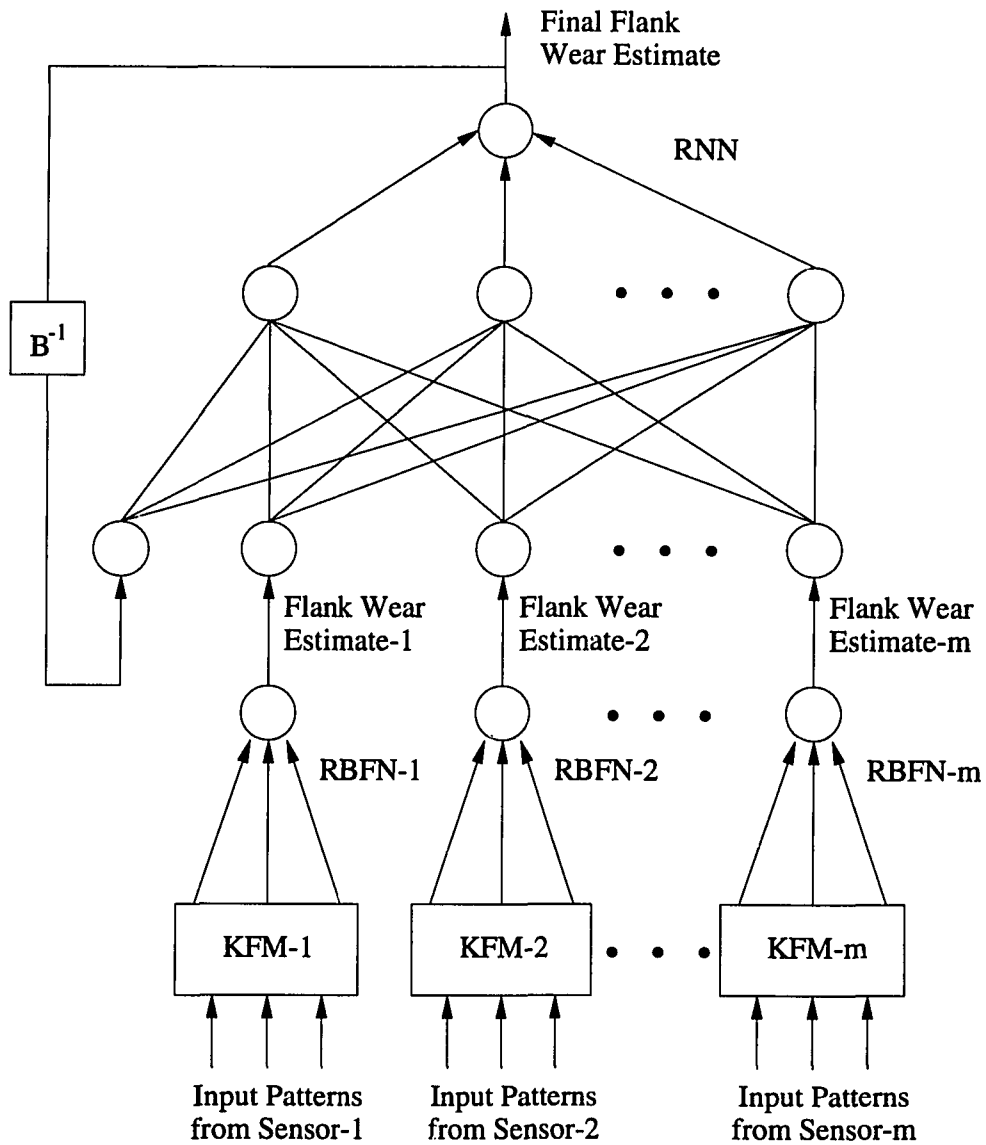
Figure 8.1: The overall neural network structure for flank wear estimation

192

nodes which are defined at the centers in the input space, and a set of weights which link the output of these processing nodes to an output node. The centers in the input space are determined through bottom-up unsupervised learning. The weights that link the processing nodes to the output node are determined through top-down supervised learning. The convergence of these weights is much faster than weights in a feedforward network because the unsupervised learning which precedes the supervised learning provides a statistically optimal form of input preprocessing (Lippman 1989, Moody and Darken 1989, Holdway 1989, Renals and Rohwer 1989, Ghosh and Chakravarty 1992). In view of these advantages, a hybrid-learning neural network architecture is adapted to compute the first level flank wear estimates. In this architecture, the bottom-up unsupervised learning is achieved through a Kohonen's feature map and the top-down supervised learning is achieved through a radial basis function.

The proposed hybrid-learning neural network architecture offers many advantages over a feedforward neural network trained using the backpropagation algorithm. Unlike a feedforward neural network, the hybrid-learning architecture does not pose the problem of finding the optimal network topology. The combination of locality of representation and linearity of learning in the hybrid-learning architecture offers tremendous speed advantages relative to the backpropagation algorithm. The backpropagation algorithm is also an inefficient algorithm for the present application where the size of input patterns is large and where the relevant input features cannot be separated from the irrelevant input features.

As the size of the input patterns increases, the sensitivity of a feedforward network to initial conditions increases. The number of fan-in connections to a node in the first hidden layer are determined by the size of the input patterns.

As the size of the input patterns increases, the fan-in weighted sum of a node in the first hidden layer may become large. In such a situation, the network learning reaches a saturation point very quickly unless the network weights are initialized to very small values or the hidden-node thresholds are selected carefully (Holdway 1989).

When the input patterns contain both the relevant and the irrelevant features, the randomly assigned values to the first-layer weights in a feedforward neural network will not result in any immediate correlation between the relevant features in the input space and activity of the first-layer hidden nodes. As a result, some useful information will be lost as it is passed on to the following layers (Holdway 1989).

8.1.2 Integration of First-Level Flank Wear Estimates

In the proposed flank wear estimation scheme, the data from each sensor is fed to a separate hybrid-learning neural network which provides the flank wear estimates based on those sensor measurements. The recurrent neural network integrates the flank wear estimates based on individual sensor measurements to provide a more accurate and robust flank wear estimate. The recurrent neural network also improves the fault-tolerant nature of the flank wear estimation scheme.

The combination of a hybrid-learning neural network an.1 a recurrent neural networks forms a new neural network architecture which provides the advantage of short training time, superior learning, and fault-tolerance.

8.2 A Radial Basis Function Network

A Radial Basis Function (RBF) network is constructed based on the theory of

194

radial basis functions. The application of Kohonen's feature maps to find the centers of a RBF network is the characteristics feature of the present implementation of a RBF network.

Traditionally, RBFs have been used as a technique for multivariate scattered data interpolation. Recently, they were found to be useful for approximating non-linear functions of several variables. Although the original applications of RBFs was in interpolation, a definition in the context of approximation is introduced by Jackson (1988).

Definition: *Given a continuous function $F : \mathbf{R}^+ \rightarrow \mathbf{R}$, and points $\{c_i : i = 1, 2, \ldots, N\}$, $c_i \in \mathbf{R}^+$, which become dense in the open region $\mathbf{D}$ of $\mathbf{R}^+$, there exists a sequence of functions*

$$f(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i \phi\left(\|\mathbf{x} - \mathbf{c}_i\|\right) \tag{8.2}$$

and some bounded open domain on which

$$|f(\mathbf{x}) - F(\mathbf{x})| \rightarrow 0 \tag{8.3}$$

as $N \rightarrow \infty$, where $\|\cdot\|$ is the Euclidean norm. The functions $\phi(\cdot)$ are termed radial basis functions, and the prototypical points c_i are referred to as the centers of the basis functions. The argument of a basis function is a scalar given by $\|\mathbf{x}_i - \mathbf{c}_i\|$.

In typical applications, RBFs are used to approximate a continuous function $F(\cdot)$ based on scattered sampling pairs $(\mathbf{x}_j, F(\mathbf{x}_j))$, where $\mathbf{x}_j \in \mathbf{D}$. In the present application, a Kohonen's feature map finds the centers c_i of input space defined by sensor measurements and the RBF utilizes these centers and approximates the function that relates sensor measurements to flank wear.

195

Compared to a multilayer feedforward neural network, the proposed RBF network requires shorter supervised training data and time (Ghosh and Chakravarthi 1992, Huang and Lippmann 1988, Moody and Darken 1989). Unlike in a multilayer neural network, the convergence of RBF weights to a global optimum is guaranteed because these weights are linearly related to RBFs in the network.

8.2.1 Approximation Properties of RBF Network

Consider a problem which has to be solved by a type of neural network. Suppose that there exists a decision function $f : \mathbf{R}^n \to \mathbf{R}^m$ whose implementation as a network plays a central role in the solution of the problem. Say that a family of functions G mapping $\mathbf{R}^n \to \mathbf{R}^m$ characterized by a certain structure and having certain elements (for example, multilayer feedforward network or RBF network) is expected to solve the problem by implementing some satisfactory member of G. The first question to consider is: Is this family G broad enough to contain f or a good approximation of f?

Several papers address this question in the case of multilayer neural network models with sigmoidal nonlinearities. It has established that multilayer neural networks can actually approximate any decision function drawn from a certain large class (Cybenko 1989, Hornik et al. 1989).

However, recently Park and Sandberg (1991) have proved that the RBF networks with only one hidden layer are also capable of universal approximation under the conditions stated in the following Theorem 1. Consider the family of RBF networks consisting of functions $f : \mathbf{R}^n \to \mathbf{R}$ represented by

$$f(x) = \sum_{i=1}^{N} w_i \cdot \phi \left(\frac{\mathbf{x} - \mathbf{c}_i}{\sigma} \right) , \tag{8.4}$$

196

where N is the number of RBF centers, $w_i \in \mathbf{R}$, $\sigma > 0$, $\mathbf{x} \in \mathbf{R}^n$, and $\mathbf{c}_i \in \mathbf{R}^n$. Let this family of RBF networks be denoted by S_K. The Theorem 1 given below establishes that, under certain mild conditions on the kernel function ϕ, RBF networks represented by Eq. (8.2) are capable of approximating arbitrarily close any function in $L_p(\mathbf{R}^n)$.

Theorem 1: *Let $\phi : \mathbf{R}^n \to \mathbf{R}$ be an integrable bounded function such that ϕ is continuous almost everywhere and $\int_{\mathbf{R}^n} \phi(x)dx \neq 0$. Then the family S_K is dense in $L_p(\mathbf{R}^n)$ for every $p \in [1, \infty]$* (Park and Sandberg 1991).

8.2.2 Construction of RBF Network

Figure 8.2 gives a schematic representation of a RBF network with n input vectors of size m and a scalar output. Such a network implements flank wear estimation using

$$w = g(y),\tag{8.5}$$

$$g(y) = \left[\frac{W_{max}}{1 + e^{-y}}\right],\tag{8.6}$$

$$y = f(x),\tag{8.7}$$

$$f(\mathbf{x}) = \lambda_0 + \sum_{i=1}^{N} \lambda_i \phi\left(\|\mathbf{x} - \mathbf{c}_i\|\right),\tag{8.8}$$

where $w \in \mathbf{R}$ is a flank wear estimate, y is an intermediate output of RBF network, $\mathbf{x} \in \mathbf{R}^n$ is an input vector of sensor measurements, $\phi(\cdot)$ is a given RBF from $\mathbf{R}^+ \to \mathbf{R}$, $\| \cdot \|$ denotes the Euclidean norm, $\lambda_0, \lambda_1, \ldots, \lambda_N$ are the RBF weights,

197

$c_1, \ldots, c_N \in \mathbf{R}^n$, are the RBF centers, and N is the number of RBF centers and W_{max} is the maximum expected flank wear in a given application.

In the present situation, $f : \mathbf{R}^n \to \mathbf{R}$ maps a sensor measurement x to a real value y which when substituted into the sigmoidal function $g : \mathbf{R} \to \mathbf{R}^+$ provides the corresponding flank wear estimate w.

In the RBF network design, the sigmoidal function $g : \mathbf{R} \to \mathbf{R}^+$ is introduced specifically to serve an important purpose in the present application. It smoothly limits flank wear estimates between 0 and W_{max}:

$$\lim_{y \to \infty} \frac{W_{max}}{1 + e^{-y}} = W_{max}, \tag{8.9}$$

$$\lim_{y \to -\infty} \frac{W_{max}}{1 + e^{-y}} = 0. \tag{8.10}$$

This property of sigmoidal function filters out the adverse effects of hardspots and nonhomogenities in the workpiece material for flank wear estimates. When hardspots and nonhomogenities in the workpiece material are encountered, y might take a large negative or positive value but the sigmoidal function smoothly binds flank wear estimates between 0 and W_{max}. In the present RBF network design, W_{max} is set to 0.0200 inches.

The middle layer in an RBF network can be considered equivalent to a hidden layer in a feedforward neural network. Similarly, the function $\phi(\cdot)$ can be considered equivalent to a hidden node activation function. In a RBF network, the functional form of $\phi(\cdot)$ and the RBF centers c_i are fixed before finding the network weights λ_i using a set of training input vectors x^p and the corresponding target outputs w^p (the superscript p refers to the pth input pattern).

198

w

$w = g(y)$

$$g(y) = \frac{W_{max}}{1 + e^{-y}}$$

$y = f(x)$

Threshold $\quad 1 \longrightarrow$

$$f(x) = \lambda_0 + \sum_{i=1}^{N} \lambda_i \, \phi(\| x - c_i \|)$$

λ_0

λ_1 $\quad$ λ_2 $\quad$ λ_N

$\phi(\| x - c_1 \|)$ $\quad$ $\phi(\| x - c_2 \|)$ $\quad \bullet \bullet \bullet \quad$ $\phi(\| x - c_N \|)$

N radial basis
function centers
with functional
form $\phi(.)$

$\bullet \bullet \bullet$

x_1 $\qquad$ x_2 $\qquad$ x_N $\quad$ Input Pattern
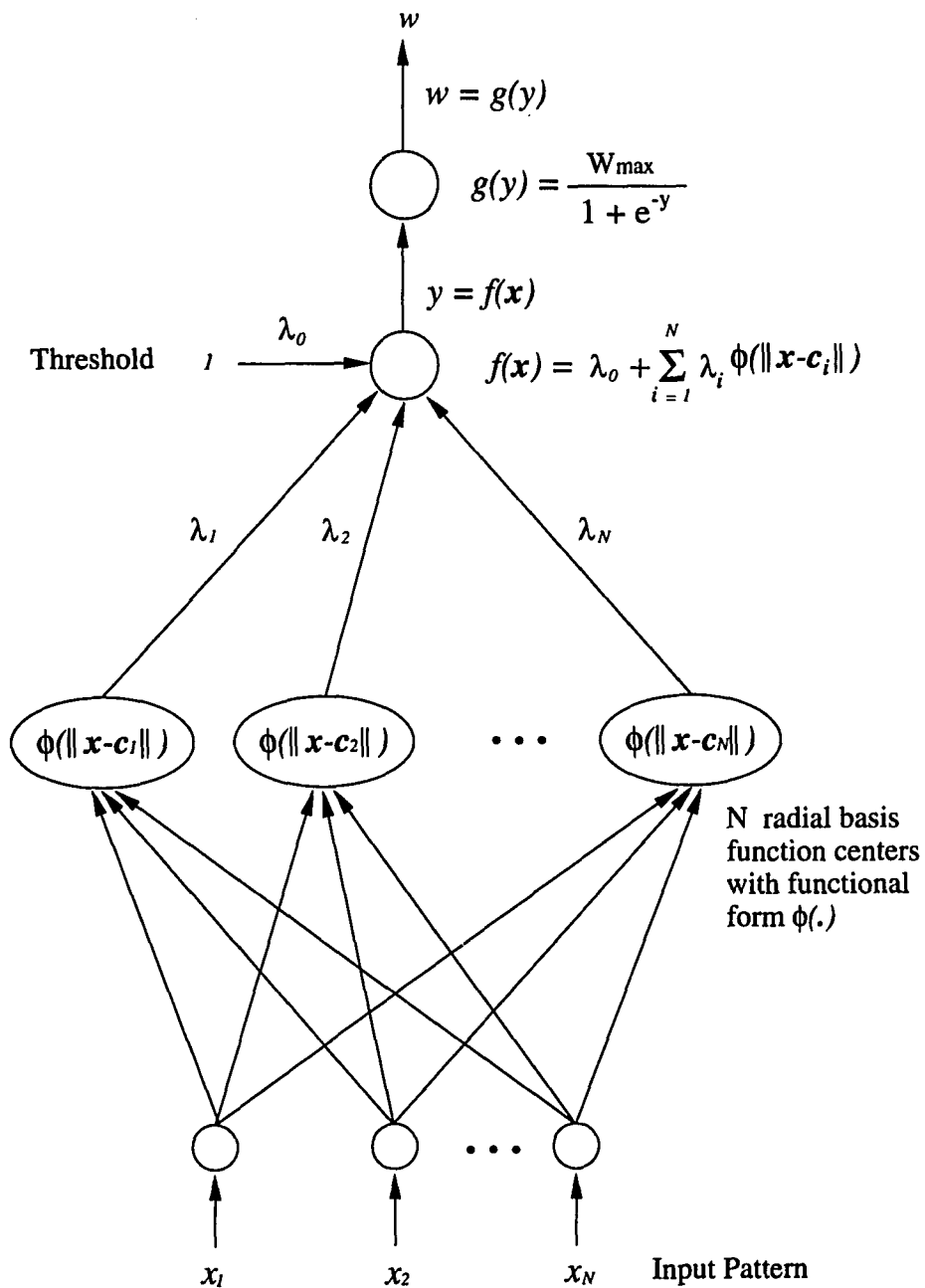
Figure 8.2: Architecture of a radial basis function network

199

An important design problem in constructing a RBF network is how to choose the functional form of $\phi(\cdot)$ and centers c_i such that the corresponding RBF network gives the best performance for a given problem. Theoretical investigations and practical results suggest that the choice of the nonlinear functional form $\phi(\cdot)$ is not crucial for the performance of a RBF network, as long as it is a radially symmetric smooth real-valued function (Chen et al. 1991). In this research a Gaussian functional form is chosen for implementing a RBF network:

$$\phi(\|\mathbf{x} - \mathbf{c}_i\|) = \exp\left[-\frac{(\mathbf{x} - \mathbf{c}_i)^T(\mathbf{x} - \mathbf{c}_i)}{2\sigma_i^2}\right], \tag{8.11}$$

where $\mathbf{c}_i$ and σ_i are centers and smoothing parameters of the ith hidden node in the RBF network as shown in Figure 8.2.

Selecting an appropriate set of centers and smoothing parameters is crucial for close approximation of the mapping function f. In practice, the centers are normally chosen from the training data points. In real world problems where the number of training data points is large, to include all the data points as centers may give rise to an extremely complex and useless model of the input space. On the other hand, arbitrarily choosing some data points as centers may result in incorrect and insufficient sampling of the input space (Chen et al. 1990).

8.2.3 Finding Centers and Smoothing Parameters of RBF Network

In this work, the centers of the RBF network are determined by training a Kohonen's feature map (Kohonen 1990) using the unsupervised competitive learning algorithm. Once the RBF centers are estimated, the smoothing parameters are found by computing the variance of input vectors from these RBF centers. The structure of a Kohonen's feature map is shown in Figure 8.3. This network

200
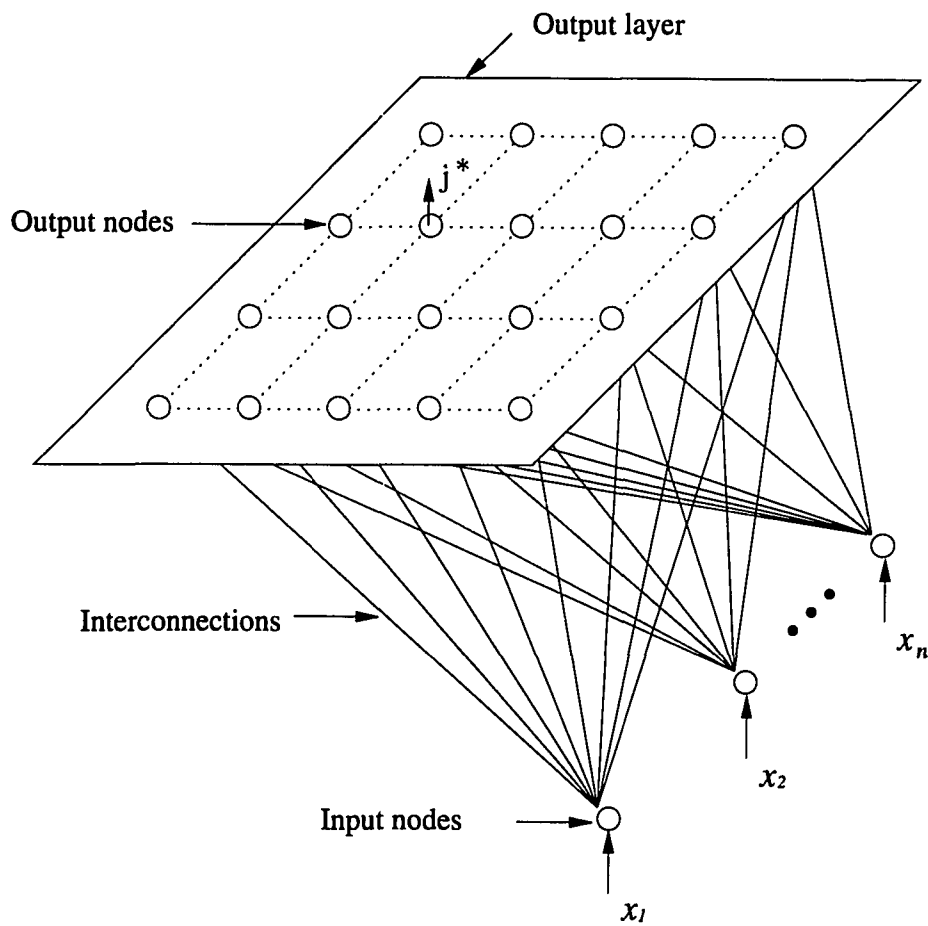
Figure 8.3: Architecture of a Kohonen's feature map

201

has an array of input nodes and a sheet-like layer of output nodes. Every output node is fully connected to input nodes. The number of input nodes is equal to the size of an input vector n. Usually, the output nodes are arranged in a square or a hexagonal grid and the size of the grid is chosen arbitrarily. In the present application, a square grid of output nodes is used. The grid size is selected such that the number of output nodes in the grid are slightly larger than the number of input patterns. This gives an opportunity to the feature map to form as many centers as the number of input patterns if they are considerably different from each other.

The Kohonen's feature map is trained in an unsupervised mode using a set of training patterns $\mathbf{x}^p \in \mathbf{R}^n$. During training, the feature map self-organizes its weight vectors $\mathbf{v}_j \in \mathbf{R}^n$ in such a way that they approximate the probability density function $p(\mathbf{x})$ of the input space in the sense of some minimal residual error. The following are the essential steps in training the Kohonen's feature map (Lippmann 1987, Kohonen 1990):

1. Initialize the feature map weights $\mathbf{v}_j$ to small random values.

2. Present an input vector $\mathbf{x}^p$.

3. Compute $d(\mathbf{x}^p, \mathbf{v}_j)$, the distance between $\mathbf{x}^p$ and $\mathbf{v}_j$. At a time instance t, the Euclidean distance $d(\mathbf{x}^p, \mathbf{v}_j)$ is computed as

$$d(\mathbf{x}^p, \mathbf{v}_j) = \sum_{k=1}^{n} \left[x_k^{p,(t)} - v_{jk}^{(t)} \right]^2 . \tag{8.12}$$

4. Select the winning node j^* such that

202

$$d(\mathbf{x}^p, \mathbf{v}_{j^*}) = \min_j \{d(\mathbf{x}^p, \mathbf{v}_j)\}. \tag{8.13}$$

8. Update weight vector of j^* and that of its neighbors according to the following learning rule:

$$\mathbf{v}_j^{(t)} = \begin{cases} \mathbf{v}_j^{(t-1)} + \alpha^{(t-1)}[\mathbf{x}^{p,(t)} - \mathbf{v}_j^{(t-1)}] & \text{if } j \in N_{j^*}^{(t-1)}, \\ \mathbf{v}_j^{(t-1)} & \text{if } j \notin N_{j^*}^{(t)}, \end{cases} \tag{8.14}$$

where $\alpha^{(t)}$ is a scalar-valued adoption gain $0 < \alpha^{(t)} < 1$, and $N_{j^*}^{(t)}$ is the neighborhood set around j^*. The adoption gain $\alpha^{(t)}$ decreases linearly or exponentially with time t. In the present work $\alpha^{(t)}$ is taken as a linearly decreasing function. The radius of neighborhood set $N_{j^*}^{(t)}$ also decreases with time. At the beginning of the training, $N_{j^*}^{(t)}$ has a wide radius and it decreases as the training progresses.

6. Repeat steps 2 to 5 for all input vectors $\mathbf{x}^p$ over a desired number of iterations.

When the training is over, let N_c be the set of nodes that are activated by input patterns and N be the total number nodes in that set. Generally N will be smaller than the total number of input patterns. The weight vectors of nodes in N_c sample the input space such that the point density function of weight vectors tends to approximate the probability density function of the input space. Thus the weight vectors of nodes in N_c provide the best set of centers in the input space. These weight vectors are used as the RBF centers in Eq. (8.11) by assigning

$$\mathbf{c}_j = \mathbf{v}_j \qquad \text{for } j \in N_c. \tag{8.15}$$

When $\mathbf{c}_j$ are fixed, σ_j^2 in Eq. (8.11) are computed as the variance of input vectors around $\mathbf{c}_j$:

203

$$\sigma_{jk}^2 = \frac{\sum_{i=1}^{P}[c_{jk} - x_{ik}]^2}{M - 1}, \tag{8.16}$$

where P is the number of input patterns and $k = 1, 2, \ldots, n$. Once $\mathbf{c}_j$ and σ_j^2 are fixed, the next step in constructing a RBF network is to determine the network weights λ_i in Eq. (8.8).

8.2.4 Finding RBF Network Weights

Let the training data set contain P pairs of input and target patterns. In the present application, an input pattern $\mathbf{x}^p$ consists of features extracted from a sensor signal and a target pattern consists of single flank wear measurement z^p. The superscript p is used to refer to the pth input pattern $\mathbf{x}^p$ and the pth target pattern z^p.

If the sigmoidal node hadn't been added to the RBF network, it would have been possible to find the RBF network weights λ_i using a Least Square (LS) estimation method. The presence of sigmoidal node adds nonlinearity between RBF center nodes and the output node. In view of this nonlinearity, a two stage training method is proposed. In the first stage, linearly optimal RBF network weights λ_i are determined using a LS estimation method. In the second stage λ_i are further adjusted using the gradient descent algorithm.

Stage 1:

Let the target pattern z^p be transformed to y^p through $y^p = g^{-1}(z^p)$, where $g(\cdot)$ is the sigmoidal function as defined in Eq. (8.6). Then

$$y^p = \ln(z^p) - \ln(W_{max} - z^p). \tag{8.17}$$

204

If y^p instead of z^p are used as the target patterns, the RBF network weights λ_i can be obtained by solving the following system of linear equations:

$$\mathbf{y} = \Phi \Lambda, \tag{8.18}$$

where

$$\Phi = \begin{bmatrix} 1 & \phi_1^1 & \phi_2^1 & \cdots & \phi_N^1 \\ 1 & \phi_1^2 & \phi_2^2 & \cdots & \phi_N^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \phi_1^P & \phi_2^P & \cdots & \phi_N^P \end{bmatrix}, \tag{8.19}$$

$$\Lambda = \begin{bmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \cdots & \lambda_N \end{bmatrix}^T, \tag{8.20}$$

$$\mathbf{y} = \begin{bmatrix} y^1 & y^2 & y^3 & \cdots & y^P \end{bmatrix}^T, \tag{8.21}$$

where Φ is a $P \times (N+1)$ matrix, Λ is a vector of size $(N+1)$, and $\mathbf{y}$ is the vector of size P. In Eq. (8.19), ϕ_i^p is the value of the ith radial basis function $\phi(\|\mathbf{x} - \mathbf{c}_i\|)$ in Eq. (8.8) when the pth pattern is presented to the RBF network.

When $P > (N+1)$, the system of equations in Eq. (8.18) can be solved for λ_i using a linear LS algorithm. The LS algorithm can be solved using

$$\Lambda = \left[\Phi^T \Phi \right]^{-1} \Phi^T \mathbf{y}, \tag{8.22}$$

or based on LU Decomposition (LUD), or Singular Value Decomposition (SVD) (Forsythe et al. 1977). In the present application, the SVD based LS method is used to compute RBF network weights. Though SVD is computationally more

205

expensive than other methods, it makes the LS algorithm more robust by handling the singularity problems caused by rank deficiencies in Φ matrix.

Stage 2:

In the second stage, the RBF network weights λ_i that were computed in the first stage are adjusted using the steepest descent algorithm to minimize the least squares criterion E:

$$E = \frac{1}{2} \sum_p (z^p - w^p)^2 \,, \tag{8.23}$$

where w^p is the output of the RBF network. This algorithm takes fixed steps in the direction of steepest descent in minimizing the criterion E:

$$\Delta\lambda_i = -\eta \frac{\partial E}{\partial \lambda_i} \qquad \text{for } i = 1, 2, \ldots, N \,, \tag{8.24}$$

$$\Delta\lambda_i = -\eta(z^p - w^p)\frac{\partial w^p}{\partial \lambda_i} \qquad \text{for } i = 1, 2, \ldots, N \,, \tag{8.25}$$

where η is the step size (or learning rate). The above equation can be further expanded using Eq. (8.5) to (8.8):

$$\lambda_i \leftarrow \lambda_i - \eta(z^p - w^p)\frac{\partial w^p}{\partial \lambda_i} \qquad \text{for } i = 1, 2, \ldots, N \,, \tag{8.26}$$

$$\lambda_i \leftarrow \lambda_i - \eta(z^p - w^p)\frac{\partial w^p}{\partial y_p}\frac{\partial y^p}{\partial \lambda_i} \qquad \text{for } i = 1, 2, \ldots, N \,, \tag{8.27}$$

$$\lambda_i \leftarrow \lambda_i - \eta(z^p - w^p)g'(y^p)\phi_i^p \qquad \text{for } i = 1, 2, \ldots, N \,, \tag{8.28}$$

206

$$\lambda_i \leftarrow \lambda_i - \eta(z^p - w^p)W_{max}w^p(1 - w^p)\phi_i^p \qquad \text{for } i = 1, 2, \ldots, N, \qquad (8.29)$$

where ϕ_i^p is the value of the ith radial basis function $\phi(\|\mathbf{x} - \mathbf{c}_i\|)$ in Eq. (8.8) when the pth pattern is presented to the RBF network. The Eq. (8.29) is used for updating the RBF weights for each input pattern $\mathbf{x}^p$ over a specified number of iterations (or until the RBF weights stabilizes).

8.3 Recurrent Neural Network

As shown in Figure 8.1, when flank wear estimates are available from different RBF networks, they are integrated together to get a more accurate and robust flank wear estimate using a recurrent neural network.

The details of this network are shown in Figure 8.4. For ease of explanation, this recurrent neural network is separated into two parts. The first part is a standard two-layer feedforward neural network with a single one-unit-time-lag feedback input. The nodes in the standard part of the recurrent network are represented by circles. The second part is a layer of preprocessing nodes. The nodes in this layer are represented by square boxes.

8.3.1 Standard Part of Recurrent Neural Network

The standard part of the network has one output node which provides the final flank wear estimates. It has $(m + 1)$ number of input nodes, where m is the number of RBF networks that provide the first level flank wear estimates. The number of hidden nodes is arbitrarily chosen to be equal to $5m$. The shape of the sigmoidal functions used by the output node and hidden nodes is different in this

207

$\hat{w}^{q,(t)}$

Final Flank
Wear Estimate

RNN

$\mathbf{B}^{-1}$

$\hat{w}^{q,(t-1)}$ $u_1^{q,(t)}$ $u_2^{q,(t)}$ $u_m^{q,(t)}$

$\hat{w}^{q,(t-1)}$ $\hat{w}^{q,(t-1)}$ $\hat{w}^{q,(t-1)}$

$w_1^{q,(t)}$ $w_2^{q,(t)}$ $w_m^{q,(t)}$
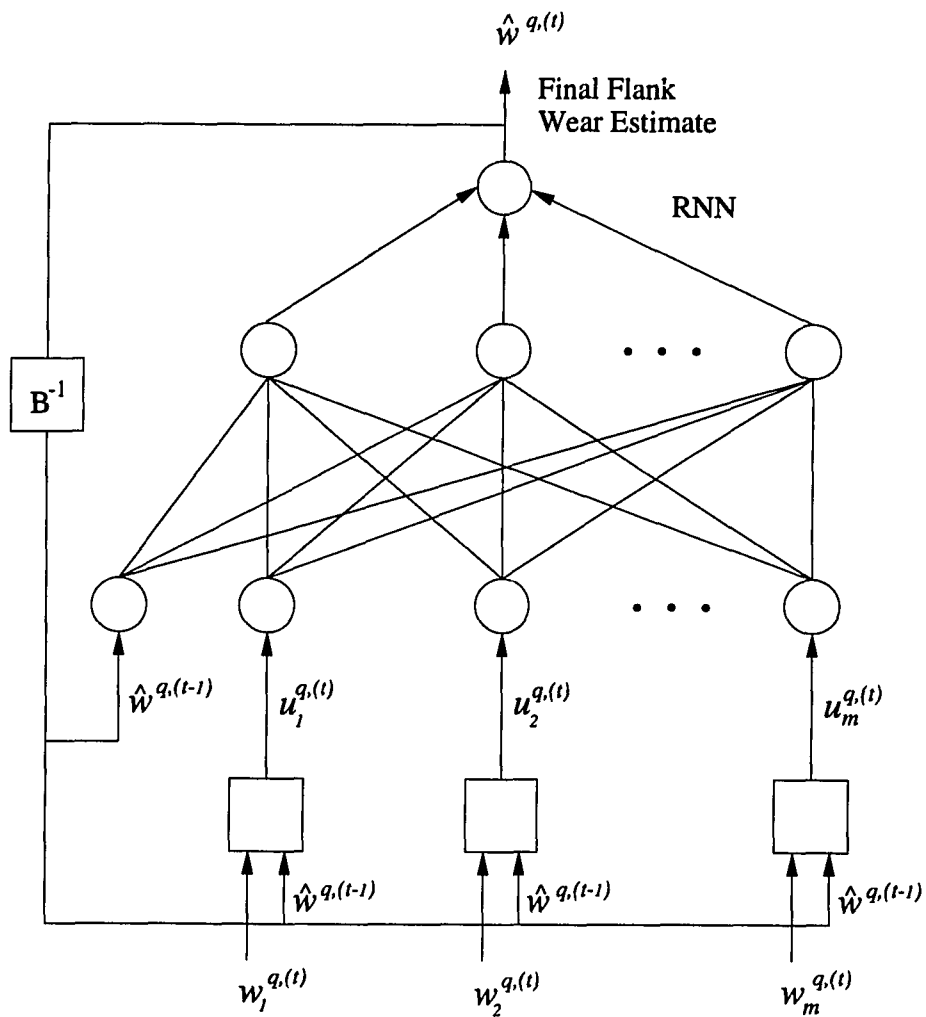
First level flank wear estimates from RBF networks

Figure 8.4: Architecture of a recurrent neural network

208

network:

$$o(y) = \frac{W_{max}}{1 + e^{-\beta y}} \qquad \text{for the output node,} \qquad (8.30)$$

$$h(y) = \frac{1}{1 + e^{-y}} \qquad \text{for a hidden node.} \qquad (8.31)$$

The parameters W_{max} and β provide a specific advantage in this application. By setting W_{max} to the maximum expected flank wear, the flank wear estimates are smoothly bound between 0 and W_{max}. The parameter β controls the steepness of the sigmoidal function. While a larger β is good for binary classification problems, a smaller β is more appropriate for continuous estimation problems. Usually, β is set to 1.0. By setting β to a value smaller than 1.0 in the present application, the accuracy of flank wear estimates is improved. However, the selection of too small a β should be avoided because it will create problems in a network convergence. In the present recurrent network, W_{max} is set to 0.0250 inches, and β is set to 0.25.

This recurrent neural network is trained using the classical backpropagation algorithm. The details of this algorithm can be found elsewhere (Rumelhart et al. 1986, Hertz et al. 1992).

8.3.2 Preprocessing Layer of Recurrent Neural Network

The layer of preprocessing nodes is added to the standard part of the recurrent neural network to serve a specific purpose in the present application. The preprocessing input nodes try to restrict the value of outlier flank wear estimates provided by individual RBF networks to a window assumed to contain approximately correct flank wear estimates. There is a one-to-one correspondence between input

209

nodes and the preprocessing nodes. A preprocessing node receives two inputs: $w_i^{q,(t)}$, the flank wear estimate from a RBF network, and $\hat{w}^{q,(t-1)}$, the feedback flank wear estimate (with one unit time lag) from the recurrent neural network. The superscript q refers the qth cutting tool. The output of the ith preprocessing node u_i^q is fed to the ith input node in the recurrent network. The function of a preprocessing node is given by

$$
u_i^{q,(t)} = \begin{cases} \hat{w}^{q,(t-1)} & \text{if } w_i^{q,(t)} \leq \hat{w}^{q,(t-1)}, \\ w_i^{q,(t)} & \text{if } \hat{w}^{q,(t-1)} < w_i^{q,(t)} < (\hat{w}^{q,(t-1)} + \Delta W), \\ \hat{w}^{q,(t-1)} + \Delta W & \text{if } w_i^{q,(t)} \geq (\hat{w}^{q,(t-1)} + \Delta W), \end{cases} \tag{8.32}
$$

where ΔW is the width of the window which is assumed to contain the acceptable flank wear estimates, and the superscript q refers to the qth cutting tool. The window width ΔW is chosen depending on the maximum possible distance between any two successive flank wear estimates. In the present application it is set to 0.004 inches. The above equation is developed based on the fact that flank wear increases monotonically with cutting time until the tip of the tool breaks down. However, a cutting tool is rarely used until the tip of the tool breaks down. In the present application, a cutting tool is considered worn out as soon as it develops a flank wear level beyond 0.0180 inches. The details of this tool life criteria were discussed in Chapter 5.

When the flank wear estimate $w_i^{q,(t)}$, given by the ith RBF network on the qth cutting tool at the tth time instance, is smaller than the flank wear estimate $\hat{w}^{q,(t-1)}$, given by the RRN at the tth time instance, then the output of the preprocessing node $u_i^{q,(t)}$ is set to $\hat{w}_i^{q,(t-1)}$. The idea is that if the current flank wear estimate $w_i^{q,(t)}$ provided by a RBF network is smaller than the previous flank wear

210

estimate $\hat{w}^{q,(t-1)}$ estimated by the recurrent neural network, then $w_i^{q,(t)}$ is assumed to be an outlier and is adjusted to $\hat{w}^{q,(t-1)}$.

When the current flank wear estimate $\hat{w}_i^{q,(t)}$ is within the window defined by $\hat{w}^{q,(t-1)}$ and $(\hat{w}^{q,(t-1)} + \Delta W)$, then $w_i^{q,(t)}$ is considered an acceptable flank wear estimate and is left unchanged ($u_i^{q,(t)}$ is set to $w_i^{q,(t)}$). The window width ΔW is chosen depending on the maximum possible distance between any two successive flank wear estimates. In the present application it is set to 0.004 inches.

When the current flank wear estimate $\hat{w}_i^{q,(t)}$ is higher than $(\hat{w}^{q,(t-1)} + \Delta W)$, then $w_i^{q,(t)}$ is considered again an outlier and is adjusted to the ceiling of the window ($u_i^{q,(t)}$ is set to $\hat{w}^{q,(t-1)} + \Delta W$).

As explained above, the preprocessing nodes detect the outlier in the flank wear estimates provided by the RBF networks and adjust them to either the floor or the ceiling of the window defined to contain acceptable flank wear estimates at time instance t.

8.4 Implementation Details

The previous sections have dealt with the theoretical details of the neural network designed for flank wear estimation. The following sections describe the implementation details.

As described in Chapter 5, the sensor data from Set 1 experiments are used for training the neural network, and the sensor data from Set 2 and Set 3 experiments are used for analyzing the flank wear estimation results. The following sections present the details of neural network training only. The analysis of flank wear estimation results is presented in Chapter 9. Set 1 contains a total of 168 data points collected from 15 cutting tools of the same type. Each cutting tool is run

211

at a unique combination of cutting speed and feed.

The neural network receives inputs from the measurements of individuals or combinations of AE, force, and vibration sensors. Seven separate neural networks are trained in the following seven sensor combinations:

1. AE sensor data.

2. Force sensor data.

3. Vibration sensor data.

4. AE, force, and vibration sensor data.

5. AE and force sensor data.

6. AE and vibration sensor data.

7. Force and vibration sensor data.

In these seven cases, force and vibration data are represented by univariate ARMA models, and AE data is represented by DWT. In addition to these seven neural networks, an eighth neural network is trained with force and vibration sensor data represented by multivariate ARMA models.

As explained in the previous sections, neural network training is conducted in three stages. These three stages are illustrated in Figure 8.1. In the first stage, a Kohonen's feature map is trained using the feature vectors extracted from sensor signals. The Kohonen's feature map provides the centers in input space. In the second stage, using these centers, a RBF network is trained. The RBF network provides the first level flank wear estimates. In the third stage, using the first level flank wear estimates provided by a set of RBF networks, a recurrent neural network is trained. The recurrent neural network provides the second level flank wear estimates. These second level approximations are considered the final flank

212

wear estimates.

In all the eight cases, output nodes of a Kohonen's feature map are arranged in a square grid size of 14×14. This amounts to 196 output nodes in a Kohonen's feature map. This grid size is chosen such that the number of output nodes are slightly higher than the number of training patterns, which happens to be 168 (total number of data points in Set 1 experiments) in the present problem. The number of input nodes in a Kohonen's feature map is equal to the size of input patterns it receives. The size of the input patterns depends on the sensor signals from which they are created. Each Kohonen's feature map is trained through 2000 iterations in the first phase and 3000 iterations in the second phase. During the first phase of the training, the radius of neighborhood nodes linearly decreases from seven to one and the learning rate linearly decreases from 1.0 to 0.1. After the total training of 5000 iterations, the feature map weights are observed to stabilize in the present application. A trained Kohonen's feature map provides centers in the input space. The number of centers in each case depends, again, on the sensor or sensors from which the input patterns are extracted.

Using the centers provided by the Kohonen's feature map, a RBF network is constructed. The initial weights of the RBF network are determined by a LS algorithm based on Singular Value Decomposition. The network weights are further adjusted further using the gradient descent algorithm. The gradient descent algorithm is run for 5000 iterations. Beyond 5000 iterations, the change in the RBF network weights is observed to be negligible. The trained RBF network provides the first level approximate flank wear estimates.

The first level flank wear estimates from a set of RBF networks are used to train a recurrent neural network. The recurrent neural network is trained using the

213

standard backpropagation algorithm through 5000 iterations. At 5000 iterations, the network weights have stabilized.

Tables 8.1 to 8.8 present the details of neural network training for eight different sensor combinations. All the programs were run on a Sparc 2 (SunOS 4.1.2) workstation with local disks (no NFS) and 32M memory. The main source code for Kohonen's feature maps is given in Appendix N. The main source code for training a RBF network and estimating the flank wear using the RBF are given Appendices O and P respectively. The main source code for training a recurrent neural network and estimating the flank wear using the recurrent neural network are given Appendices Q and R respectively.

8.5 Conclusions

This chapter described the construction, training, and implementation of a neural network for flank wear estimation. This neural network is constructed by integrating Kohonen's feature maps, radial basis function, and recurrent neural networks. The design of the neural network satisfies the following primary objectives:

1. The network utilizes unsupervised training data and reduces demand on the supervised training data and supervised training time.

2. The network provides fault-tolerance to fluctuation and non-uniformities in metal cutting.

The performance of this network is studied using the data from Set 2 and Set 3 experiments. Chapter 9 presents the analysis of flank wear estimation results obtained using the neural network architecture proposed in this chapter.

214

Table 8.1: Details of neural network training using AE data

Sensor: AE	
Number of input streams	*2*
Inputs to KFM-1	*Scale-2 DWT coefficients of AE signals*
Inputs to KFM-2	*Scale-3 DWT coefficients of AE signals*
Size of input patterns for KFM-1	*64*
Size of input patterns for KFM-2	*32*
Number of center created by KFM-1	*111*
Number of center created by KFM-2	*120*
CPU time required to train KFM-1	*420.38 min*
CPU time required to train KFM-2	*215.91 min*
CPU time required to train RBF-1	*2.94 min*
CPU time required to train RBF-2	*3.13 min*
CPU time required to train RRN	*2.55 min*
CPU time required per estimate	*0.15 sec*

Table 8.2: Details of neural network training using force data

Sensor: Force	
Number of input streams	*3*
Inputs to KFM-1	*ARMA coefficients of cutting force*
Inputs to KFM-2	*ARMA coefficients of feed force*
Inputs to KFM-3	*ARMA coefficients of radial force*
Size of input patterns for KFM-1	*6*
Size of input patterns for KFM-2	*6*
Size of input patterns for KFM-3	*6*
Number of center created by KFM-1	*122*
Number of center created by KFM-2	*125*
Number of center created by KFM-3	*128*
CPU time required to train KFM-1	*46.53 min*
CPU time required to train KFM-2	*46.49 min*
CPU time required to train KFM-3	*46.93 min*
CPU time required to train RBF-1	*3.29 min*
CPU time required to train RBF-2	*3.23 min*
CPU time required to train RBF-3	*3.39 min*
CPU time required to train RRN	*3.87 min*
CPU time required per estimate	*0.18 sec*

216

Table 8.3: Details of neural network training using vibration data

Sensor: Vibration	
Number of input streams	*2*
Inputs to KFM-1	*ARMA coefficients of cutting vibration*
Inputs to KFM-2	*ARMA coefficients of feed vibration*
Size of input patterns for KFM-1	*10*
Size input patterns for KFM-2	*10*
Number of center created by KFM-1	*122*
Number of center created by KFM-2	*120*
CPU time required to train KFM-1	*70.67 min*
CPU time required to train KFM-2	*71.86 min*
CPU time required to train RBF-1	*5.54 min*
CPU time required to train RBF-2	*5.50 min*
CPU time required to train RRN	*2.61 min*
CPU time required per estimate	*0.16 sec*

217

Table 8.4: Details of neural network training using AE, force, and vibration data

Sensors: AE, Force, Vibration	
Number of input streams	*7*
Inputs to KFM-1	*Scale-2 DWT coefficients of AE signals*
Inputs to KFM-2	*Scale-3 DWT coefficients of AE signals*
Inputs to KFM-3	*ARMA coefficients of cutting force*
Inputs to KFM-4	*ARMA coefficients of feed force*
Inputs to KFM-5	*ARMA coefficients of radial force*
Inputs to KFM-6	*ARMA coefficients of cutting vibration*
Inputs to KFM-7	*ARMA coefficients of feed vibration*
CPU time required to train RRN	*12.63 min*
CPU time required per estimate	*0.29 sec*

218

Table 8.5: Details of neural network training using AE and force data

Sensors: AE, Force	
Number of input streams	*5*
Inputs to KFM-1	*Scale-2 DWT coefficients of AE signals*
Inputs to KFM-2	*Scale-3 DWT coefficients of AE signals*
Inputs to KFM-3	*ARMA coefficients of cutting force*
Inputs to KFM-4	*ARMA coefficients of feed force*
Inputs to KFM-5	*ARMA coefficients of radial force*
CPU time required to train RRN	*7.55 min*
CPU time required per estimate	*0.22 sec*

219

Table 8.6: Details of neural network training using AE and vibration data

Sensors: AE, Vibration	
Number of input streams	*4*
Inputs to KFM-1	*Scale-2 DWT coefficients of AE signals*
Inputs to KFM-2	*Scale-3 DWT coefficients of AE signals*
Inputs to KFM-3	*ARMA coefficients of cutting vibration*
Inputs to KFM-4	*ARMA coefficients of feed vibration*
CPU time required to train RRN	*5.54 min*
CPU time required per estimate	*0.20 sec*

220

Table 8.7: Details of neural network training using force and vibration data (through univariate ARMA modeling)

Sensors: Force and Vibration	
Number of input streams	*5*
Inputs to KFM-1	*ARMA coefficients of cutting force*
Inputs to KFM-2	*ARMA coefficients of feed force*
Inputs to KFM-3	*ARMA coefficients of radial vibration*
Inputs to KFM-4	*ARMA coefficients of cutting vibration*
Inputs to KFM-5	*ARMA coefficients of feed vibration*
CPU time required to train RRN	*7.56 min*
CPU time required per estimate	*0.95 sec*

221

Table 8.8: Details of neural network training using force and vibration data (through multivariate ARMA modeling)

Sensors: Force and Vibration	
Number of input streams	*2*
Inputs to KFM-1	*multivariate ARMA coefficients of cutting, feed, radial forces*
Inputs to KFM-2	*multivariate ARMA coefficients of cutting, feed vibrations*
Size of input patterns for KFM-1	*54*
Size of input patterns for KFM-2	*40*
Number of center created by KFM-1	*114*
Number of center created by KFM-2	*113*
CPU time required to train KFM-1	*359.86 min*
CPU time required to train KFM-2	*121.30 min*
CPU time required to train RBF-1	*2.96 min*
CPU time required to train RBF-2	*2.94 min*
CPU time required to train RRN	*2.57 min*
CPU time required per estimate	*0.17 sec*

Chapter 9

RESULTS AND ANALYSIS

9.1 Introduction

This chapter analyzes the performance of the proposed flank wear estimation method. This analysis primarily compares the estimated flank wear values with the measured flank wear values for different sensor combinations at different operating conditions. The flank wear estimation results vary depending on the combination of sensors used in the estimation method. In this method, seven possible sensor combinations are used, namely:

1. AE sensor alone.

2. Force sensors only.

3. Vibration sensors only.

4. AE, force, and vibration sensors together.

5. AE and force sensors together.

6. AE and vibration sensors together.

7. Force and vibration sensor together.

The performance of the flank wear estimation method for these seven sensor combinations is studied on two test-data sets: Set 2 and Set 3. Set 2 has 95 data points collected over nine cutting tools, and Set 3 has 41 data points collected over six cutting tools. A data point contains the measurements of flank wear and data from five sensors: cutting force, feed force, radial force, cutting vibration, feed

223

vibration, and AE. The data points in Set 2 are collected at cutting speeds (100, 130, 160 sfpm) and feeds (0.0088, 0.0112, 0.0136 ipr) that were used for generating training data. The data points in Set 3 are collected at cutting speeds (190, 220 sfpm) higher than those used for generating the training data but at the same feeds (0.0088, 0.0112, 0.0136 ipr) as those used for generating the training data. The details of these test-data sets have been discussed in section 5.7 in Chapter 5. Tables 5.6, 5.7, and 5.8 list the operating conditions used for generating the data points in Set 2 and Set 3. The next section describes the point and interval estimators that are used in this chapter for evaluating the flank wear estimation results.

9.2 Statistical Estimators

To carry out the analysis of the flank wear estimation results, three statistical estimators are computed: mean, and standard deviation, and root mean square (RMS) of the paired differences of the measured and the estimated flank wear values. The confidence intervals for mean and variance of paired differences are computed to evaluate the accuracy of point estimators: mean and variance. The lower and upper quantiles of estimation errors are also computed to strengthen the inferences that can be drawn using confidence intervals.

9.2.1 Paired Differences

Let n be the total number of data points in a test-data set. The paired differences of the measured and the estimated flank wear values are defined as

$$d_i = \hat{w}_i - w_i \qquad \text{for } i = 0, \ldots, n, \qquad (9.1)$$

224

where $\hat{w}_i$ and w_i respectively are the ith estimated and measured flank wear values. In this chapter the paired differences are also sometimes referred to by flank wear estimation errors or simply estimation errors.

9.2.2 The Mean, Standard Deviation, and RMS

The mean $\bar{d}$ and the standard deviation s of the paired differences d_i are computed as follows:

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}, \qquad (9.2)$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}. \qquad (9.3)$$

The RMS of paired differences d_i is computed as the square root of average of the squares of the paired difference:

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}}. \qquad (9.4)$$

To evaluate the accuracy of the point estimates of population (flank wear estimates), confidence intervals for the population mean and variance are computed using t and χ^2 statistics respectively (Neter et al. 1985).

9.2.3 Confidence Intervals

The confidence intervals for mean and variance of flank wear estimation errors are computed when d_i s are assumed to be random samples from a normal population.

225

Let $d_i \sim n(\mu_d, \sigma_d^2)$, where μ_d and σ_d^2 are the population mean and variance respectively. The sample mean $\bar{d}$ and the sample standard deviation s are computed using Eq. (9.2) and (9.3). The standard deviation of the sampling distribution of $\bar{d}$ is estimated from

$$s_d = \frac{s}{\sqrt{n}}. \tag{9.5}$$

Since d_is are assumed to be random samples from a normal population, $\frac{(\bar{d}-\mu_d)}{s_d} \sim t_{(n-1)}$ and $\frac{(n-1)s^2}{\sigma_d^2} \sim \chi_{(n-1)}$. When Type I error is α, the $100(1-\alpha)\%$ confidence intervals for μ_d and σ_d^2 are given by

$$\bar{d} - t_{(1-\alpha/2,\, n-1)}s_d \leq \mu_d \leq \bar{d} + t_{(1-\alpha/2,\, n-1)}s_d, \tag{9.6}$$

$$\frac{(n-1)s}{\chi^2(1-\alpha/2,\, n-1)} \leq \sigma_d \leq \frac{(n-1)s}{\chi^2(\alpha/2,\, n-1)}. \tag{9.7}$$

In the present problem, α is set at 0.05 and 95% confidence intervals for mean and variance of flank wear estimation errors are computed.

9.2.4 Normal Probability Plots

The correctness of the confidence limits depends on the requirement that d_i follow a normal distribution. This requirement on d_i is verified by constructing a normal probability plot. The normal probability plot is a graphical technique to visually ensure that the points in a data set approximate random samples from a normal population. The following steps explain the construction and use of a normal probability plot (Box 1978):

1. Sort the elements d_i in the test-data set in ascending order.

226

2. Plot d_i versus x_i, where $x_i = F^{-1}(p_i)$, $p_i = \frac{i-0.5}{N}$, and $F(x)$ is the theoretical normal cumulative distribution function.

3. If the points on the plot are aligned close to a straight line, then it confirms that the points in the test-data set follow the normal distribution.

9.2.5 Lower and Upper Quantiles

The lower and upper quantiles of flank wear estimation errors are computed when no assumptions regarding the distribution of d_i are made.

Let d_i be the ith element in the set of paired differences sorted in ascending order. Let a discrete fraction p_i be defined as

$$p_i = \frac{i-0.5}{N} \qquad i = 1, 2, \ldots, N. \tag{9.8}$$

While p is a fraction between 0 and 1, the p quantile of the set of paired differences is denoted by $Q(p)$ and is defined as (Chambers et al. 1983)

$$Q(p) = \begin{cases} d_i & \text{if } p = p_i, \\ (1-f)Q(p_i) + fQ(p_{i+1}) & \text{if } p_i < p < p_{i+1}, \end{cases} \tag{9.9}$$

where

$$f = \frac{p - p_i}{p_{i+1} - p_i}. \tag{9.10}$$

In the present situation, the lower quantile $Q(0.025)$ and an upper quantile $Q(0.975)$ of flank wear estimation errors are computed. In addition to these two quantiles, $Q(0.5)$, which is known as the median of the data, is also computed.

227

9.3 Estimation Results for Test Data in Set 2

This section presents the flank wear estimation results for test data in Set 2. The analysis and discussion of the results is based on the mean, standard deviation, and RMS of estimation errors for each sensor combination. The test data in Set 2 are collected at the same cutting speed and feed combinations that were used for generating the neural network training data. This data set has a total of 95 data points collected over 9 cutting tools at different levels of flank wear. See Table 5.6 for details. As listed in this table, the cutting tools numbered from 15 to 24 are used to generate data in Set 2.

The measured and the estimated flank wear values are graphically shown in Figures 9.1 to 9.8. These plots show the measured and the estimated flank wear values over the life of the cutting tool. Table 9.1 gives the mean, standard deviation, and RMS of estimation errors for different sensor combinations. These parameters are graphically depicted in Figure 9.9.

Table 9.1: Statistics of estimation errors on Set 2 tools. These statistics are computed from 95 pairs of measured and estimated flank wear values over nine cutting tools.

Case #	Sensor Combination	Mean $\times 10^{-2}$ in.	Std. Dev. $\times 10^{-2}$ in.	RMS $\times 10^{-2}$ in.
1	AE	0.1241	0.1168	0.17000
2	Force (F)	-0.0179	0.1364	0.13689
3	Vibration (V)	0.0390	0.1338	0.13872
4	AE + F + V	-0.0057	0.1142	0.11370
5	AE + F	0.0018	0.1033	0.10280
6	AE + V	0.0424	0.1140	0.12104
7	F + V (ARMA)	-0.0137	0.1441	0.16574
8	F + V (MARMA)	0.1707	0.1592	0.23281

228

9.3.1 Estimation Results using AE Sensor

Figure 9.1 presents the plots of flank wear estimation results obtained using an AE sensor alone. In this case, the AE signals are represented by DWTs. There are nine plots in the figure, each one of which corresponds to the flank wear estimation over the tool life a cutting tool. Each one of these nine cutting tools is run at a different combination of cutting speed and feed. The cutting speed and feed for each tool are indicated at the right bottom corner of the corresponding plot.

From these plots, it can be concluded that the AE sensor tends to overestimate flank wear values at all combinations of cutting speed and feed. This can be attributed to an increase in AE energy due to factors other than flank wear. The most likely factors in the present experiment that could increase AE energy are built-up edge, chip fracture, and friction between the tangled chip and the workpiece.

Compared with force and vibration sensors, the AE sensor has a relatively large mean and RMS of estimation errors. This is because the AE sensor is sensitive to many factors other than flank wear in the cutting process. However, the standard deviation of estimation errors is smaller than those of either force or vibration sensors. This indicates that the AE sensor gives more consistent flank wear estimates over a range of operational conditions than either force or vibration sensors.

9.3.2 Estimation Results using Force Sensors

Figure 9.2 presents the plots of flank wear estimation results when force measurements in cutting, feed, and radial direction are used. In this case, the individual force signals are represented by separate univariate ARMA models. When

229

the cutting speed is 100 sfpm, a clear trend appears in the flank wear estimates. The flank wear estimates at the initial and final stages of tool life are closer to the measured flank wear values. In between the initial and final stages of tool life, flank wear values are underestimated. This behavior in the flank wear estimates can be attributed to the change in the tool rake angle due to the presence of a built-up edge. While conducting the actual experiments, formation of built-up edge was prevalent at the cutting speed of 100 sfpm. As the cutting speed increased to 130 sfpm or 160 sfpm, the formation of built-up edge was not noticed. When the built-up edge is absent at cutting speeds of 130 sfpm and 160 sfpm, the flank wear estimates are closer to measured flank wear values.

The flank wear estimates on cutting tool# 19 (speed 130 sfpm, and feed 0.0088 ipr) are steadily higher than the measured flank wear values. Such a trend might have been caused by some unobserved phenomenon, such as variation in material hardness, that occurred during this specific stretch of cutting.

In the case of force sensors, the mean of the estimation errors is negative. This indicates that force sensors have a tendency to underestimate the flank wear values over the range of operating conditions that were used for generating the data in Set 2. Nevertheless, individually, force sensors gave the best estimation results. Both the mean and the RMS of estimation errors are closer to zero than they are in the case of AE and vibration sensors.

9.3.3 Estimation Results using Vibration Sensors

Figure 9.3 presents the plots of flank wear estimation results when vibration sensors in feed and cutting direction are used. In this case, the individual vibration signals are represented by separate univariate ARMA models. At the cutting speed

230

of 100 sfpm, the flank wear estimates tend to be lower than the measured flank wear values. This trend disappears at higher cutting speeds, 130 and 160 sfpm.

Similar to force sensors, vibration sensors also overestimated the flank wear values on cutting tool# 19. This effect might be again attributed to some unobserved change in workpiece material properties during this specific stretch of cutting.

A comparison of the flank wear estimation results given by force and vibration sensors, shows that they follow the same trend. This similarity in trend is expected because both force and vibration measurements are affected by the same tool-workpiece dynamics (Rao 1986).

The accuracy of flank wear estimates given by vibration sensors is a little lower than that given by force sensors. The RMS and the standard deviation of estimation errors given by vibration sensors are almost comparable to those of force sensors. The mean of estimation errors given by vibration sensors is positive and farther away from zero than the mean of estimation errors given by force sensors. However, in a practical situation, vibration sensors may be more preferable to force sensors because adaptive control schemes and tool change strategies based on overestimated flank wear values are preferable to those based on underestimated flank wear values.

9.3.4 Estimation Results using AE, Force, and Vibration Sensors

Figure 9.4 presents the plots of flank wear estimation results when all AE, force and vibration sensors are used. In this case AE signals are represented by DWTs, and force and vibration signals are represented by univariate ARMA models. From these plots one can observe that flank wear estimation results obtained

231

using all three sensors are better than the results obtained using any one of them individually. Since force and vibration sensors give the same flank wear estimation trend, they induce a similar trend in the estimation results obtained using AE, force, and vibration sensors. At the cutting speed of 100 sfpm, flank wear estimates are underestimated. At other operating conditions, the flank wear estimates are closer to those given by either force or vibration sensors than to those given by the AE sensor.

All three parameters—mean, standard deviation, and RMS of flank wear estimation errors—show considerable reduction when AE, force, and vibration sensors are used together. This means that the flank wear estimates in this case are more accurate and consistent than those given by AE, force, and vibration sensors individually.

9.3.5 Estimation Results using AE and Force Sensors

Figure 9.5 presents the plots of flank wear estimation results when AE and force sensors are used. In this case, AE signals are represented by DWTs and force signals are represented by univariate ARMA models. Since the AE sensor tends to overestimate and force sensors tends to under-estimate flank wear, the combination of these two sensors gave the most accurate flank wear estimation results.

The mean, standard deviation, and RMS of estimation errors indicate that the AE-force sensor combination gives the most accurate and consistent flank wear estimates compared to the estimates given by any other sensor combination.

9.3.6 Estimation Results using AE and Vibration Sensors

232

Figure 9.6 presents the plots of flank wear estimation results when AE and vibration sensors are used. In this case, AE signals are represented by DWTs and vibration signals are represented by univariate ARMA models. At the cutting speed of 100 sfpm, where vibration sensors underestimate flank wear, the results are significantly improved due the compensatory effect from AE sensor estimates. At other cutting speeds, the over estimating nature of the AE sensor is reduced by the presence of vibration sensors. Because of this, the mean of estimation errors for the AE-vibration combination is lower than that of AE sensor alone but higher than that of vibration sensors.

The RMS and standard deviation of estimation errors indicate that the combination of AE and vibration sensors improves accuracy and consistency of flank wear estimates compared to estimation results obtained using either AE or vibration sensors.

9.3.7 Estimation Results using Force and Vibration Sensors

Figures 9.7 and 9.8 present the plots of flank wear estimation results when force and vibration sensors are used. The results shown in Figure 9.7 were obtained when individual force and vibration signals are modeled by separate univariate ARMA models. This is how force and vibration sensors are modeled in all the six cases discussed in the previous sections. In contrast, the results in Figure 9.8 were obtained when all force signals are modeled by a multivariate ARMA model and all vibration signals are modeled by another multivariate ARMA model.

233

Results for Univariate ARMA Case

Since force and vibration sensors have a similar flank wear estimation pattern, the performance of the force-vibration combination deteriorated compared to the performance of either force or vibration sensors separately. This can be observed from Figure 9.7 as well as from the standard deviation and RMS of estimation errors. This shows that by combining sensors of similar behavior, the estimation results will not improve.

Results for Multivariate ARMA Case

In this case, three force signals (in cutting, feed, and radial direction) are modeled by a multivariate ARMA model and two vibration signals (in feed and cutting directions) are modeled by another multivariate ARMA model. The flank wear estimation results obtained through multivariate ARMA modeling are quite different from those obtained through univariate ARMA modeling. The multivariate ARMA modeling case shows a consistent trend of overestimation. The flank wear estimation results obtained in this case are similar to those obtained using the AE sensor.

From the RMS and standard deviation of estimation errors, it can be observed that both the accuracy and the consistency of flank wear estimates in the multivariate ARMA case deteriorated compared to those of univariate ARMA modeling case. These statistics also indicate that the performance of force-vibration sensors through multivariate ARMA modeling is the poorest among all the eight cases listed in Table 9.1.

234

9.3.8 Confidence Intervals for Mean and Variance

In the present situation, the confidence intervals provide an estimate of the accuracy of the mean and variance that are computed from a set of data. The narrower the confidence intervals for mean and variance, the more reliable are the mean and variance of flank wear estimates.

As explained in section 9.2.3, the computation of confidence intervals for mean and variance assumes that the data approximate a random sample from a normal population. In order to verify the normality of flank wear estimation errors, normal probability plots are constructed for all cases of sensor combinations listed in Table 9.1. The procedure for construction and interpretation of normal probability plots was briefly explained in section 9.2.4. These normal probability plots are presented in Figures 9.10 to 9.17.

In the case of the AE sensor, AE-force-vibration sensors, and AE-force sensors, the points on the normal probability plots are fairly closely aligned to the straight lines that pass through them. This indicates that in these cases, the estimation errors closely follow the normal distribution. In other cases, the points on the normal probability plots show some minor deviations from a straight line. This indicates that in these cases, the estimation errors deviate from the normal distribution. Nevertheless, it can be concluded that in all cases of sensor combinations, there is no serious violation of normality. This confirms that confidence intervals for mean and variance of flank wear estimation errors can be computed with accuracy.

Tables 9.2 and 9.3 give the 95% confidence intervals for both mean and variance of estimation errors. These confidence intervals are computed as explained in section 9.2.3. Figure 9.18 and 9.19 depict the position and the width of these

235

confidence intervals. In these figures, each vertical bar corresponds to a confidence interval. The top and the bottom horizontal lines of each bar represent the upper and the lower confidence intervals. The horizontal line at the center of each bar represents the sample mean in Figure 9.18 and the sample variance in Figure 9.19.

Table 9.2: Confidence intervals for the mean of estimation errors. These confidence intervals are computed at 95% significance level from 95 pairs of measured and estimated flank wear values.

Case #	Sensor Combination	Lower Limit $\times 10^{-2}$ in.	Upper Limit $\times 10^{-2}$ in.	Interval $\times 10^{-2}$ in.
1	AE	0.10011	0.14805	0.04794
2	Force (F)	-0.04592	0.01006	0.05599
3	Vibration (V)	0.01150	0.06643	0.05492
4	AE + F + V	-0.02913	0.01771	0.04685
5	AE + F	-0.01936	0.02303	0.04240
6	AE + V	0.01900	0.06577	0.04677
7	F + V (ARMA)	-0.04331	0.01584	0.05915
8	F + V (MARMA)	0.1380	0.20335	0.06535

Table 9.3: Confidence intervals for the variance of estimation errors. These confidence intervals are computed at 95% significance level from 95 pairs of measured and estimated flank wear values.

Case #	Sensor Combination	Lower Limit $\times 10^{-4}$ in.	Upper Limit $\times 10^{-4}$ in.	Interval $\times 10^{-4}$ in.
1	AE	0.10863	0.19542	0.08679
2	Force (F)	0.14816	0.26653	0.11837
3	Vibration (V)	0.14258	0.25650	0.11392
4	AE + F + V	0.10373	0.18661	0.08288
5	AE + F	0.08498	0.15288	0.06790
6	AE + V	0.10339	0.18598	0.08259
7	F + V (ARMA)	0.16537	0.29749	0.13212
8	F + V (MARMA)	0.20163	0.36272	0.16109

236

These confidence intervals indicate that the AE-force sensor combination gives the narrowest confidence intervals for both mean and variance while the force-vibration combination leads to the broadest limits. Among univariate and multivariate ARMA representations of force-vibration signals, the latter representation leads to broader confidence intervals. The narrow confidence intervals for population mean and variance (mean and variance of flank wear estimation errors) infer that the estimated sample mean and variance (mean and variance of estimation errors computed from the data points in Set 2) are close to the true population mean and variance. Considering mean, variance, and their confidence intervals together, three sensor combinations give more accurate and reliable flank wear estimates: AE-force, AE-force-vibration, and AE-vibration, in the order of their performances. From the same point of view, force-vibration is the least effective sensor combination for flank wear estimation.

The force and vibration sensors individually show almost the same level of performance in giving flank wear estimates. Individually, their performance is superior to that of the AE sensor. The RMS and standard deviation of estimation errors are approximately the same for force and vibration sensors. The confidence intervals for variance of estimation errors in the case of force and vibration sensors are close to each other. The force and vibration sensors also give almost the same level of mean of estimation errors but with an opposite sign. The force sensors give negative mean of estimation errors where as the vibration sensors give positive mean of estimation errors. The force sensors have tendency to underestimate flank wear values while the vibration sensors have a tendency to overestimate.

Individually, the AE sensor shows a strong tendency of overestimation. However, when it is integrated with either force or vibration sensors, the accuracy

237

and consistency of flank wear estimates provided by those sensor combinations are greatly improved. The performance of AE-force-vibration, AE-force, or AE-vibration sensors is superior to AE, force, or vibration sensors alone.

The above analysis through confidence intervals rests on the assumption that flank wear estimation errors follow a normal distribution. Since the normal probability plots show that, in some cases, the estimation errors slightly deviate from the normality, the empirical quantile limits on the flank wear estimation errors are computed. The computation of empirical quantile limits does not make any assumption on the distribution of the estimation errors. These quantile limits are computed to reinforce the conclusions drawn using confidence intervals.

9.3.9 Lower and Upper Quantiles of Estimation Errors

The lower and upper quantiles of flank wear estimation errors are computed as explained in section 9.2.5. The lower quantile is computed at 0.025 fraction. The upper quantile is computed at 0.975 fraction. Defined at these fractions, the lower and the upper quantiles provide a window of 95% for flank wear estimation errors. Table 9.4 lists the median, lower and upper quantiles for each sensor combination. Figure 9.20 graphically displays this information. In this figure, each vertical bar corresponds to a 95% window for flank wear estimation errors. The top and the bottom horizontal lines of each bar represent the upper and the lower quantiles of estimation errors. There are two horizontal lines in between the top and the bottom horizontal lines: the thin horizontal line represents the median of the estimation errors and the thick horizontal line represents the sample mean of estimation errors.

238

Table 9.4: Lower and upper quantiles of estimation errors. Lower and upper quantiles here computed at 0.025 and 0.972 fractions from 95 pairs of measured and estimated flank wear values.

Case #	Sensor Combination	Median $\times 10^{-2}$ in.	Lower Quantile $\times 10^{-2}$ in.	Upper Quantile $\times 10^{-2}$ in.	Interval $\times 10^{-2}$ in.
1	AE	0.1186	-0.1483	0.3787	0.5270
2	Force (F)	0.0036	-0.3253	0.2391	0.6506
3	Vibration (V)	0.0265	-0.2125	0.3816	0.5941
4	AE + F + V	0.0032	-0.2240	0.2556	0.4796
5	AE + F	0.0008	-0.1916	0.2246	0.4162
6	AE + V	0.0408	-0.2018	0.2913	0.4931
7	F + V (ARMA)	0.0016	-0.3236	0.2815	0.6051
8	F + V (MARMA)	0.1324	-0.1015	0.5831	0.6846

The empirical quantile limits listed in Table 9.4 support the conclusions drawn based on the confidence intervals. The AE-force sensors provide the narrowest empirical window for the flank wear estimates, while the force-vibration sensors give the widest empirical window. The median in the case of the AE-force sensors is closer to zero than in any other case. Individually, the AE sensor has the highest median of flank wear estimates compared with the median of either force or vibration sensors. However, the integration of the AE sensor with either force sensors or vibration sensors narrows the empirical-window width compared to the window width of individual type of sensors.

9.4 Estimation Results for Test Data in Set 3

This section presents the flank wear estimation results for test data in Set 3. The test data in Set 3 is collected at the cutting speeds higher than those used for generating the neural network training data. This data set has a total of 41 data points collected using six cutting tools. See Table 5.7 for details. As listed in the

239

table, the cutting tools numbered from 25 to 30 are used for generating data in Set 3.

Figures 9.21 to 9.27 show the plots of measured and estimated flank wear values for all cases of sensor combinations: AE, force, vibration, AE-force-vibration, AE-force, AE-vibration, and force-vibration. From these plots the following two observations can be made:

1. In all cases of sensor combinations, there is uniformity in the characteristics of flank wear estimates. In the initial stage of a tool life, the estimated flank wear values are closer to the measured ones. As the cutting time increases, the flank wear estimates deviate gradually from the measured values. This deviation is always negative, indicating that the actual flank wear is being underestimated.

2. In all cases of sensor combinations, the deviation of flank wear estimates from the measured values seems to be influenced by cutting speed as well as by feed.

This behavior in the flank wear estimates can be attributed to crater wear which develops at higher cutting speeds. The development of crater wear causes a drop in AE energy, cutting forces and tool vibration. That means the characteristics of the signals are altered by the presence of crater wear. A neural network that is trained with signals from crater-wear-free cutting tools is not able to separate the effect of crater wear on signal characteristics. This might be the reason for the gradual under-estimation of flank wear values with progressing cutting time.

Table 9.5 gives the mean, standard deviation, and RMS of flank wear estimation errors for all sensor combinations. These parameters are graphically depicted

240

in Figure 9.28. These statistics indicate that all sensor combinations give almost the same level of mean, standard deviation, and RMS of estimation errors. These values are much higher compared with the mean, standard deviation, and RMS of estimation errors for the test data in Set 3. That means flank wear estimates are relatively less accurate at the operating conditions used for generating data in Set 3 than at the operating conditions used for generating data in Set 2. This difference in the estimation results for data in Set 2 and Set 3 indicates that the accuracy of flank wear estimates differ with respect to cutting speed and feed. The next section studies how and why the accuracy of flank wear estimates vary at different cutting speeds and feeds.

Table 9.5: Statistics of estimation errors on set 3 tools. These statistics are computed from 41 pairs of measured and estimated flank wear values using six cutting tools.

These statistics are computed from 41 pairs of measured and estimated flank wear values over six cutting tools				
Case #	Sensor Combination	Mean $\times 10^{-2}$ in.	Std. Dev. $\times 10^{-2}$ in.	RMS $\times 10^{-2}$ in.
1	AE	-0.2058	0.1827	0.27377
2	Force (F)	-0.1939	0.1984	0.27567
3	Vibration (V)	-0.2227	0.1964	0.29536
4	AE + F + V	-0.2224	0.1884	0.28998
5	AE + F	-0.2001	0.1813	0.26858
6	AE + V	-0.2337	0.1938	0.30214
7	F + V	-0.2159	0.1934	0.28831

9.5 Influence of Speed and Feed on Flank Wear Estimation

Tables 9.6, 9.7, and 9.8 list mean and standard deviation of estimation errors for each cutting tool in Set 2 and Set 3 for AE, force, and vibration sensors respec-

241

tively. These tables also list the respective cutting speed, feed, and the number of data points recorded for a cutting tool. To study the influence of speed and feed on flank wear estimation result, bubble charts are drawn. These charts are shown in Figures 9.29 to 9.34 for mean and standard deviation for AE, force, and vibration sensors. In these bubble charts, the diameter of a bubble is proportional to the mean or the standard deviation it represents. A bubble is shaded if it represents a negative value. From these bubble charts the following conclusions can be drawn:

Table 9.6: Mean and standard deviation of estimation errors at different speeds and feeds for AE sensor

Tool #	Feed ipr	Speed sfpm	Number of Points	Mean $\times 10^{-2}$ in.	Std. Dev. $\times 10^{-2}$ in.
16	0.0088	100	11	0.1936	0.0279
17	0.0112	100	12	0.2309	0.0357
18	0.0136	100	12	0.1173	0.0122
19	0.0088	130	11	0.1722	0.0299
20	0.0112	130	10	0.0943	0.0227
21	0.0136	130	11	0.1648	0.0248
22	0.0088	160	9	0.0156	0.0314
23	0.0112	160	9	0.0832	0.0280
24	0.0136	160	10	-0.0059	0.0389
25	0.0088	190	8	-0.1542	0.0512
26	0.0112	190	8	-0.0824	0.0259
27	0.0136	190	7	-0.1491	0.0322
28	0.0088	220	6	-0.3663	0.0760
29	0.0112	220	6	-0.1611	0.0589
30	0.0136	220	6	-0.3896	0.0993

1. At cutting speeds 100, 130, and 160 sfpm (i.e., for data in Set 2) for AE, force, and vibration sensors, the magnitude of both the mean and the standard deviation of estimation errors show no particular increasing or decreasing pattern with respect to either cutting speed or feed.

242

Table 9.7: Mean and standard deviation of estimation errors at different speeds and feeds for force sensors

Tool #	Feed ipr	Speed sfpm	Number of Points	Mean $\times 10^{-2}$ in.	Std. Dev. $\times 10^{-2}$ in.
16	0.0088	100	11	-0.0589	0.0376
17	0.0112	100	12	-0.1296	0.0451
18	0.0136	100	12	-0.1043	0.0450
19	0.0088	130	11	0.1233	0.0380
20	0.0112	130	10	-0.0643	0.0195
21	0.0136	130	11	0.0342	0.0221
22	0.0088	160	9	-0.0255	0.0257
23	0.0112	160	9	0.0638	0.0298
24	0.0136	160	10	0.0317	0.0311
25	0.0088	190	8	-0.0032	0.0415
26	0.0112	190	8	-0.1591	0.0420
27	0.0136	190	7	-0.1930	0.0458
28	0.0088	220	6	-0.3813	0.1043
29	0.0112	220	6	-0.2186	0.0902
30	0.0136	220	6	-0.2835	0.0724

2. At cutting speeds 190 and 220 sfpm (i.e., for data in Set 3) for AE, force, and vibration sensors, the mean of estimation errors consistently becomes negative.

3. At cutting speeds 190 and 220 sfpm for AE and force sensors, the magnitude of both the mean and the standard deviation of estimation errors shows an increasing trend with respect to cutting, speed but no regular pattern with respect to feed.

4. At cutting speeds 190 and 220 sfpm for AE and vibration sensors, the magnitude of both the mean and the standard deviation of estimation errors show a gradually increasing trend with respect to cutting speed as well as feed.

243

Table 9.8: Mean and standard deviation of estimation errors at different speeds and feeds for vibration sensors

Tool #	Feed ipr	Speed sfpm	Number of Points	Mean $\times 10^{-2}$ in.	Std. Dev. $\times 10^{-2}$ in.
16	0.0088	100	11	-0.0407	0.0378
17	0.0112	100	12	-0.0686	0.0334
18	0.0136	100	12	0.0604	0.0201
19	0.0088	130	11	0.2117	0.0547
20	0.0112	130	10	0.0733	0.0385
21	0.0136	130	11	0.0994	0.0333
22	0.0088	160	9	0.0251	0.0232
23	0.0112	160	9	-0.0270	0.0200
24	0.0136	160	10	0.0110	0.0188
25	0.0088	190	8	-0.1129	0.0329
26	0.0112	190	8	-0.1199	0.0265
27	0.0136	190	7	-0.2070	0.0633
28	0.0088	220	6	-0.2907	0.0849
29	0.0112	220	6	-0.3191	0.0842
30	0.0136	220	6	-0.3600	0.1248

Two main reasons can be ascribed to this discrepancies between the flank wear estimation results at cutting speeds 100, 130, 160 sfpm and 190, 220 sfpm:

1. Neural networks used for predicting flank wear are trained with the signals collected at operating conditions 100, 130, and 160 sfpm but not with the signals collected at the operating conditions 190 and 220 sfpm. That means that the neural networks used in this application are not able to give correct flank wear estimates for the operating conditions outside those used for generating the training data. This pattern in estimation results is expected because neural networks, in general, are good at interpolation but not at extrapolation.

2. The cutting tools develop crater wear in addition to flank wear at the cutting speeds 190 and 220 sfpm. As explained in a previous section, the development

244

of crater wear causes a drop in AE energy, cutting forces, and tool vibration. This will alter the sensor-signal characteristics accordingly. A neural network that is trained with signals from crater-wear-free cutting tools will misinterpret the drop in the level of sensor signals to the reduction in flank wear. This explanation is further supported by the observation that flank wear estimates are closer to the measured values at the initial stage of the tool life when the crater wear is usually absent.

9.6 Conclusions

This chapter analyzed the flank wear estimation results for different sensor combinations using two test data sets. The three parameters—mean, standard deviation, and RMS of estimation—were used as the point estimators for evaluating the estimation results for different sensor combinations. The confidence intervals and quantile limits were used as the interval estimators for evaluating the same results. From the analysis of the results, the following conclusions can be drawn:

1. The proposed flank wear estimation method provides accurate flank wear estimates within the range of operating conditions that were used during neural network training data generation. However, the development of crater wear on a cutting tool will adversely affect the accuracy of flank wear estimates.

2. Among all possible sensor combinations, AE-force sensors provide the most accurate and consistent flank wear estimates. In contrast, force-vibration sensors provide the least accurate flank wear estimates.

3. The accuracy of flank wear estimates indicate that discrete wavelet transforms are very effective for representing AE signals, and univariate ARMA models are more appropriate than multivariate ARMA models for representing force

245

and vibration signals.

4. The RMS of estimation errors for all sensor combinations fall in range of 0.0010 to 0.0017 inches. This is a considerably low estimation error, making the proposed estimation method very attractive to the real-world applications.

The results that are presented in this chapter cannot be generalizable to other workpiece-tool combinations. However, the flank wear estimation methodology is generalizable to any tooling setup and workpiece-tool combination.

246
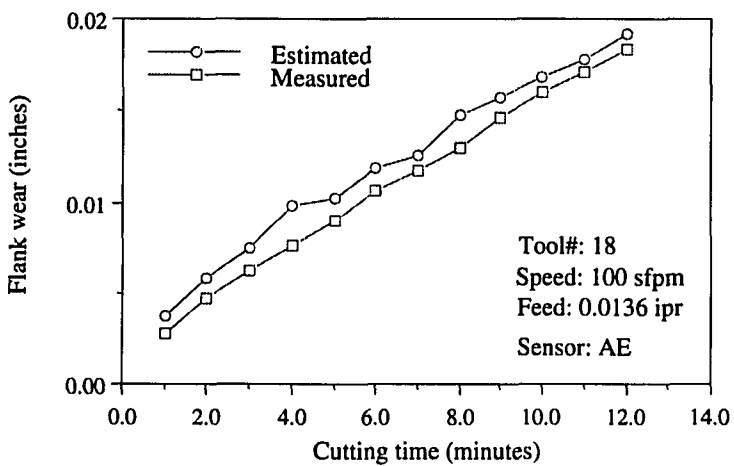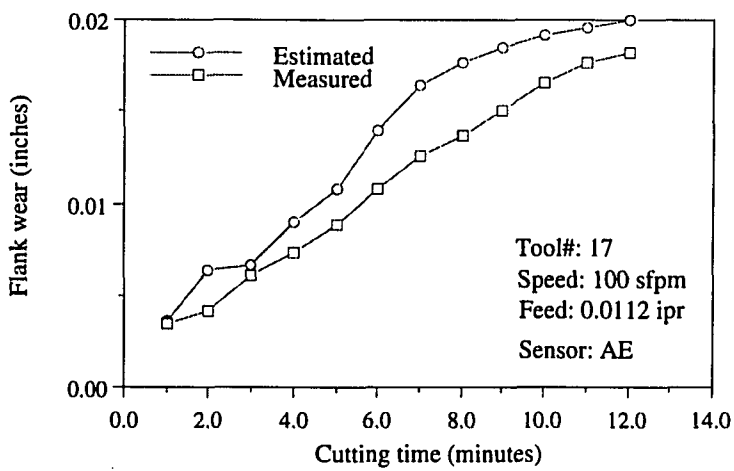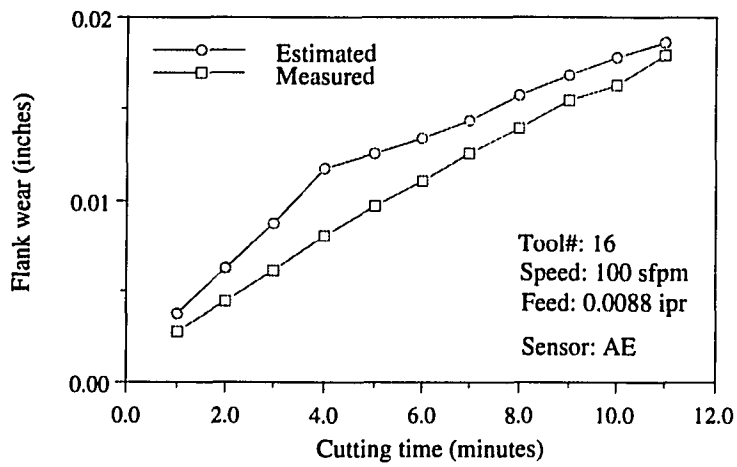
Figure 9.1: Flank wear estimation results using AE measurements

(Figure Continues)

247

Figure 9.1: Continued

248

Figure 9.1: Continued

249
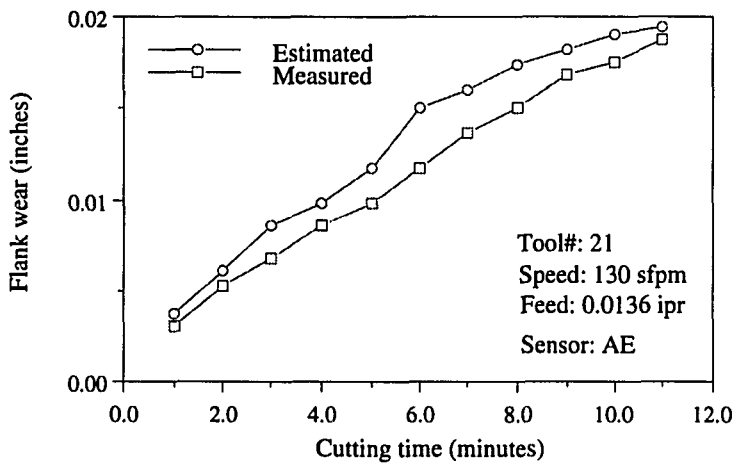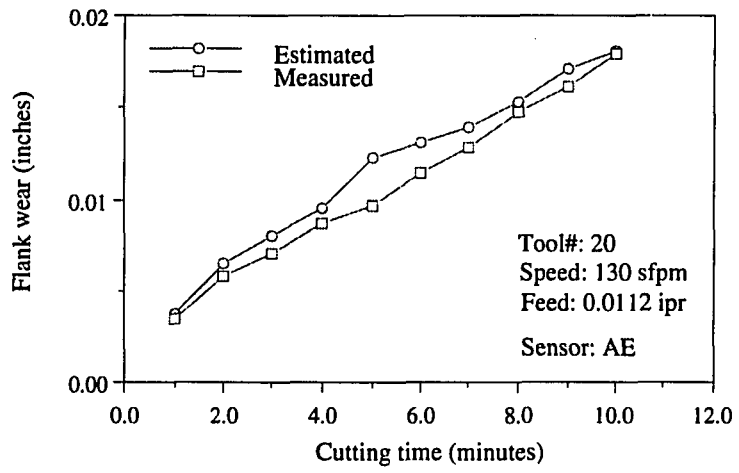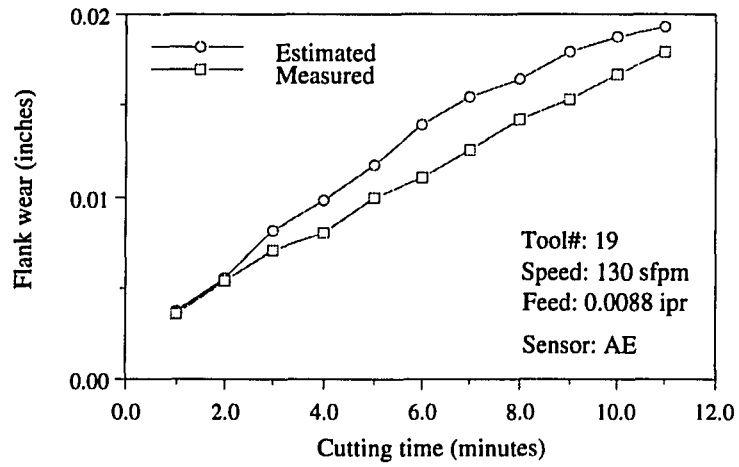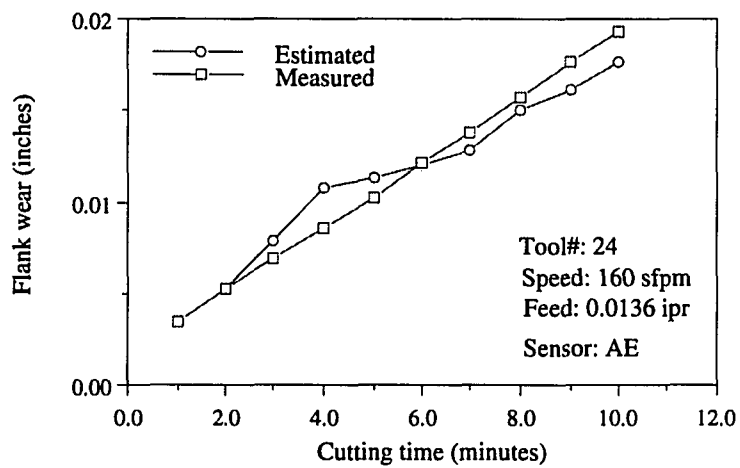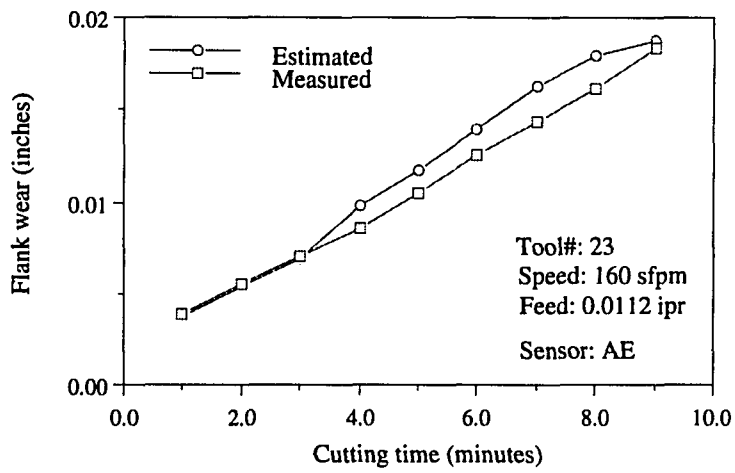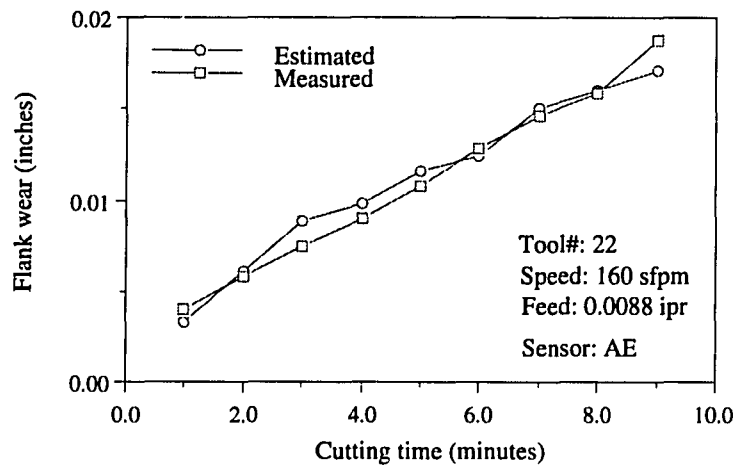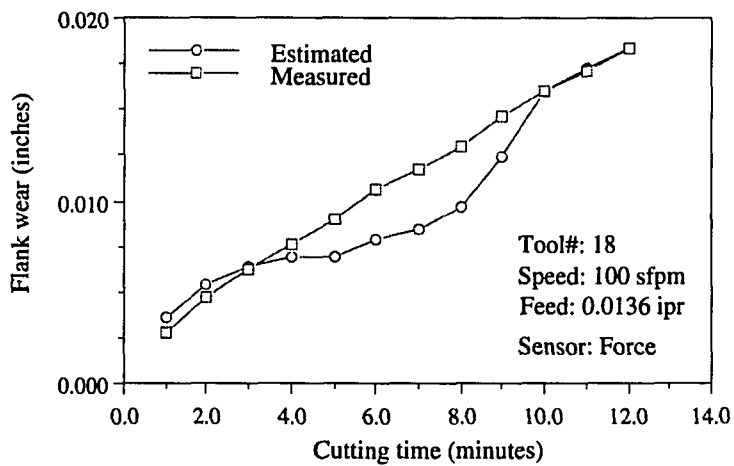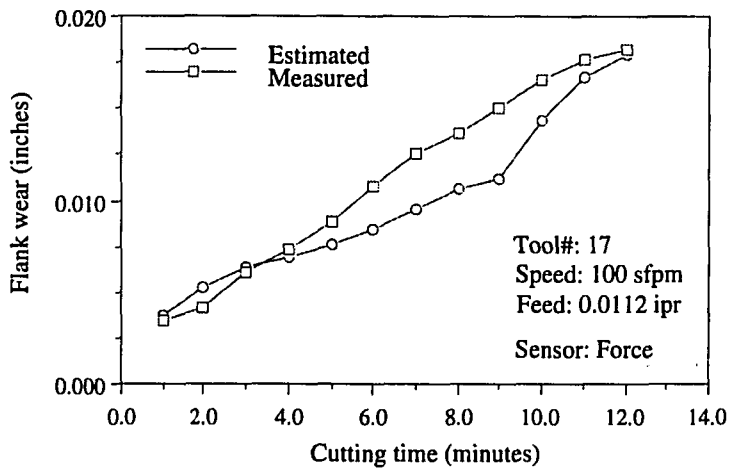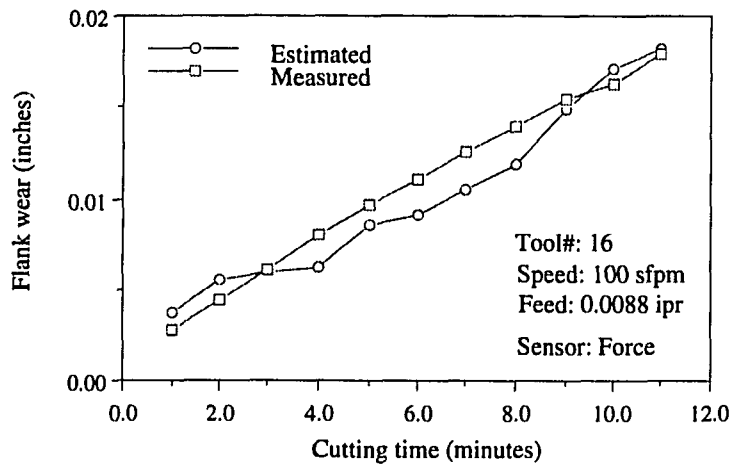
Figure 9.2: Flank wear estimation results using force measurements

(Figure Continues)

250

Figure 9.2: Continued

251

Figure 9.2: Continued

252
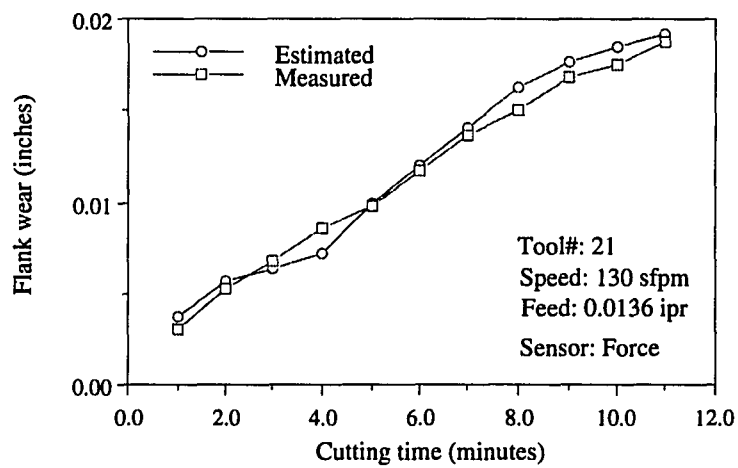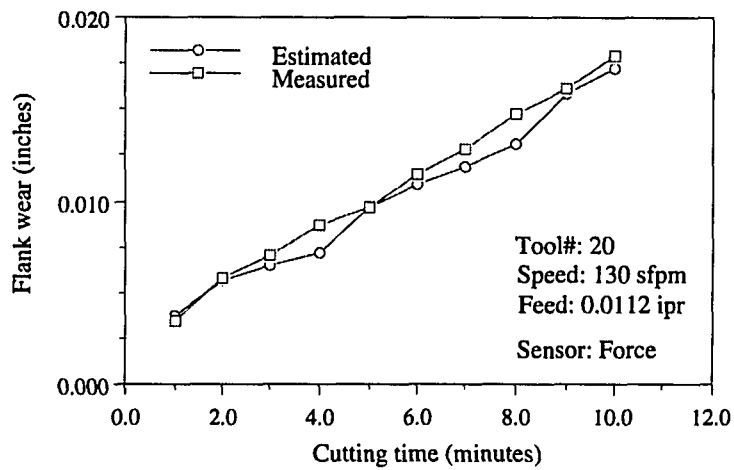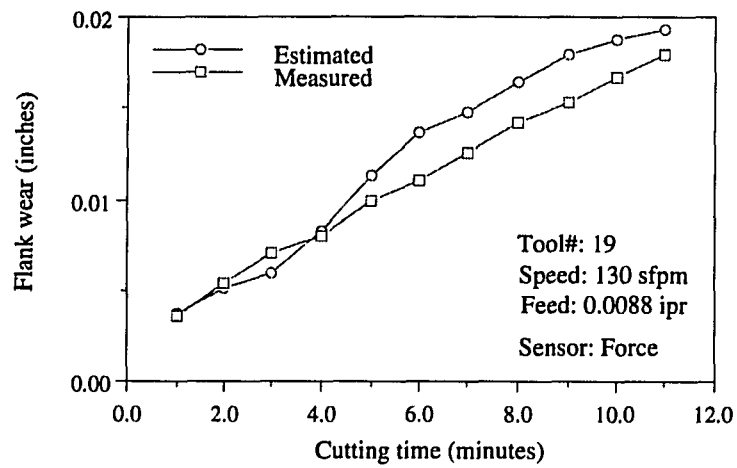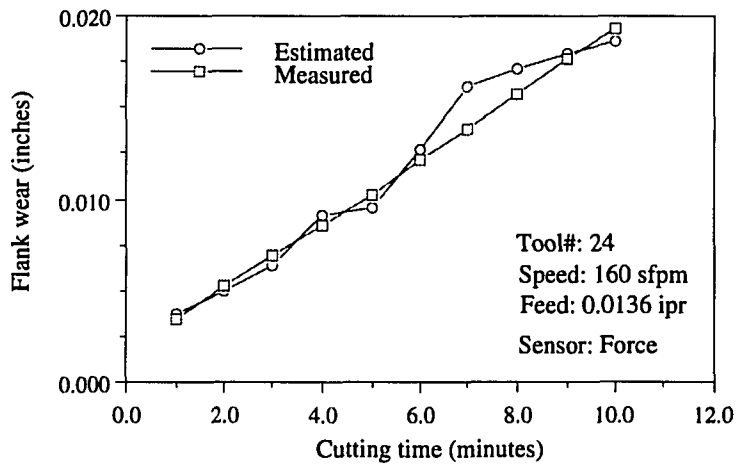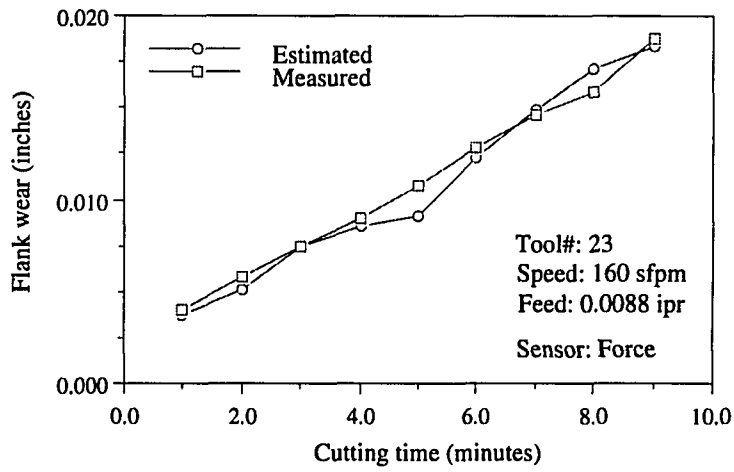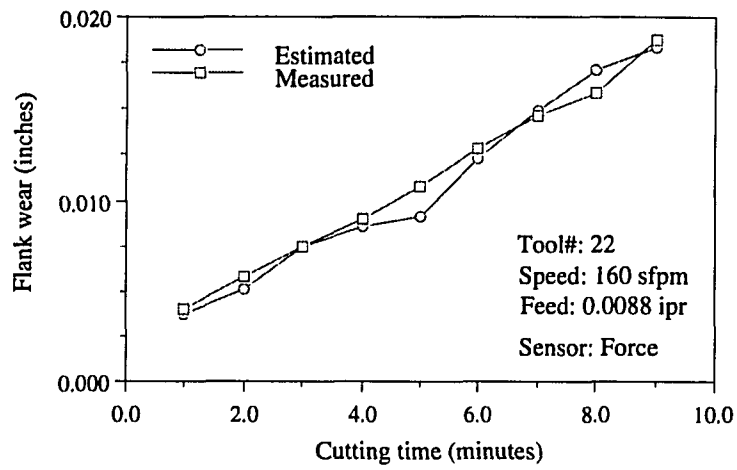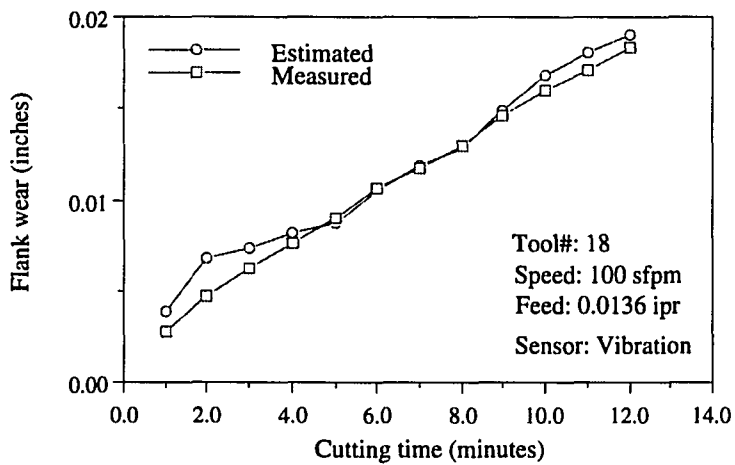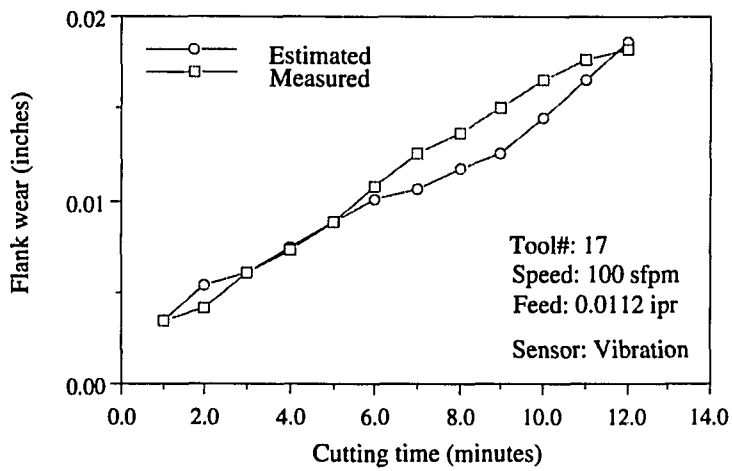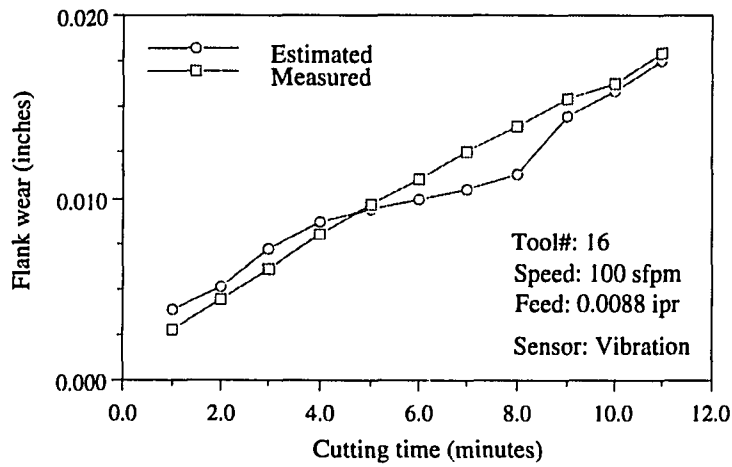
Figure 9.3: Flank wear estimation results using vibration measurements

(Figure Continues)

253

Figure 9.3: Continued

254

Figure 9.3: Continued

255
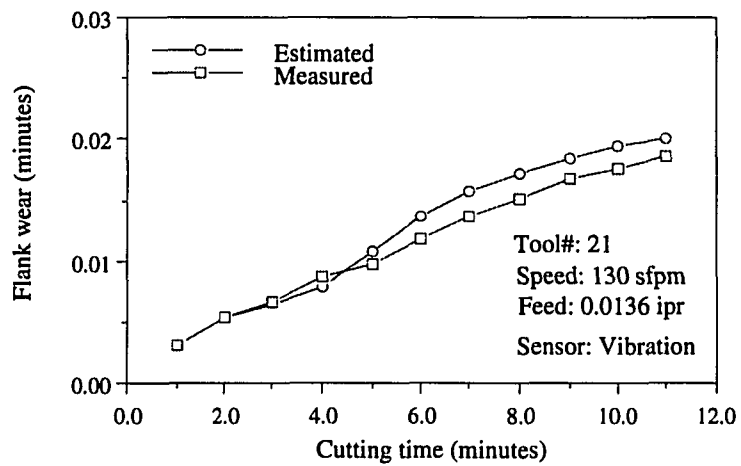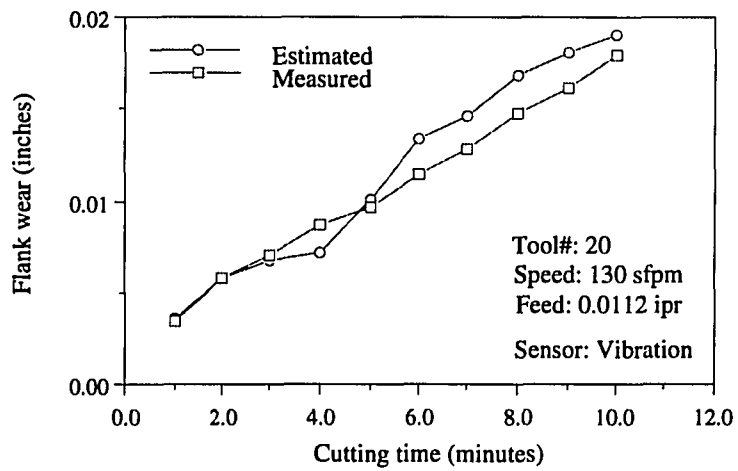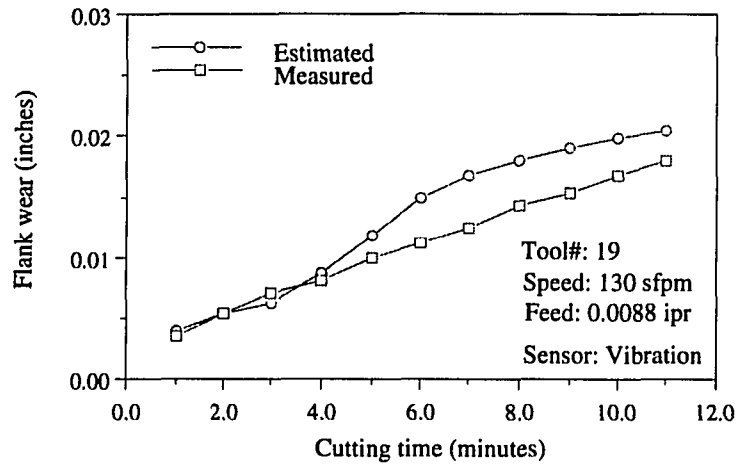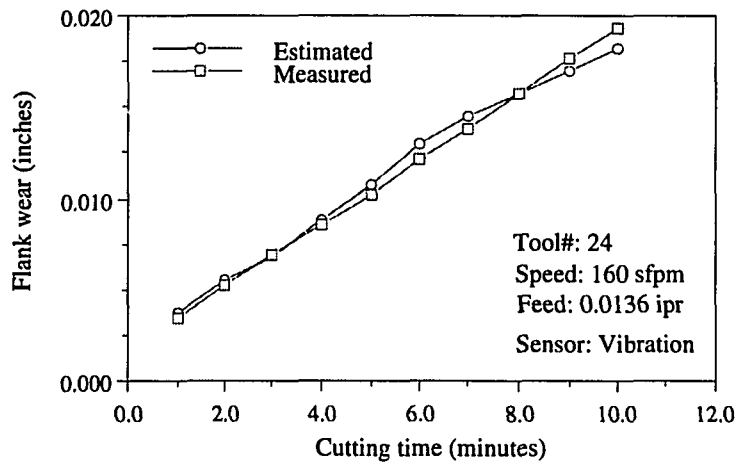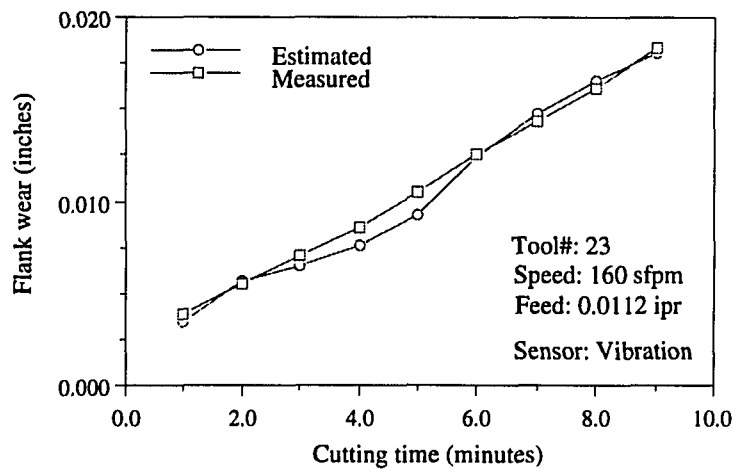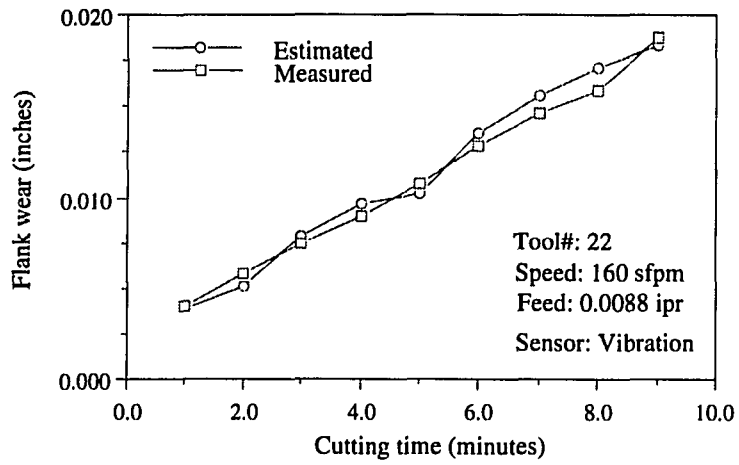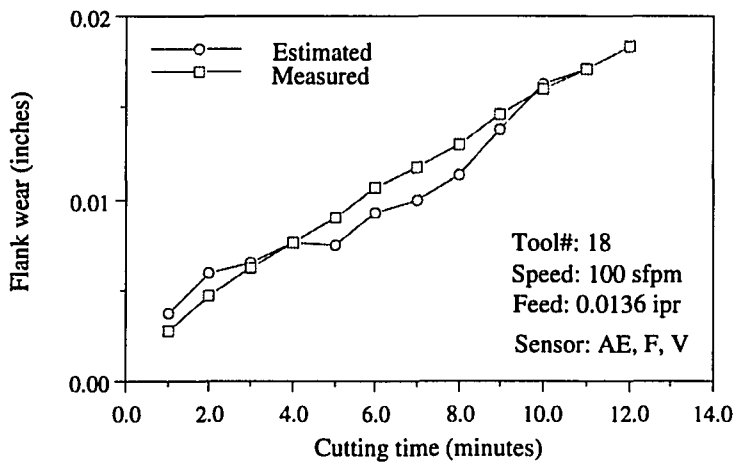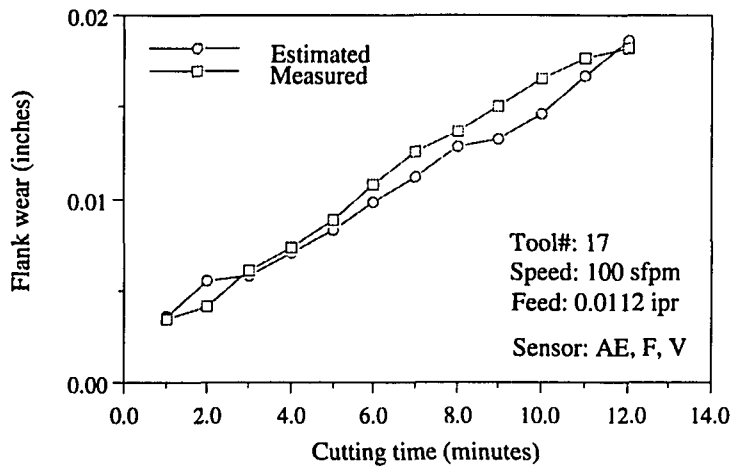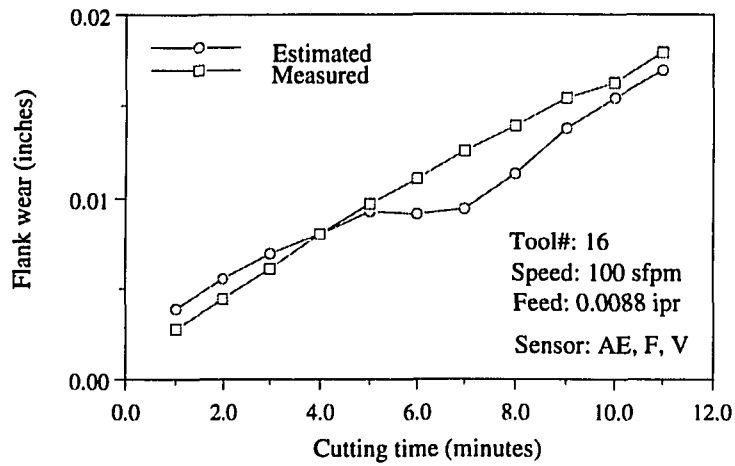
Figure 9.4: Flank wear estimation results using AE, force, and vibration measurements

(Figure Continues)

256

Figure 9.4: Continued

257

Figure 9.4: Continued

258
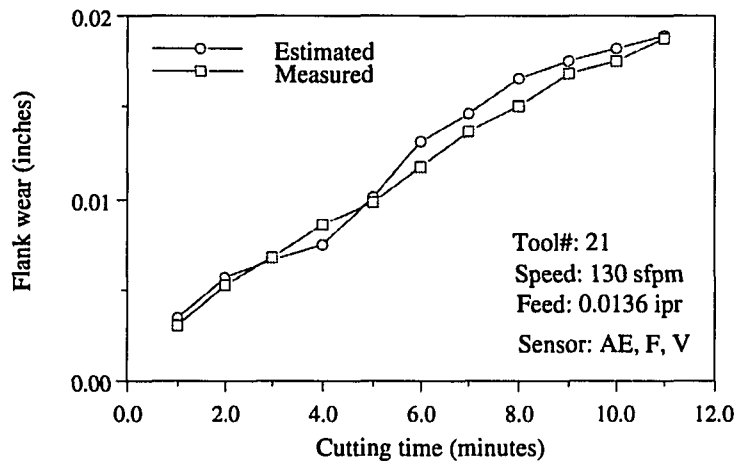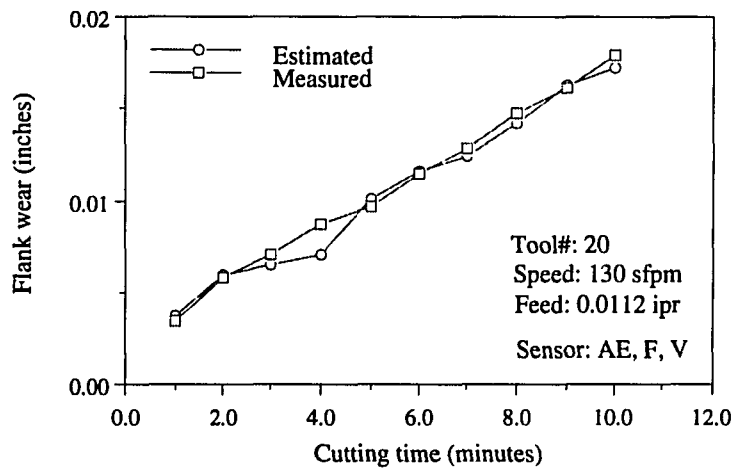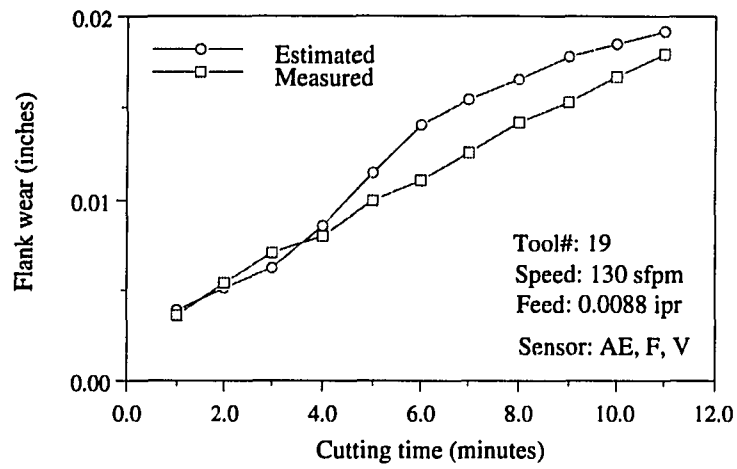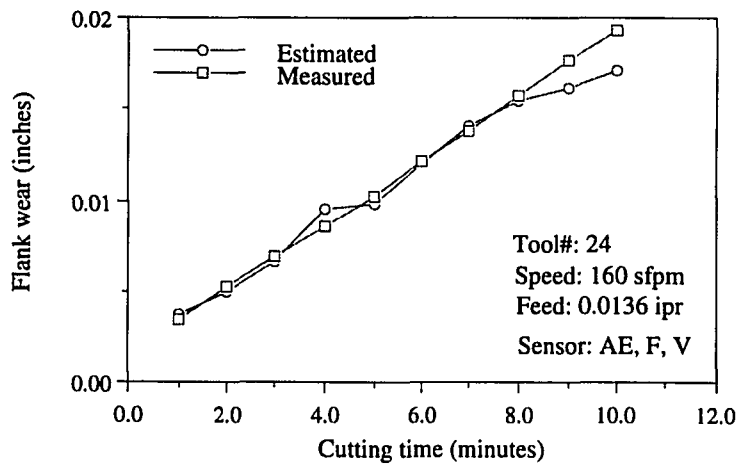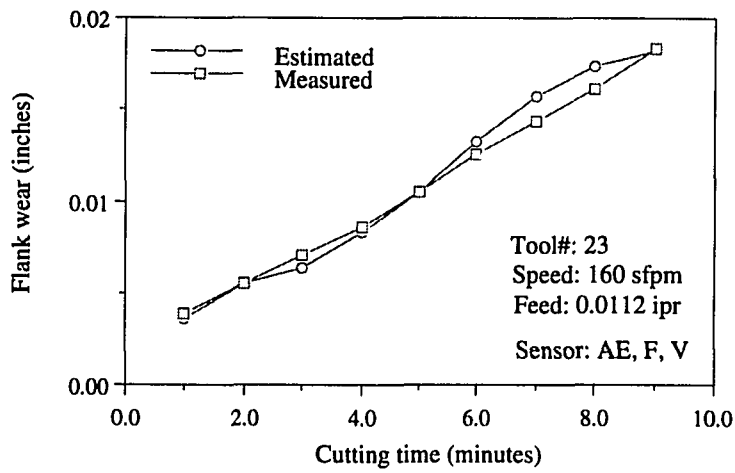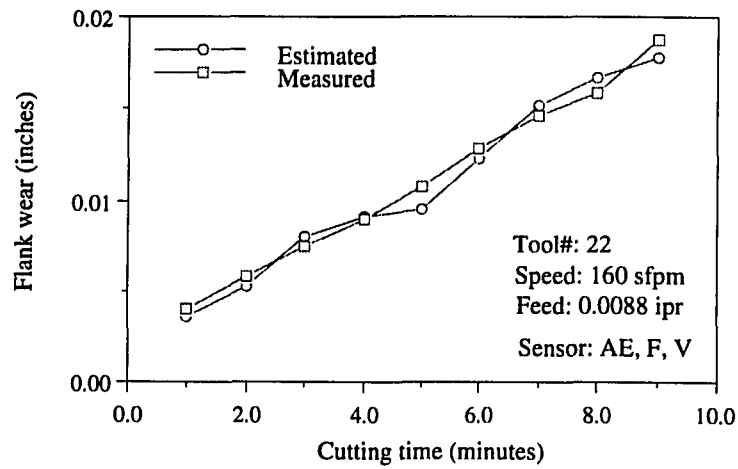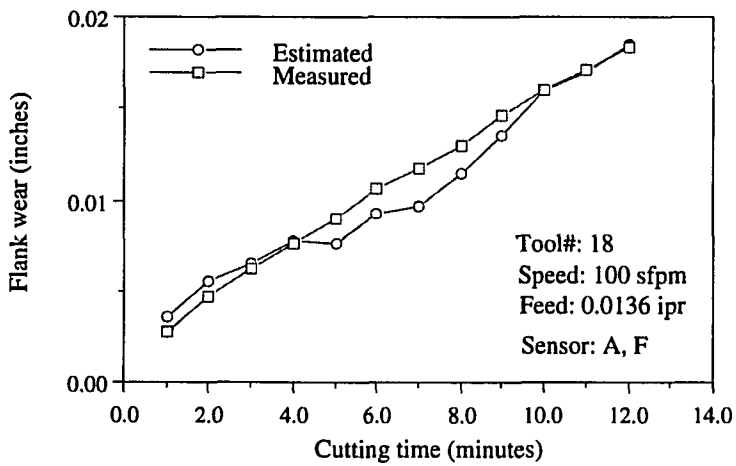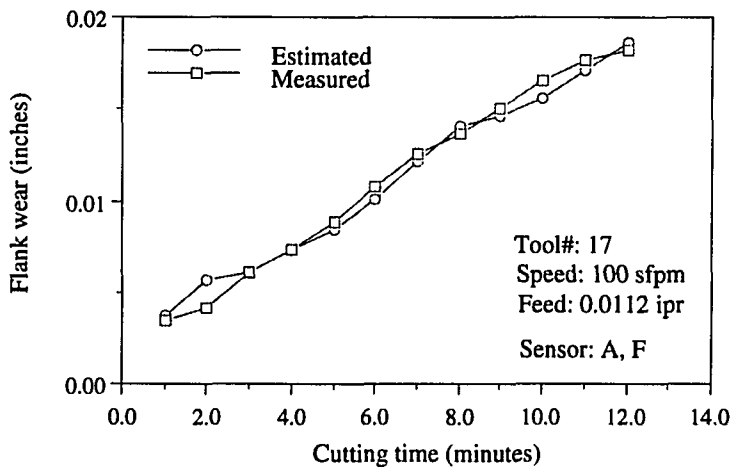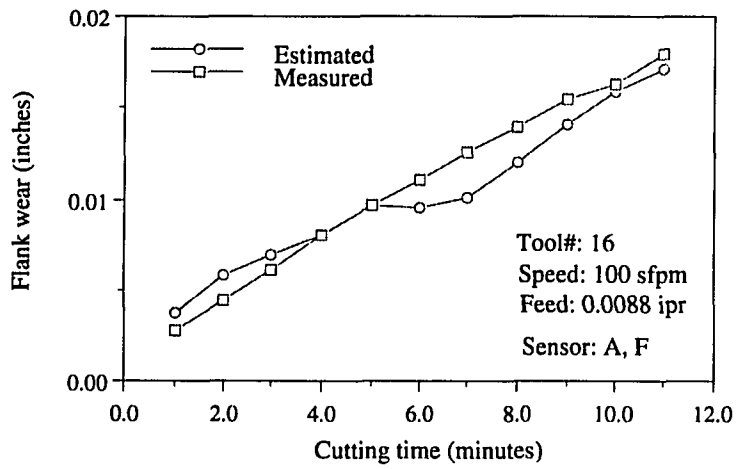
Figure 9.5: Flank wear estimation results using AE and force measurements

(Figure Continues)

259

Figure 9.5: Continued

260

Figure 9.5: Continued

261
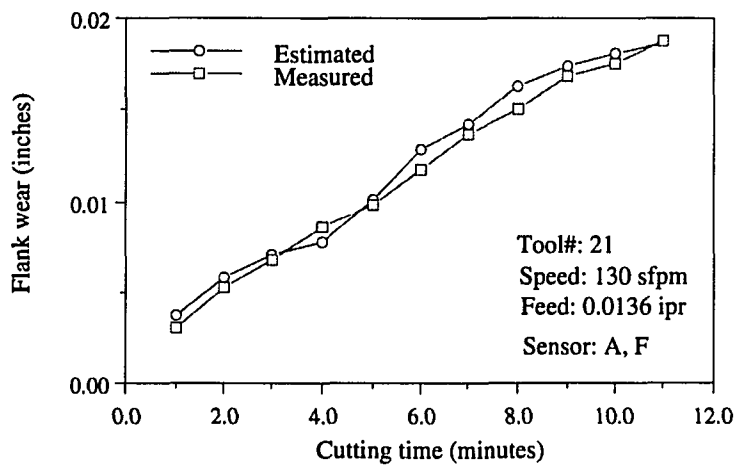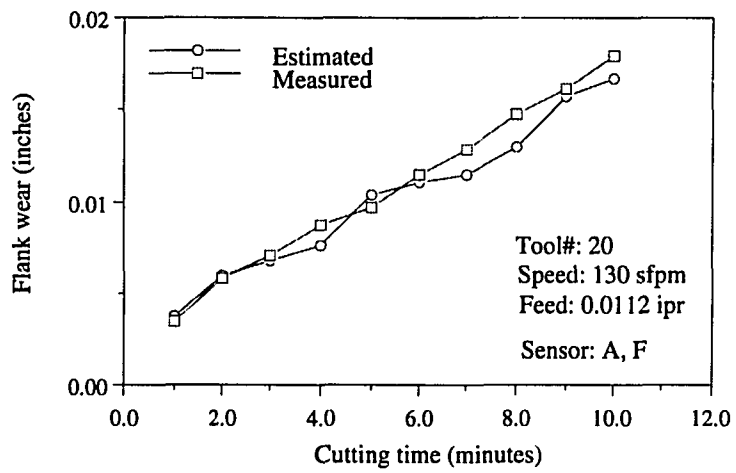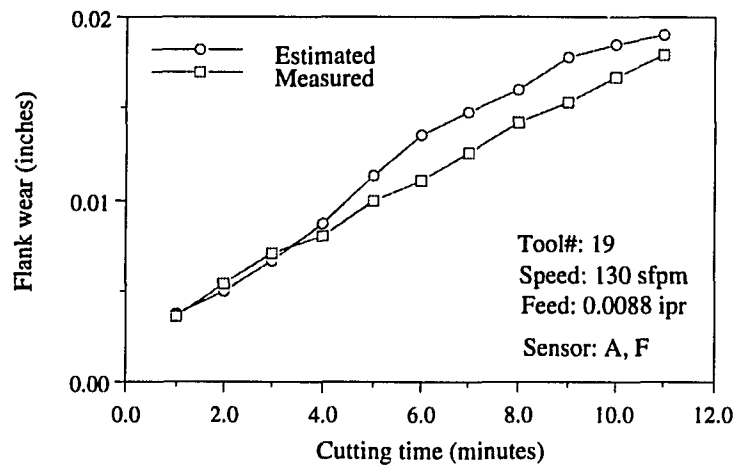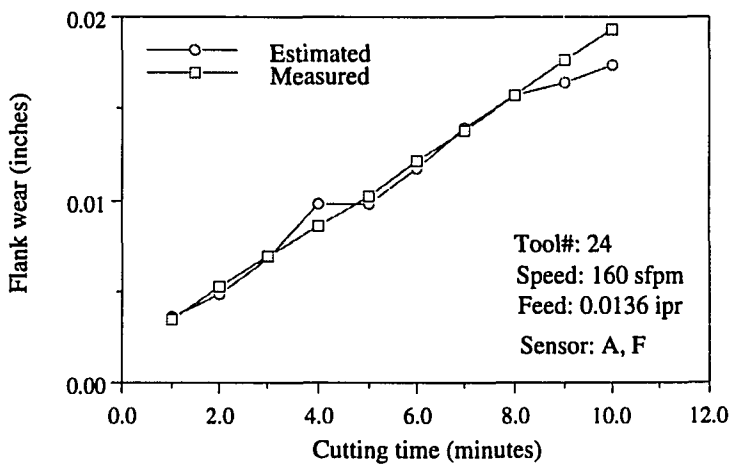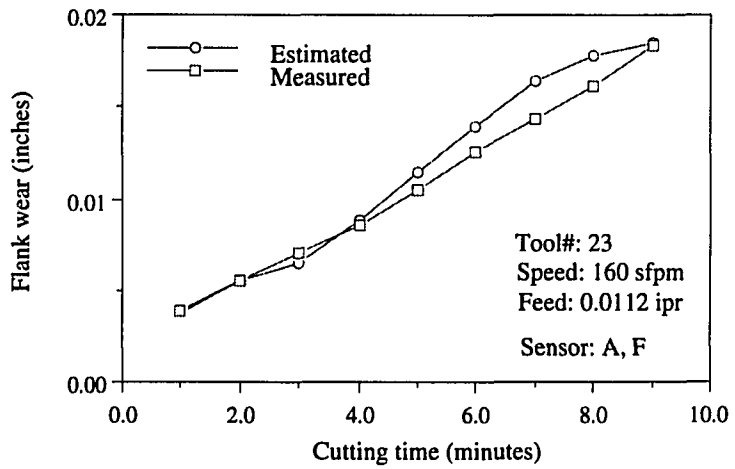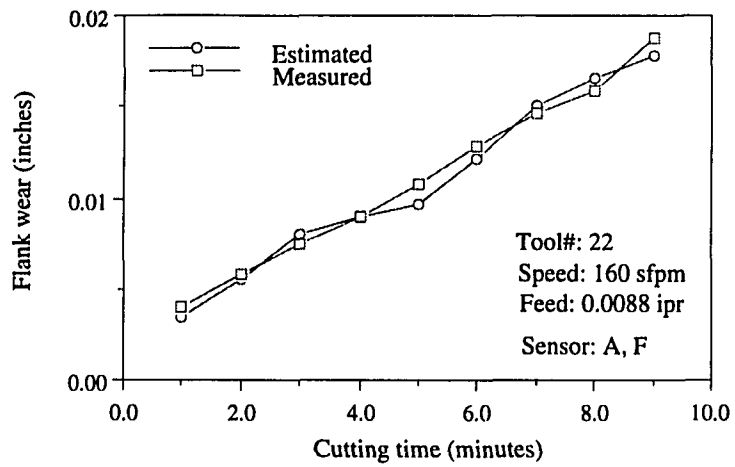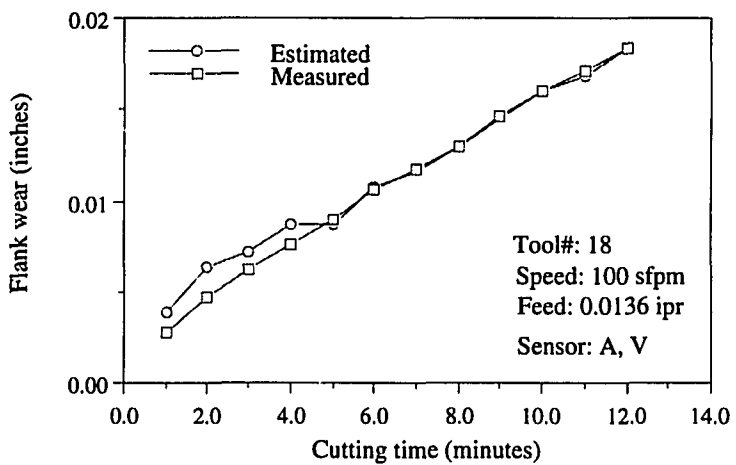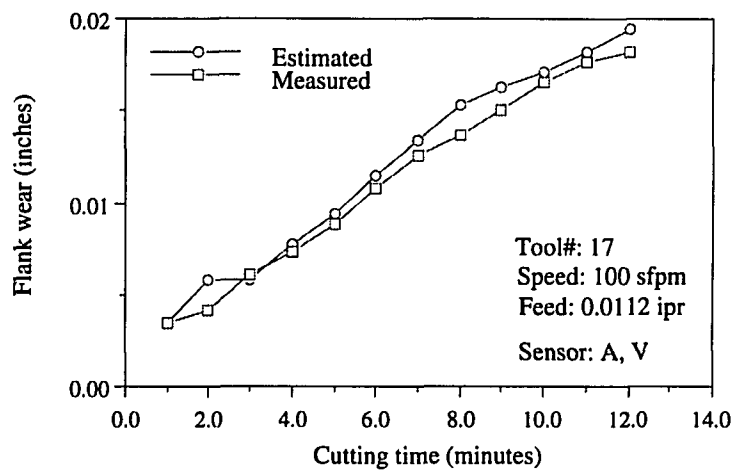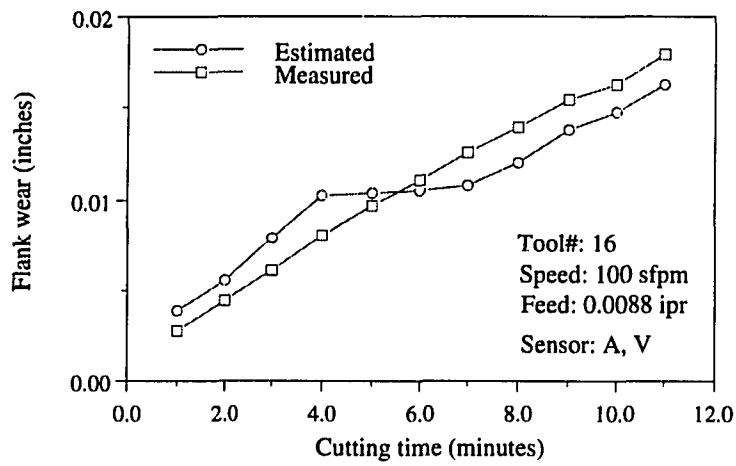
Figure 9.6: Flank wear estimation results using AE and vibration measurements

(Figure Continues)

262

Figure 9.6: Continued

263

Figure 9.6: Continued

264
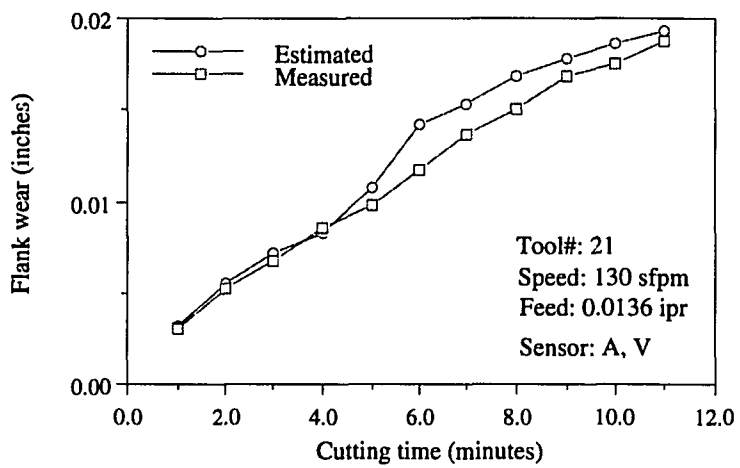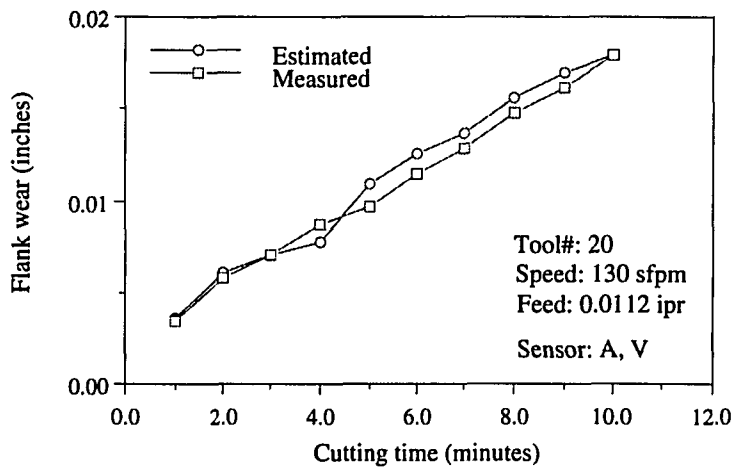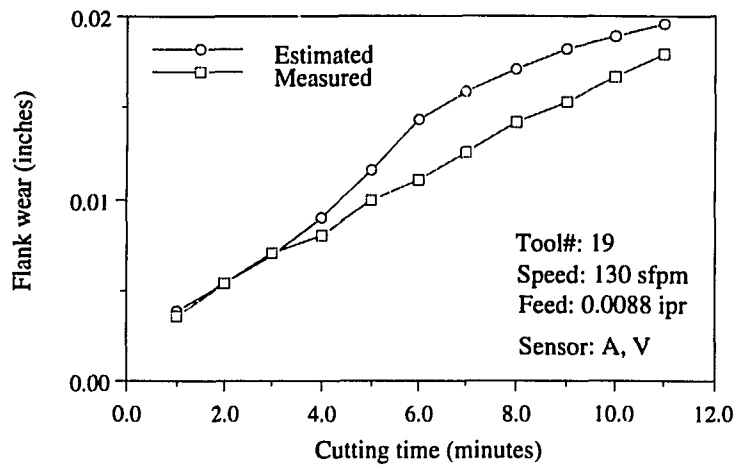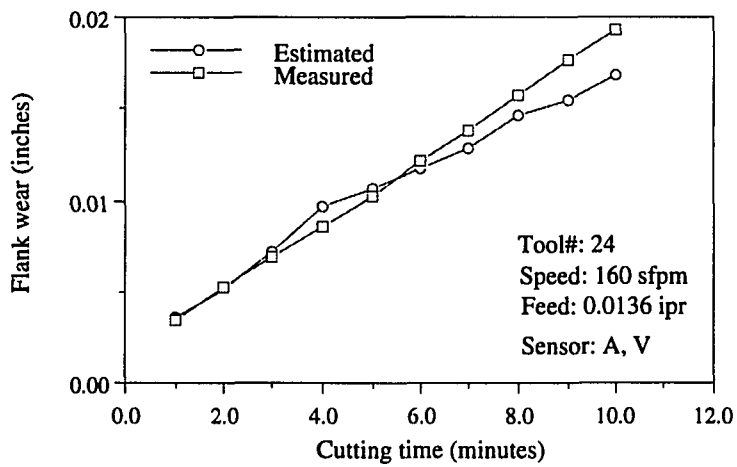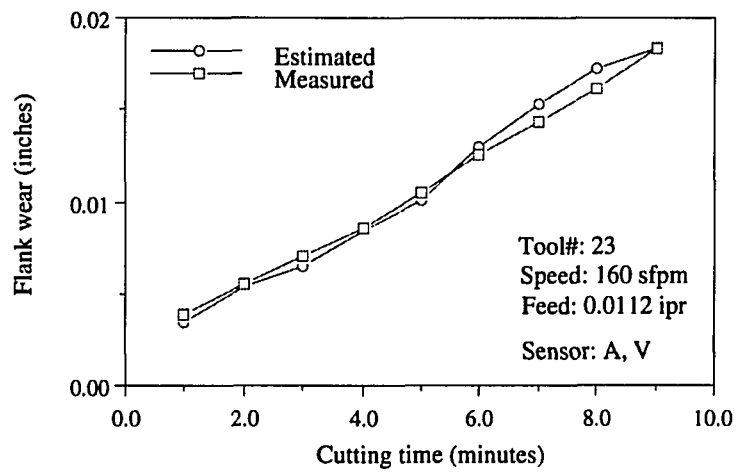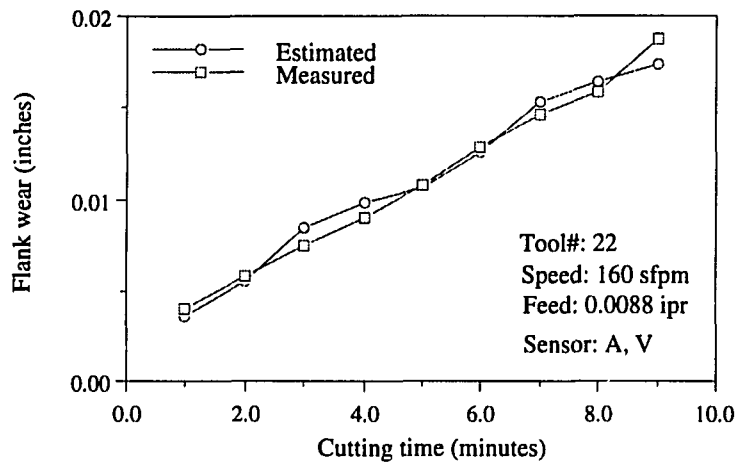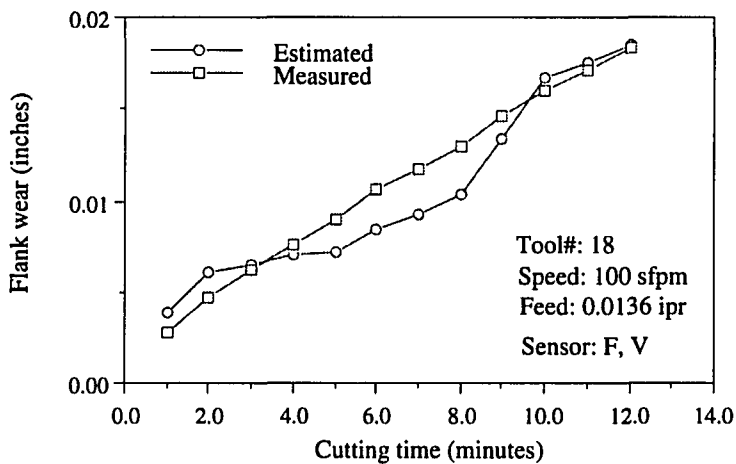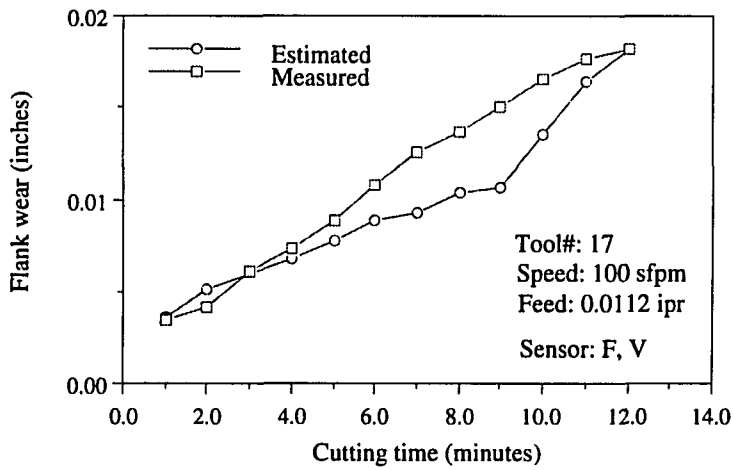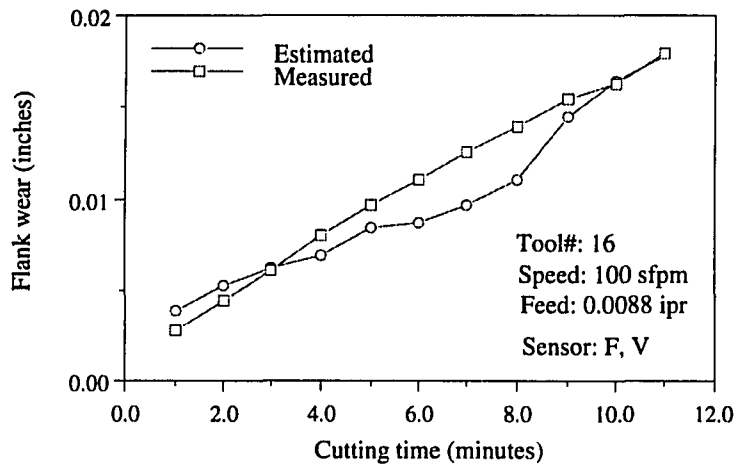
Figure 9.7: Flank wear estimation results using force and vibration measurements through univariate ARMA models

(Figure Continues)

265

Figure 9.7: Continued

266

Figure 9.7: Continued

267
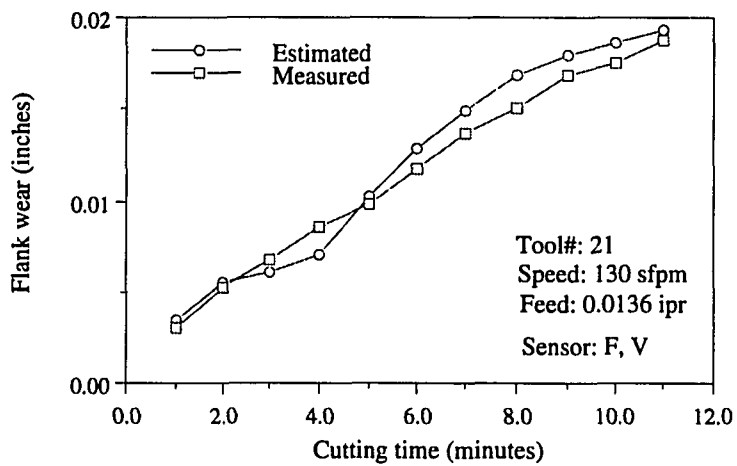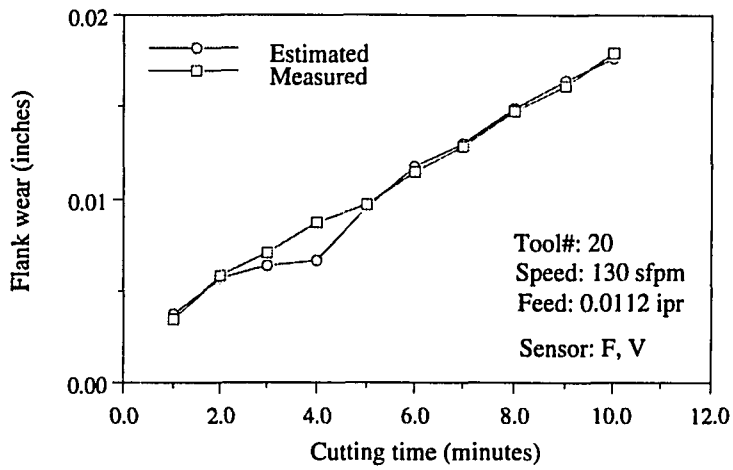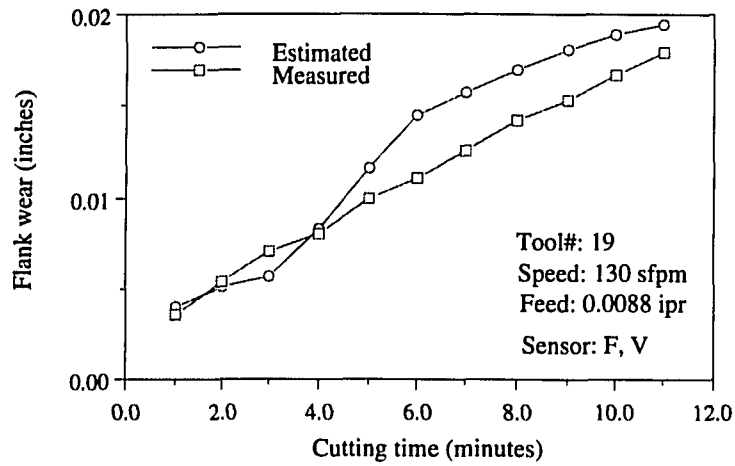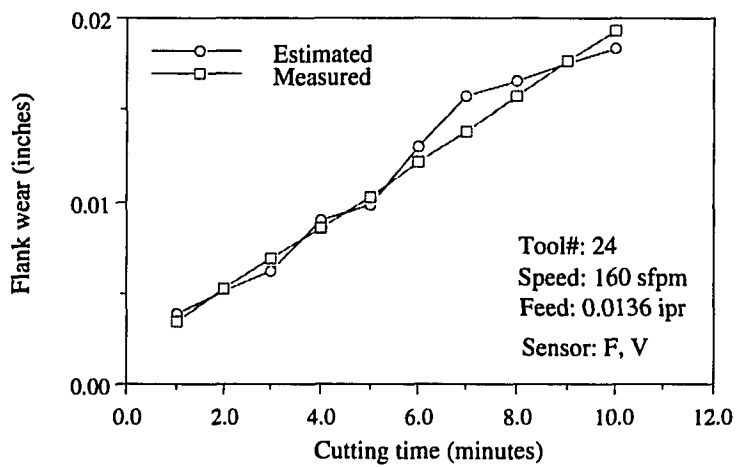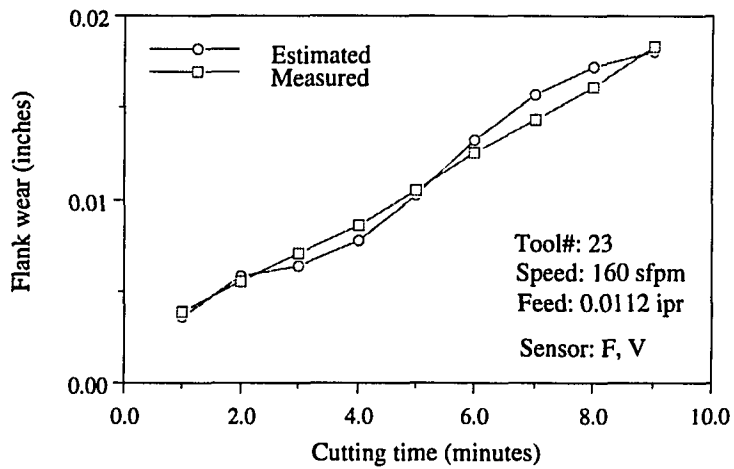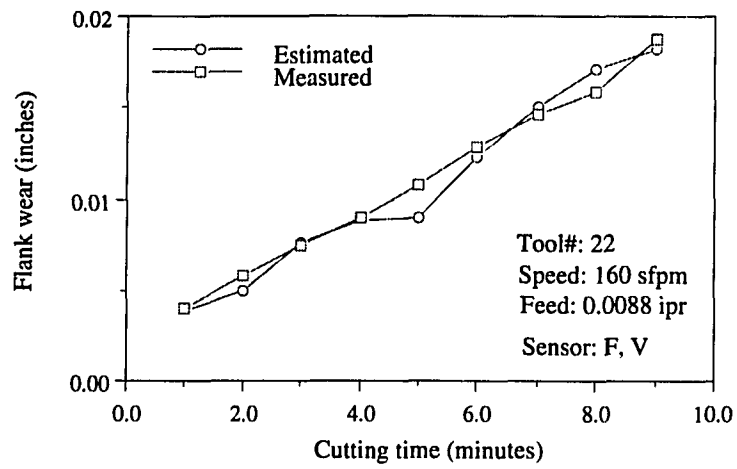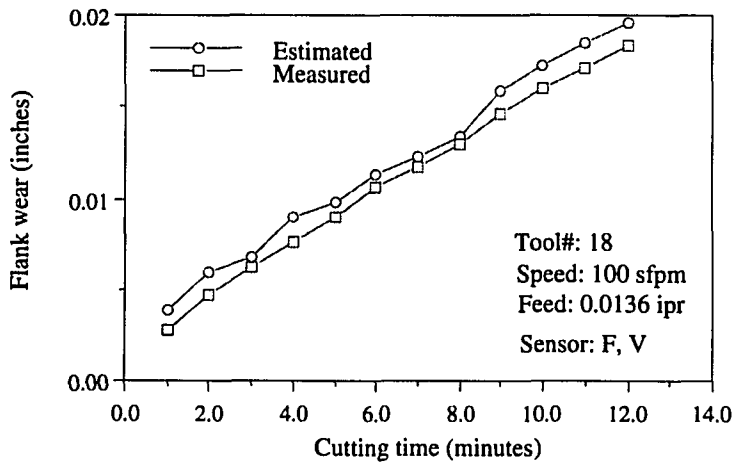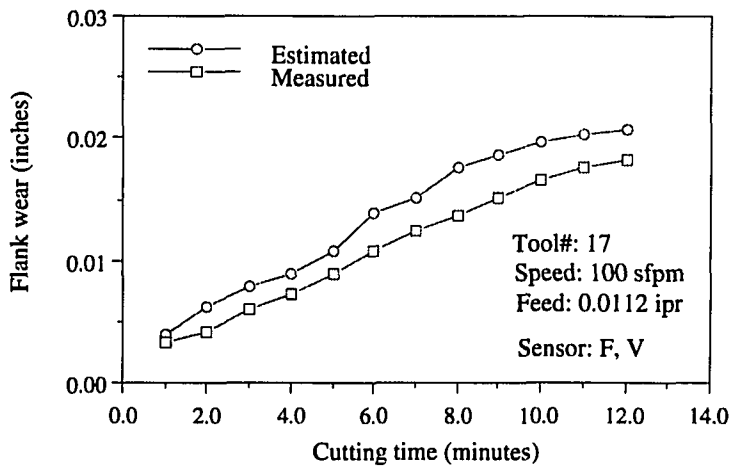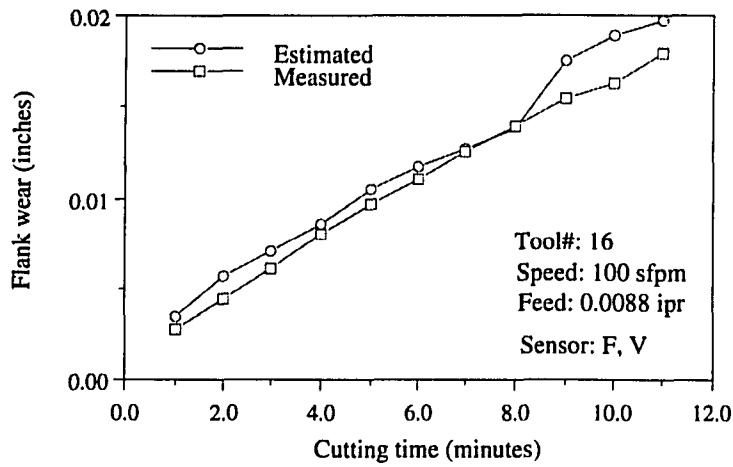
Figure 9.8: Flank wear estimation results using force and vibration measurements through multivariate ARMA models

(Figure Continues)

268

Figure 9.8: Continued

269

Figure 9.8: Continued

270
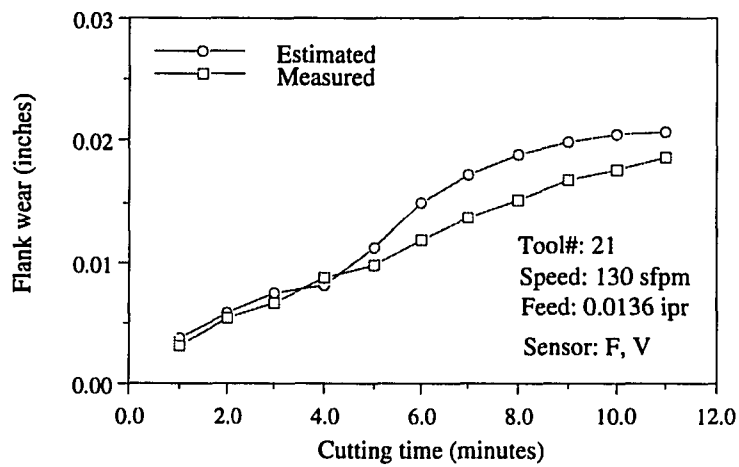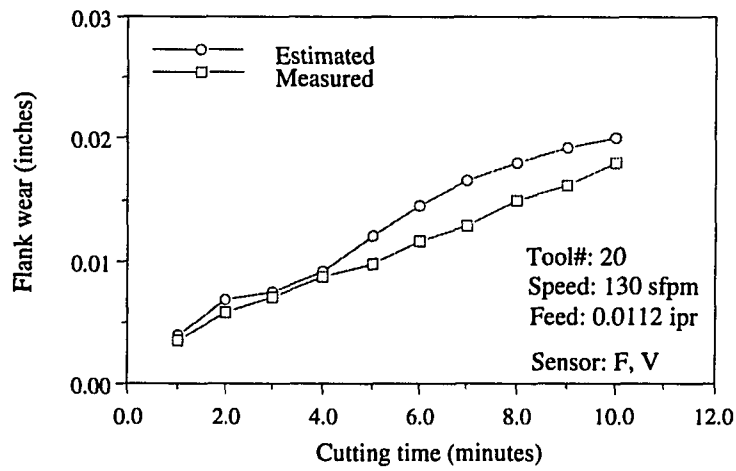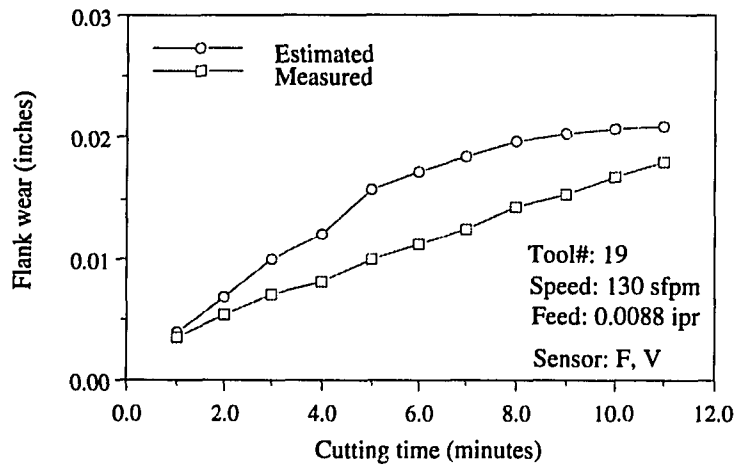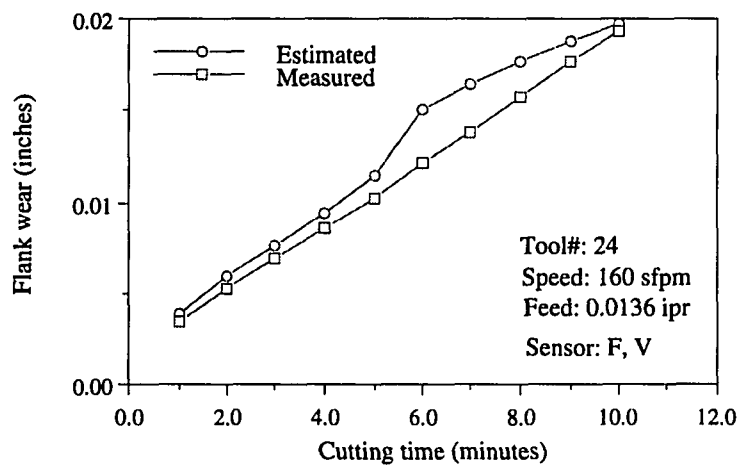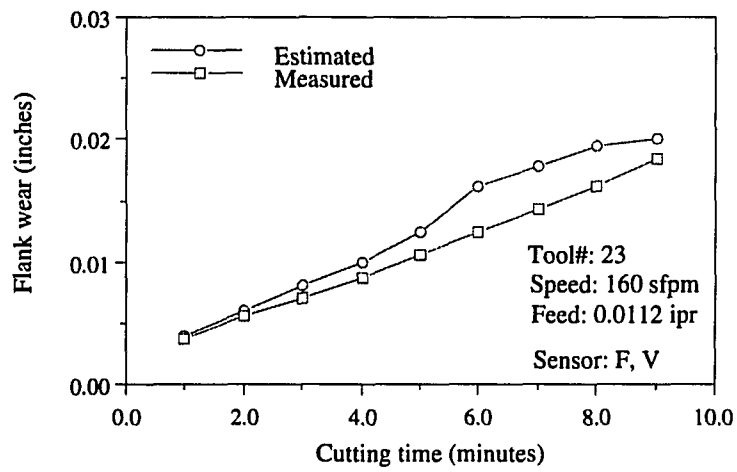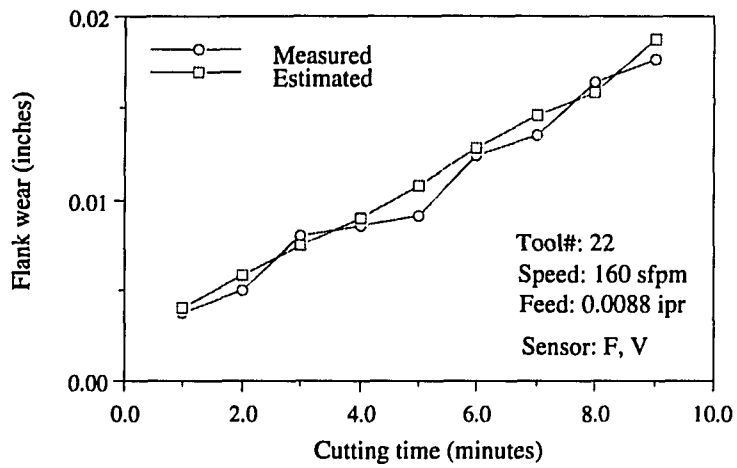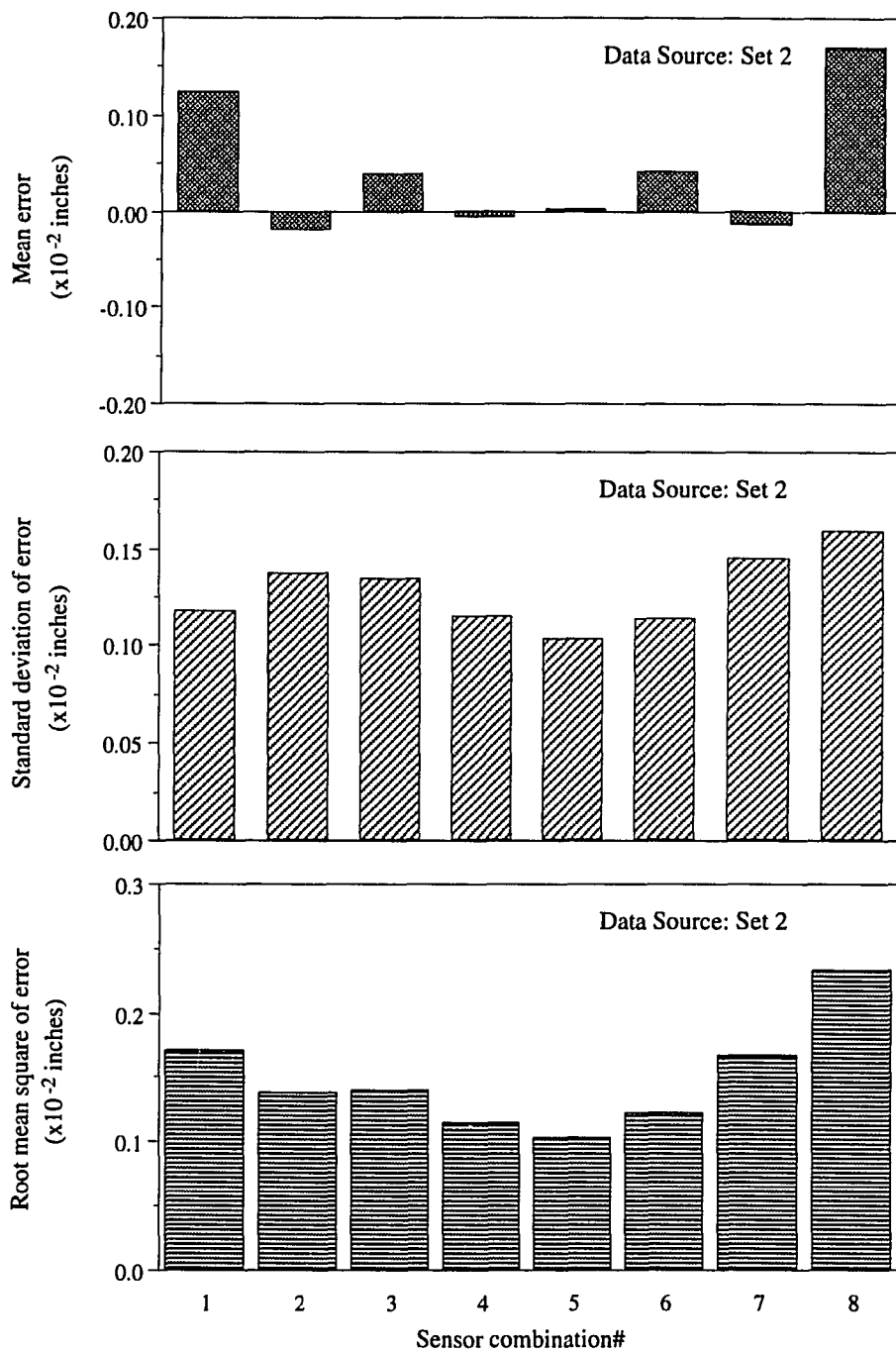
Figure 9.9: Mean, standard deviation, and RMS of
estimation errors for data in Set 2

271
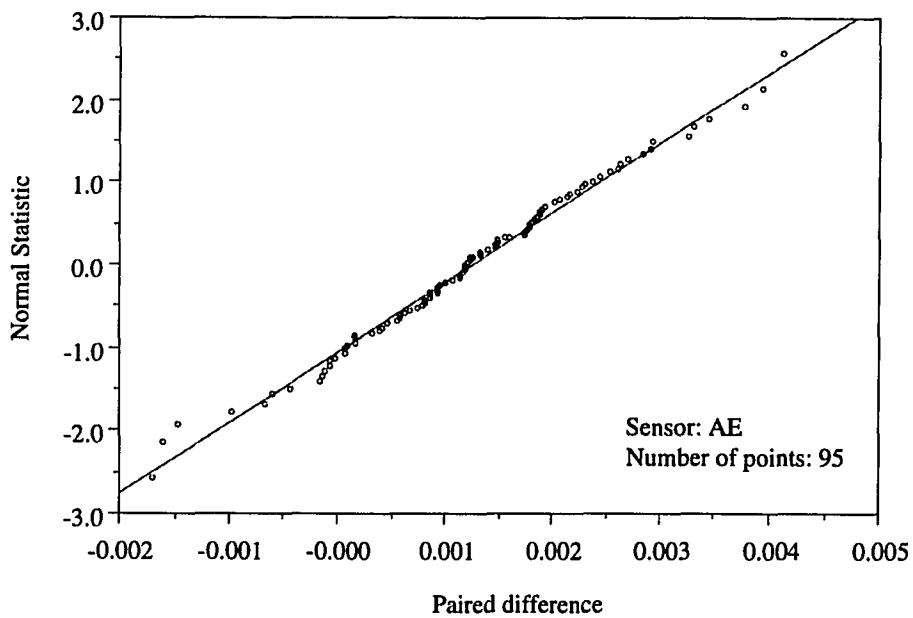
Figure 9.10: Normal probability plot of estimation errors for AE sensor

272
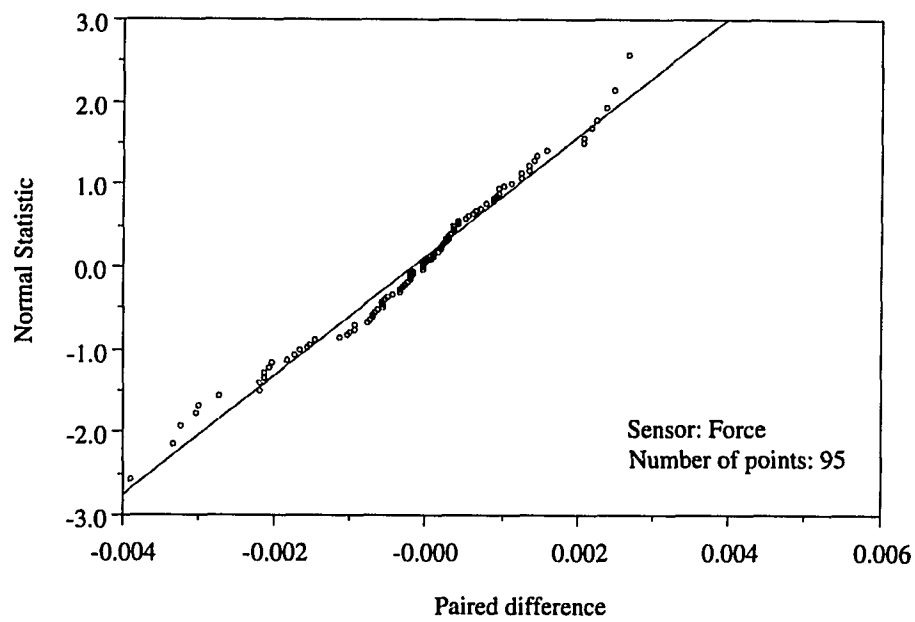
Figure 9.11: Normal probability plot of estimation errors for force sensors

273
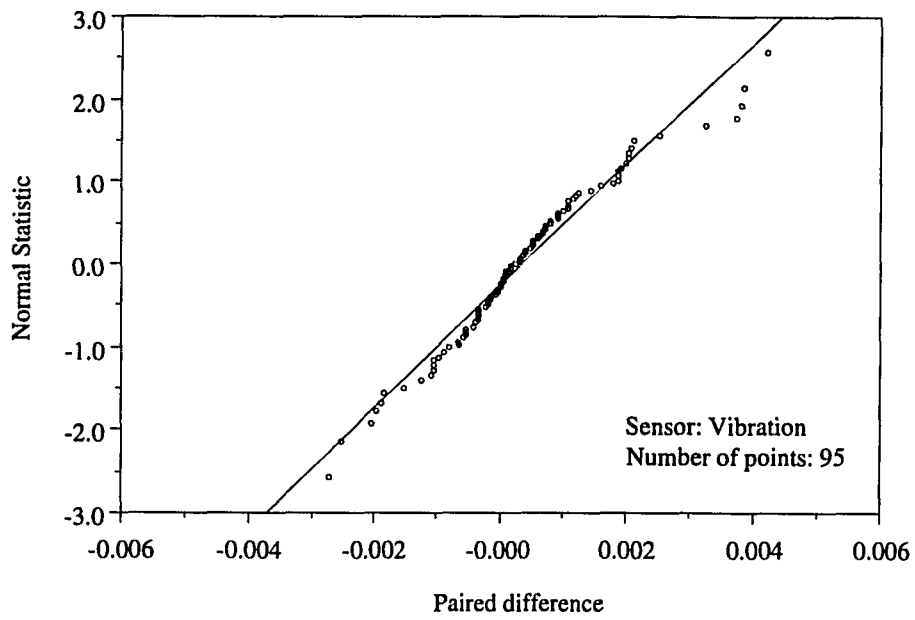
Figure 9.12: Normal probability plot of estimation errors for vibration sensors

274
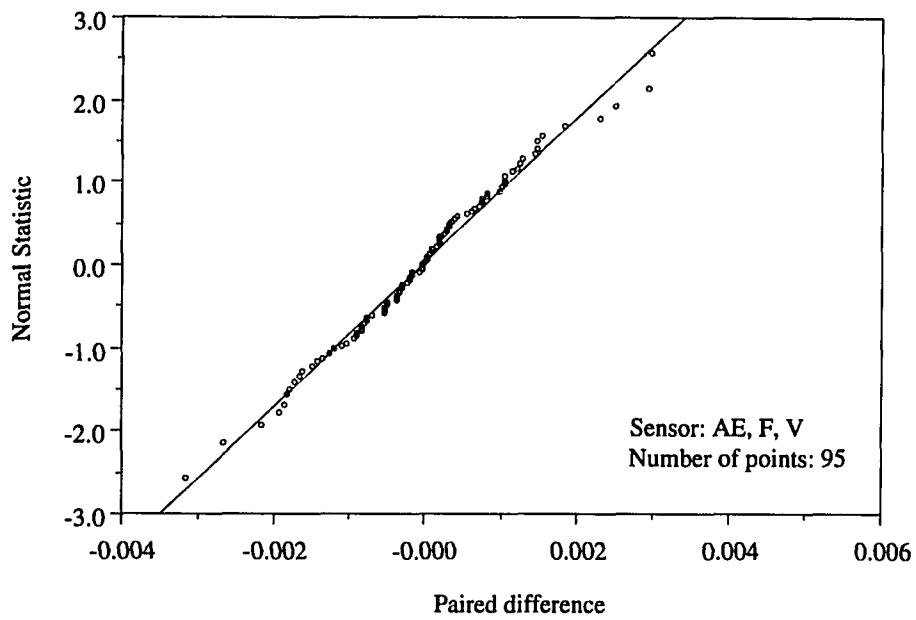
Figure 9.13: Normal probability plot of estimation errors for AE, force, and vibration sensors

275
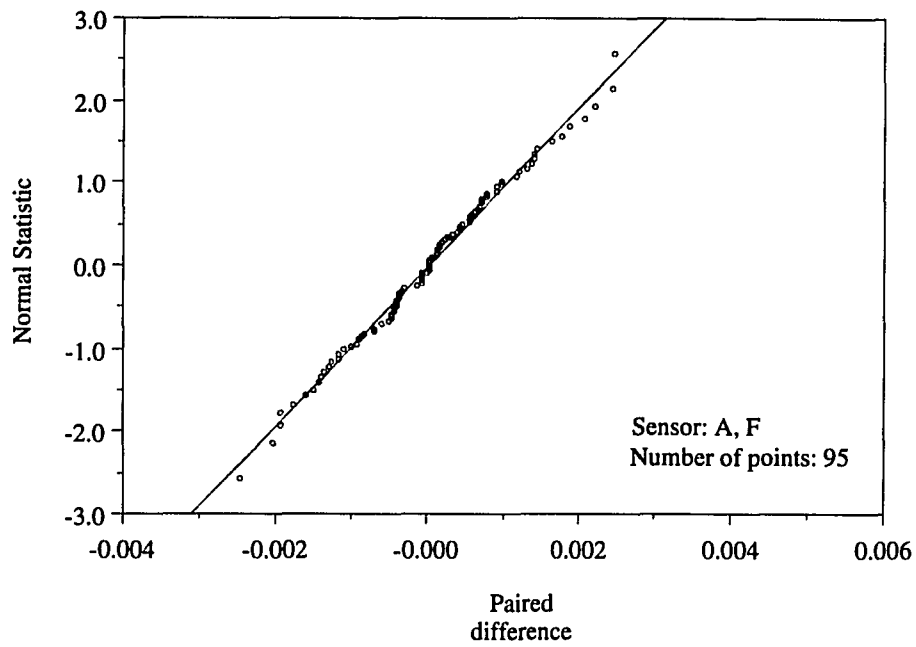
Figure 9.14: Normal probability plot of estimation errors for AE and force sensors

276
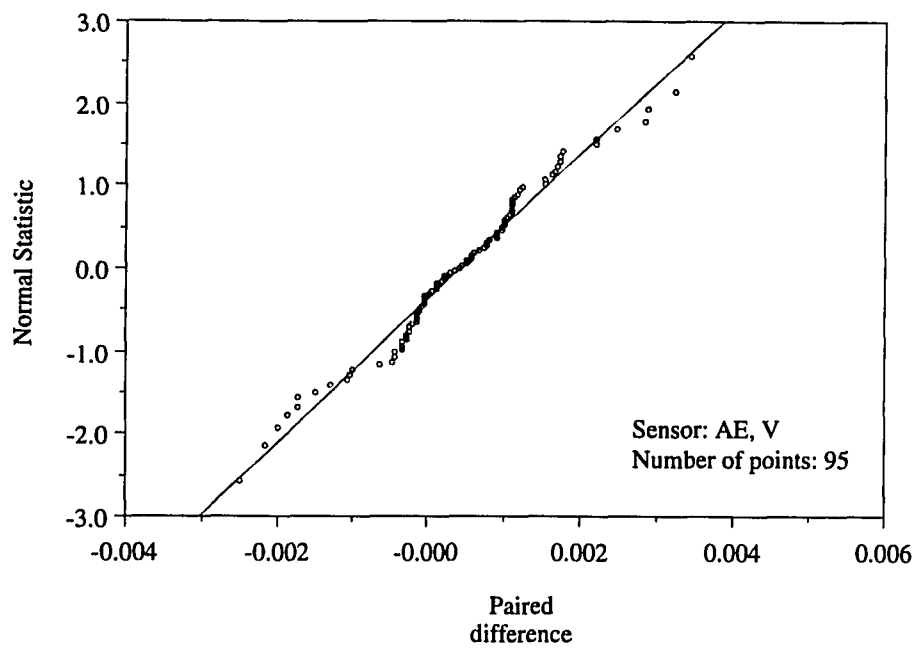
Figure 9.15: Normal probability plot of estimation errors for AE and vibration sensors

The figure's axes are labelled "Normal Statistic" (vertical) and "Paired difference" (horizontal), with an inset reading "Sensor: AE, V / Number of points: 95".

277
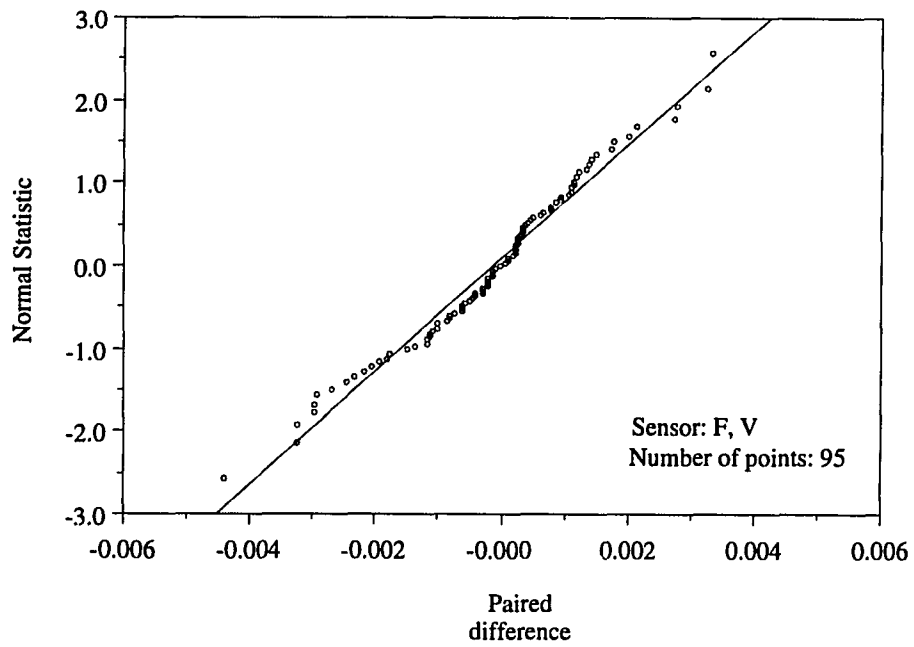
Figure 9.16: Normal probability plot of estimation errors for force and vibration sensors through univariate ARMA models

278
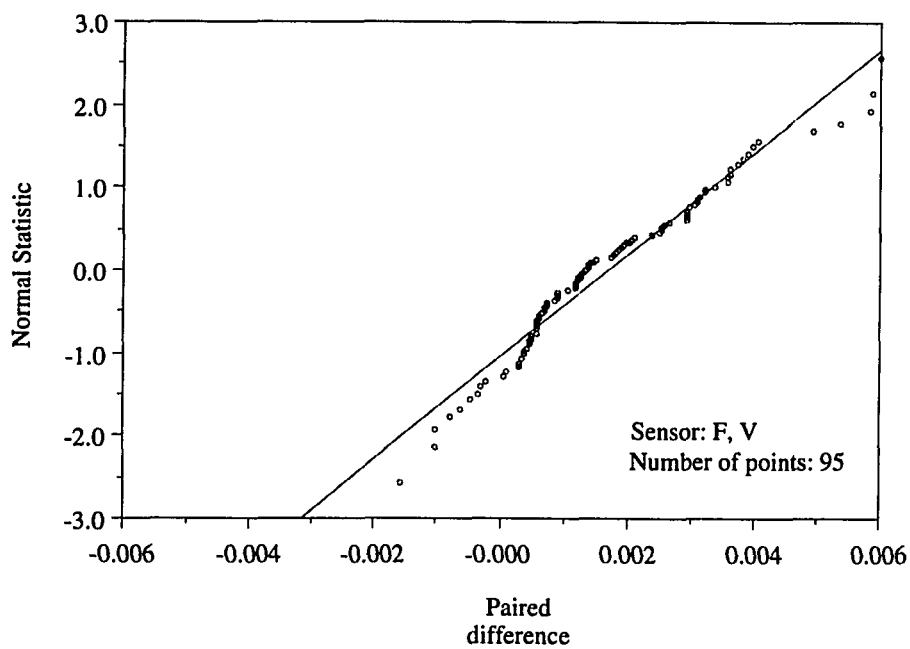
Figure 9.17: Normal probability plot of estimation errors for force and vibration sensors through multivariate ARMA models

279
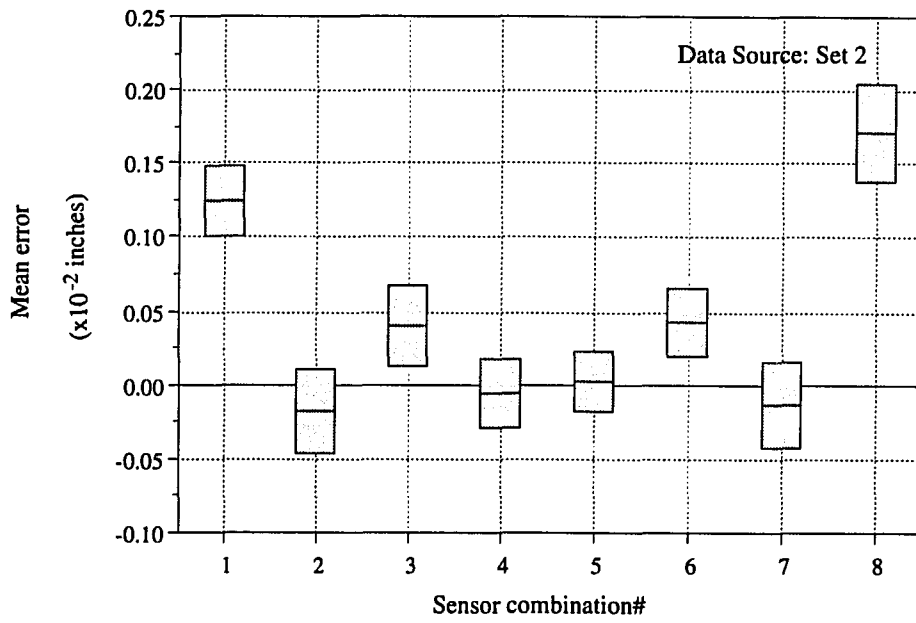
Figure 9.18: 95% confidence intervals for the mean of estimation errors for data in Set 2

280
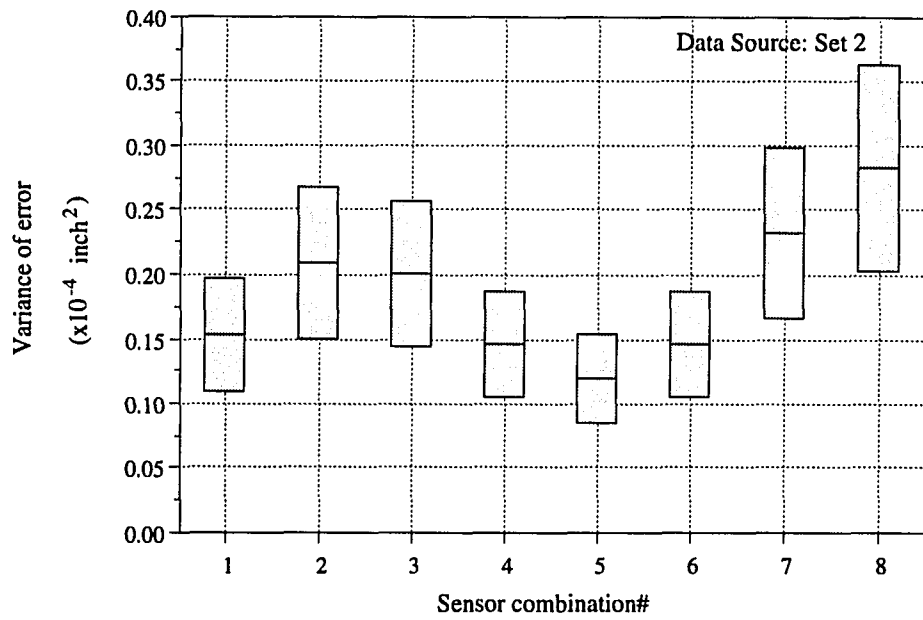
Figure 9.19: 95% confidence intervals for the variance of estimation errors for data in Set 2

281
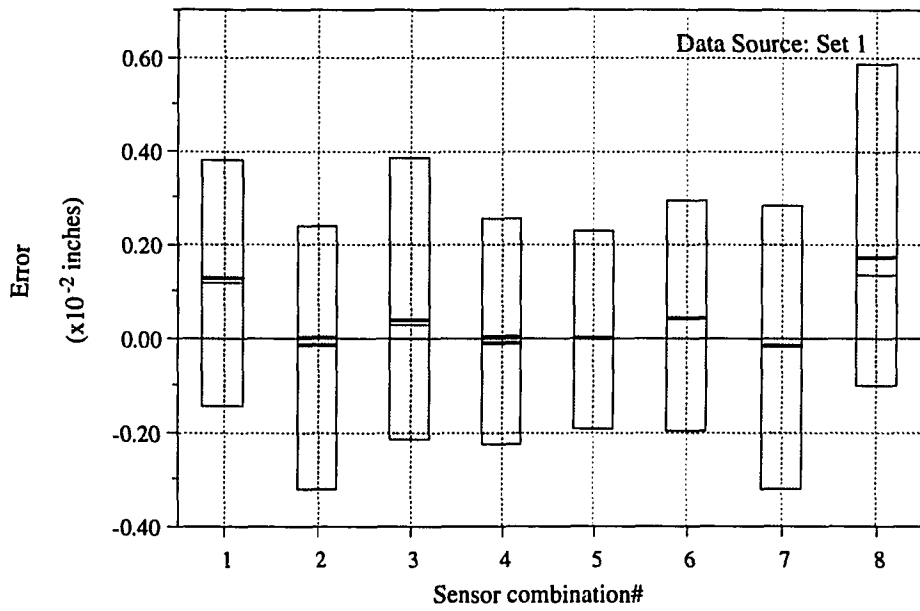
Figure 9..20: Empirical lower quantile (0.025) and upper quantiles (0.975) for flank wear estimation error for data in Set 2
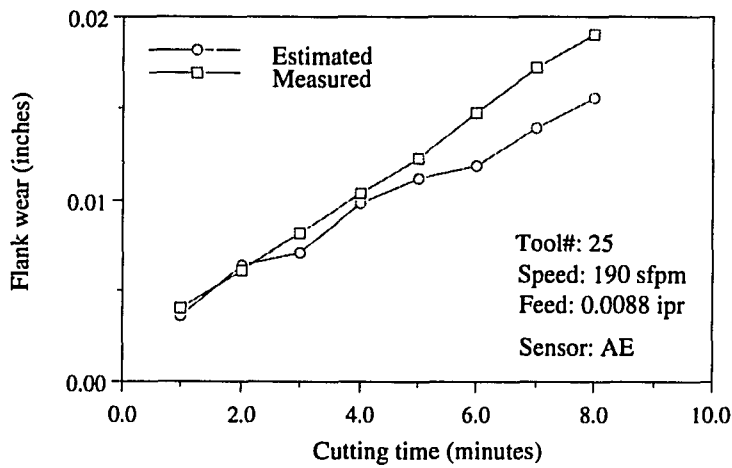
282
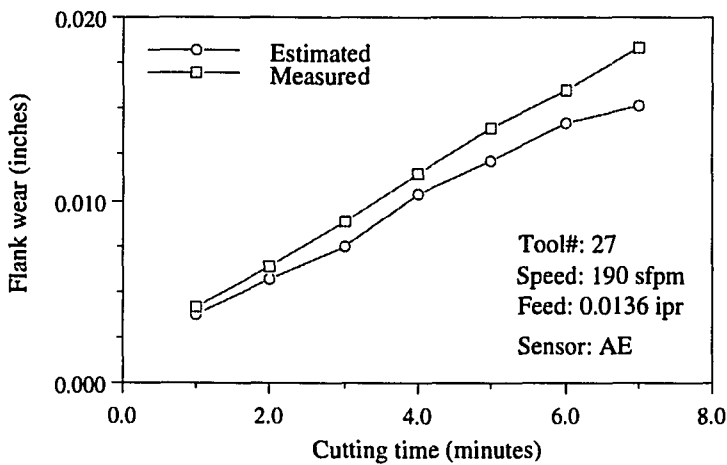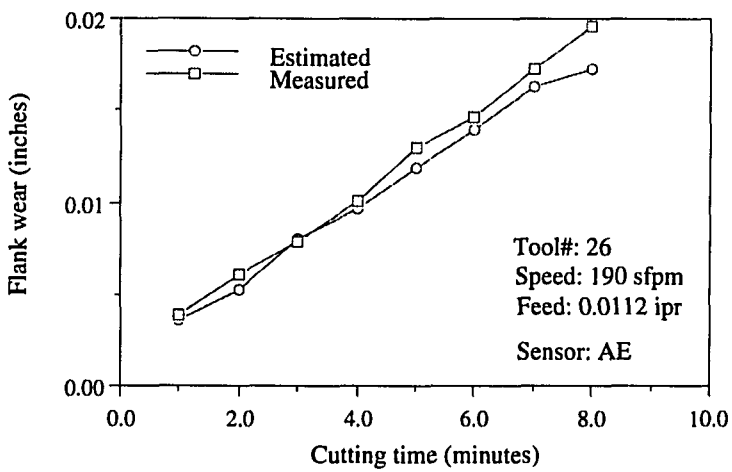
Figure 9.21: Flank wear estimation results using AE measurements

(Figure Continues)

283
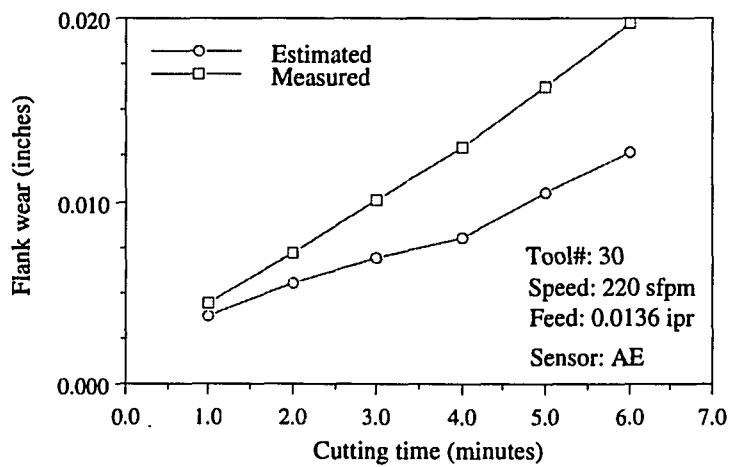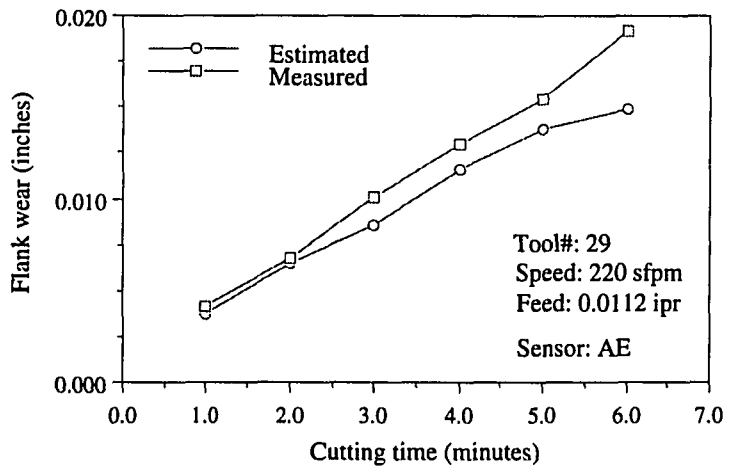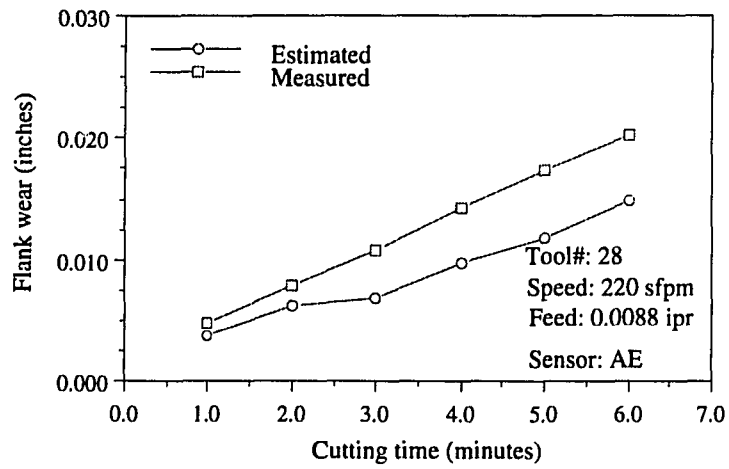
Figure 9.21: Continued

284

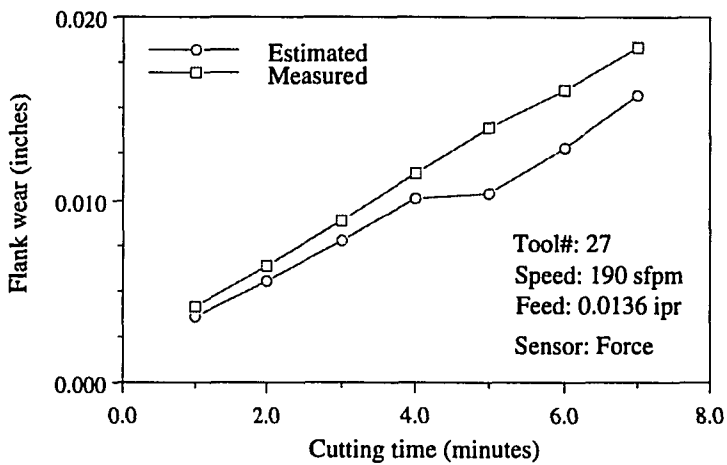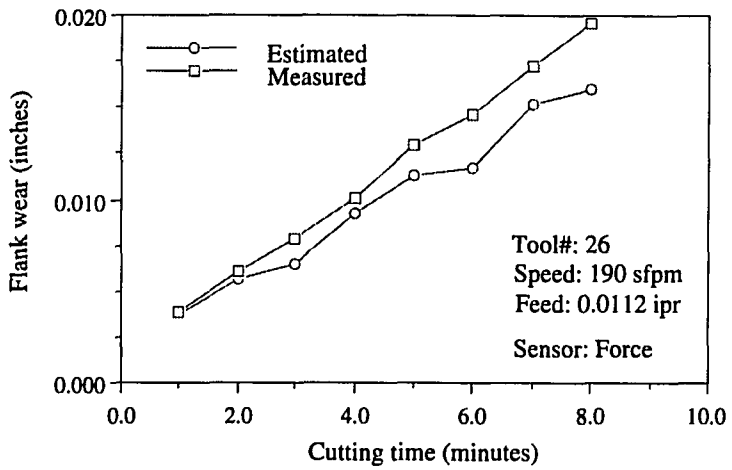**Figure 9.22:** Flank wear estimation results using force measurements

(Figure Continues)

285
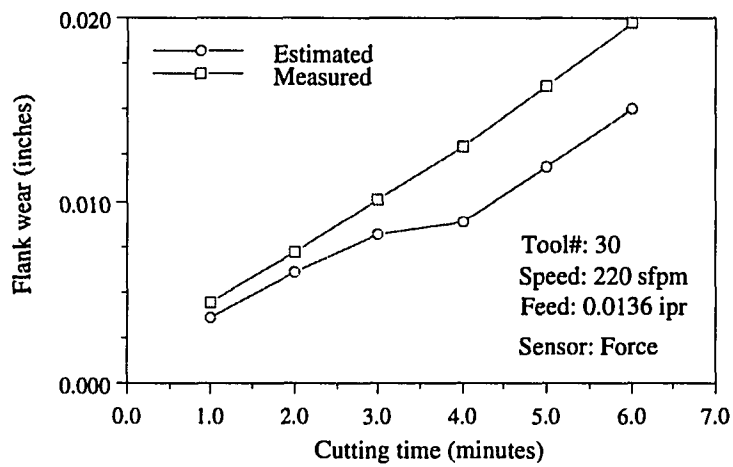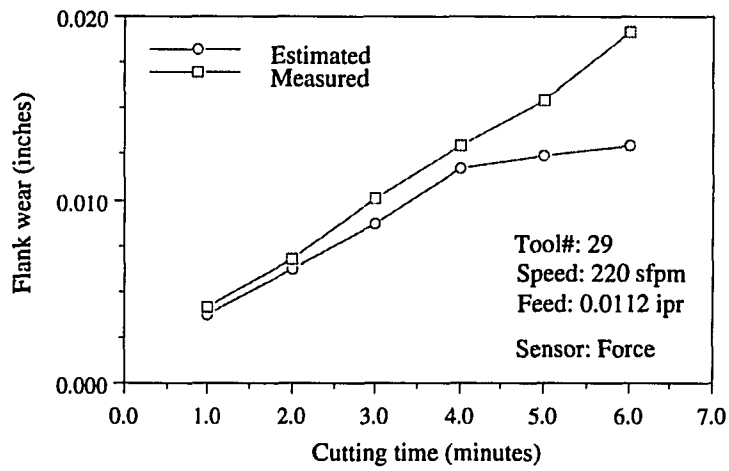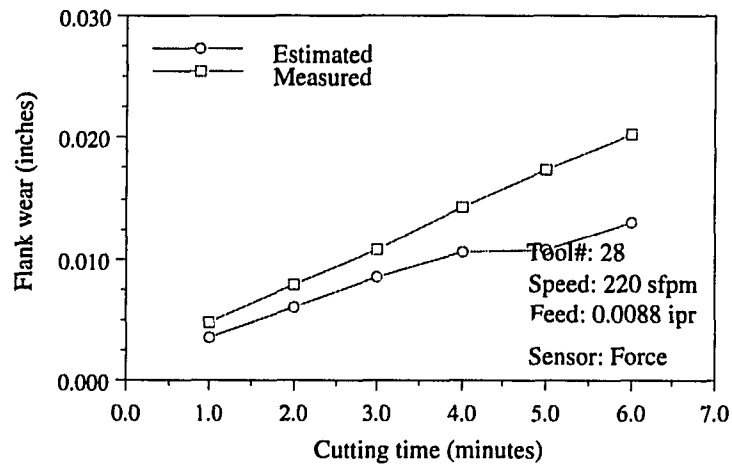
Figure 9.22: Continued

286
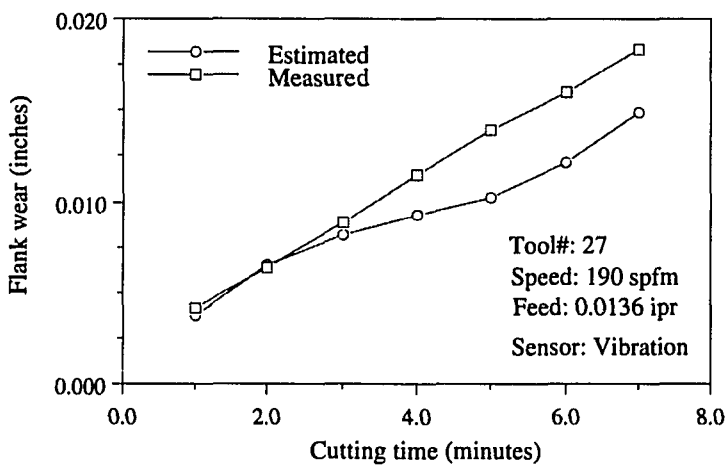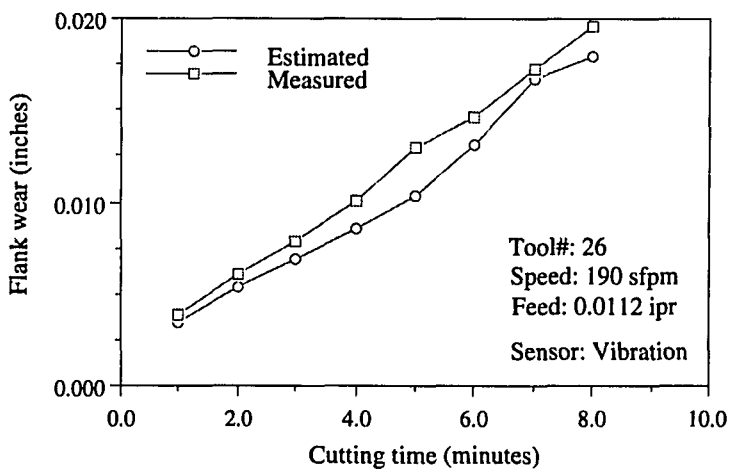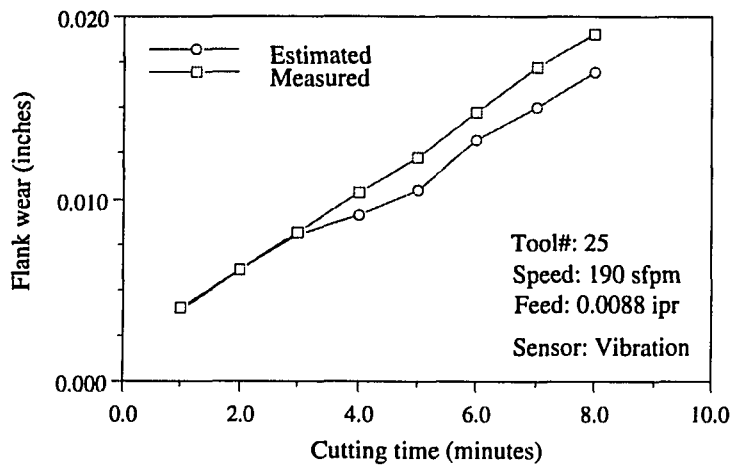
Figure 9.23: Flank wear estimation results using vibration measurements

(Figure Continues)

287
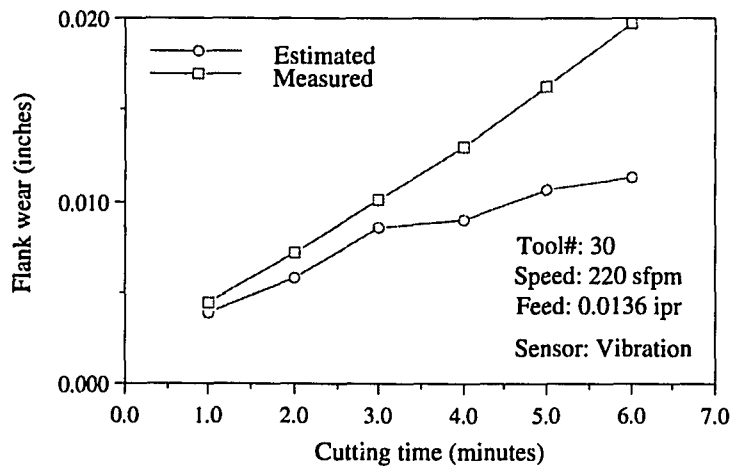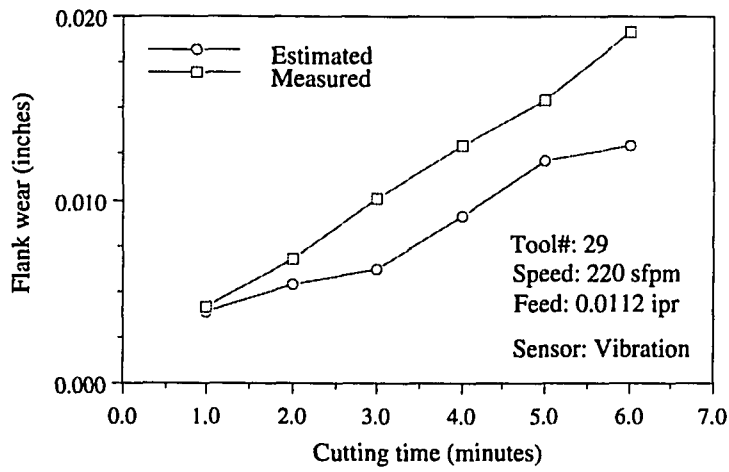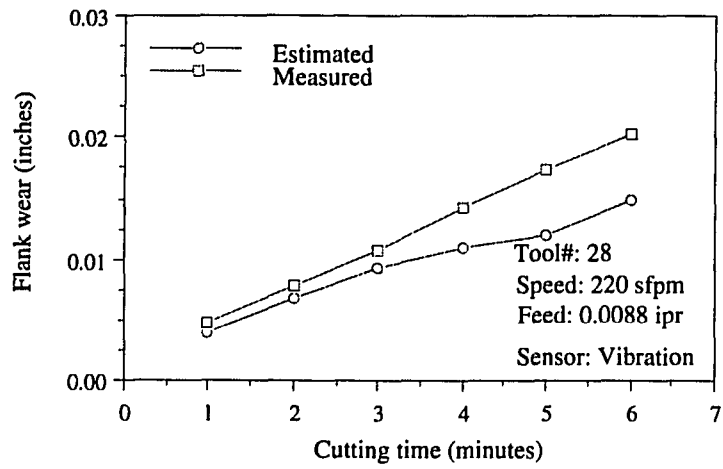
Figure 9.23: Continued

288
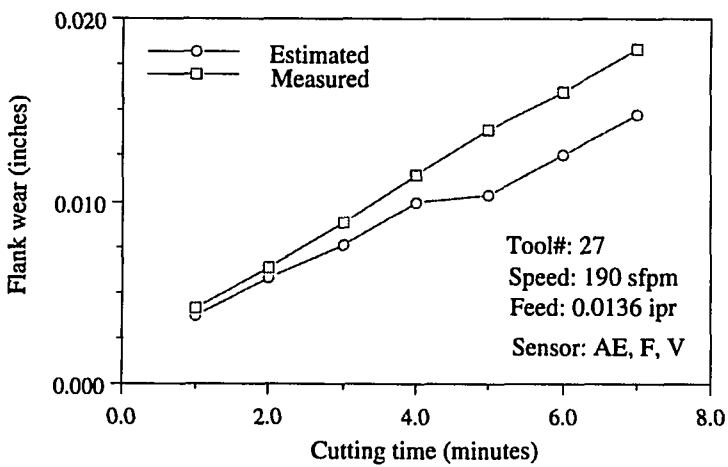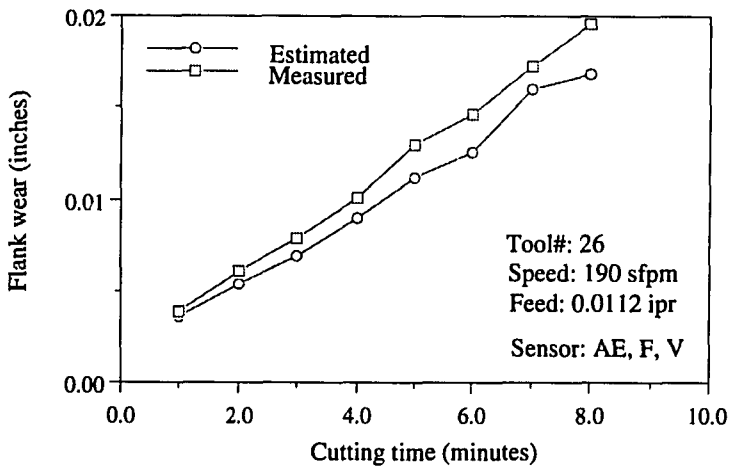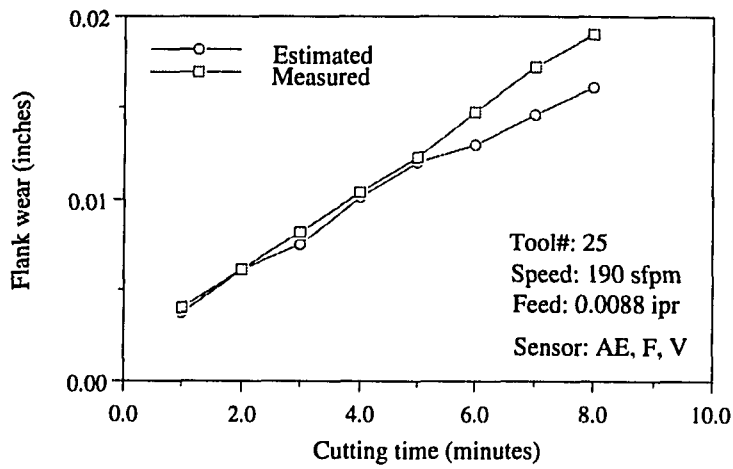
Figure 9.24: Flank wear estimation results using AE, force, and vibration measurements

(Figure Continues)
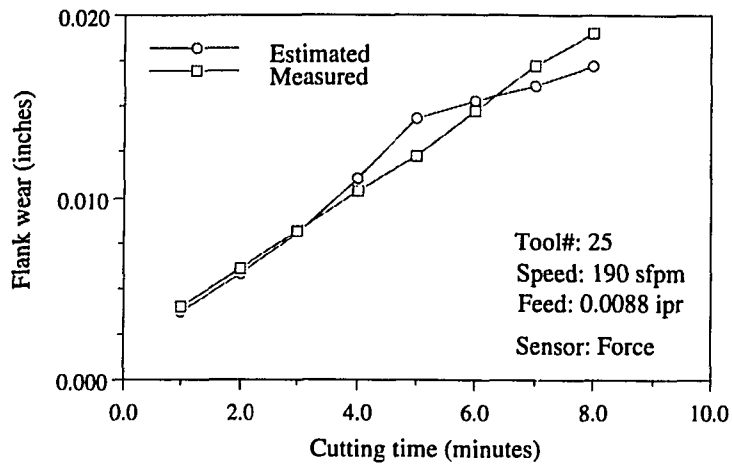
289
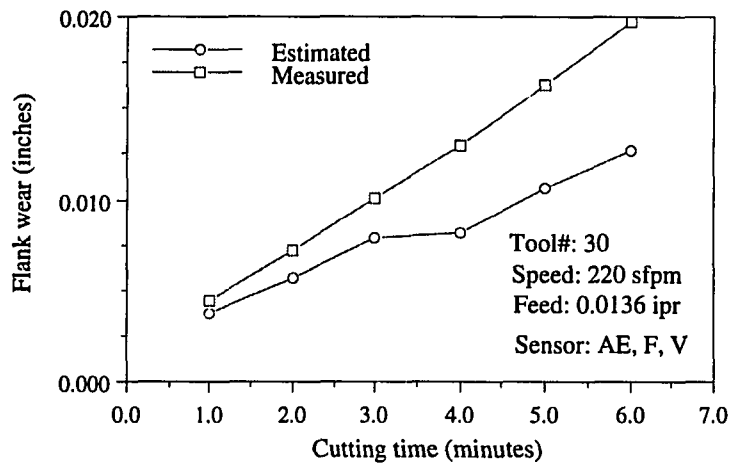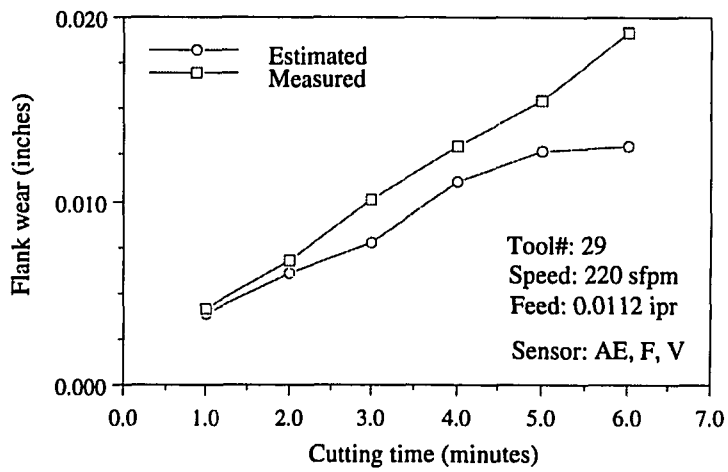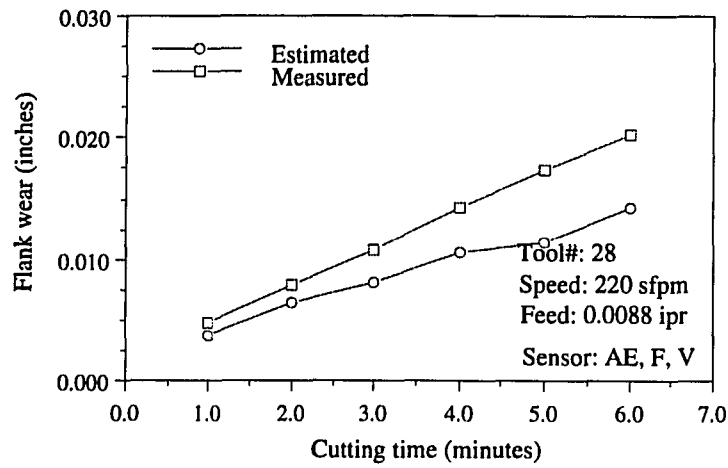
Figure 9.24: Continued

290
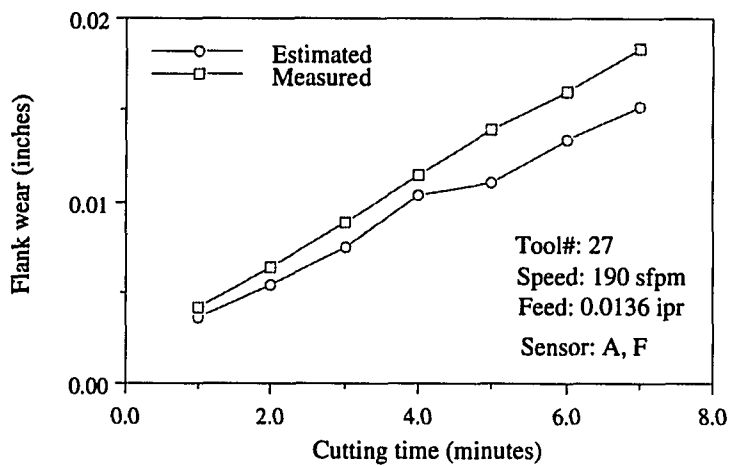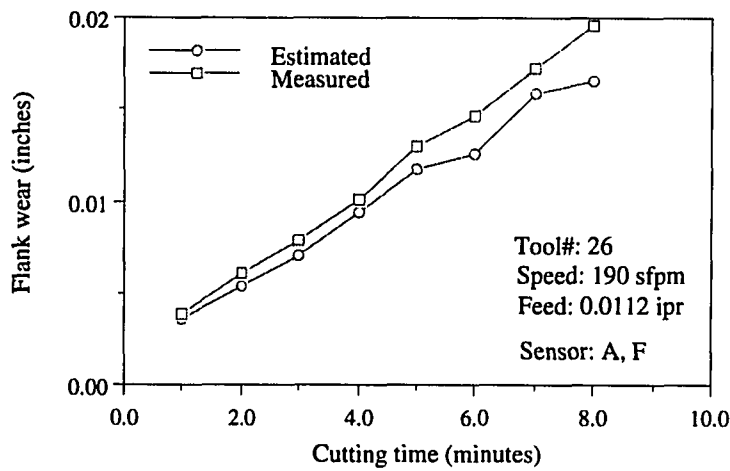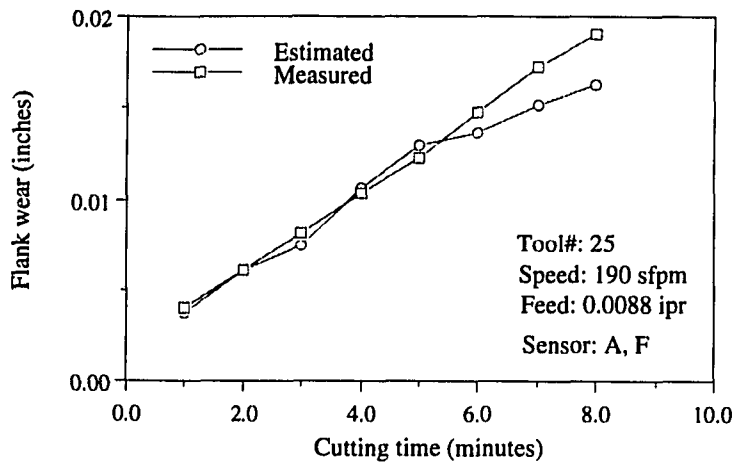
Figure 9.25: Flank wear estimation results using AE and force measurements

(Figure Continues)

291
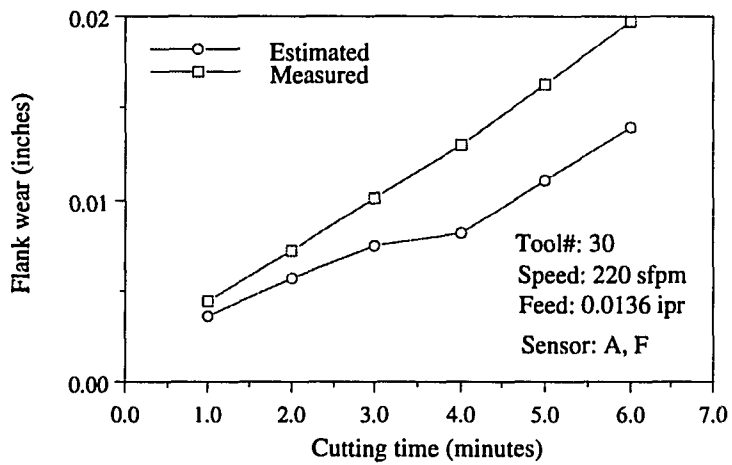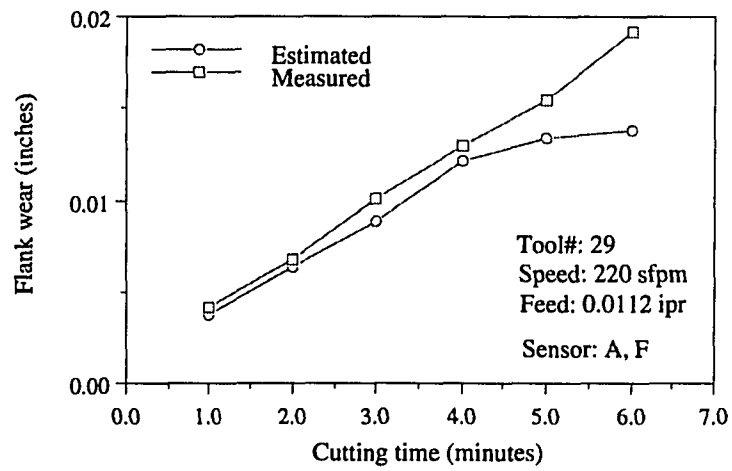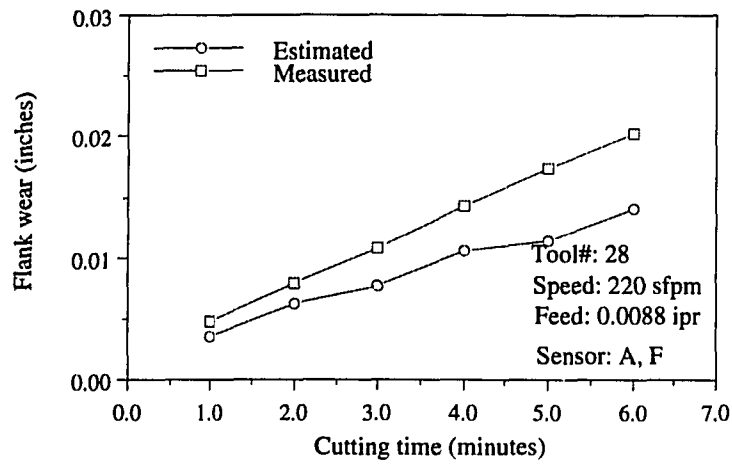
Figure 9.25: Continued

292
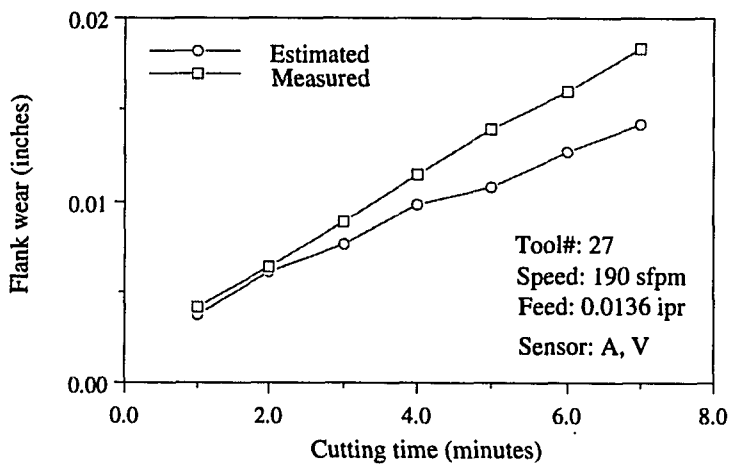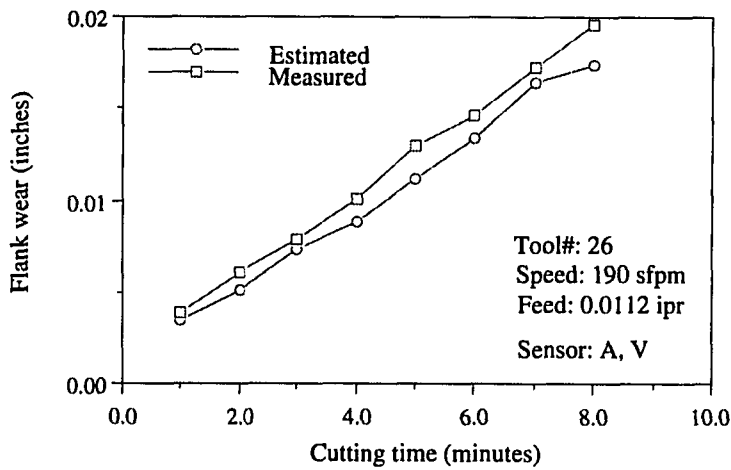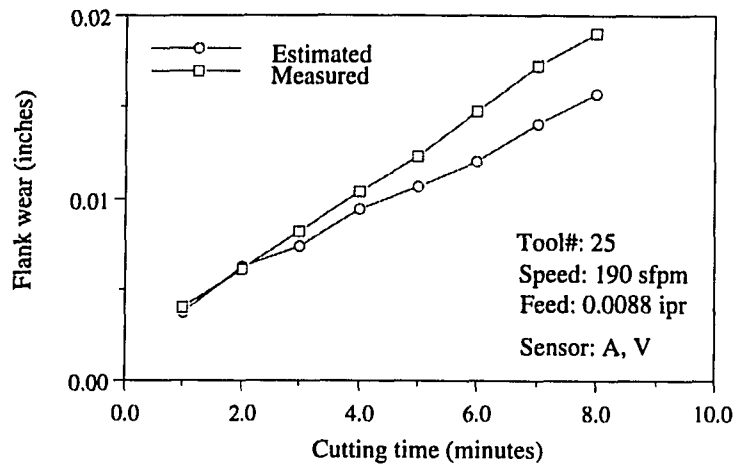
Figure 9.26: Flank wear estimation results using AE and vibration measurements

(Figure Continues)

293
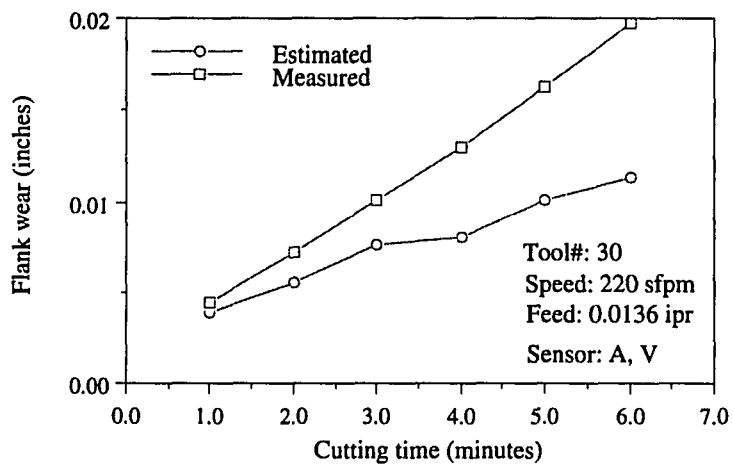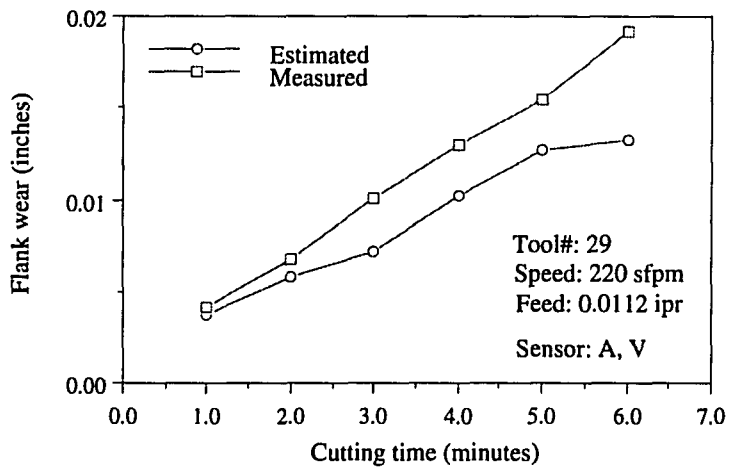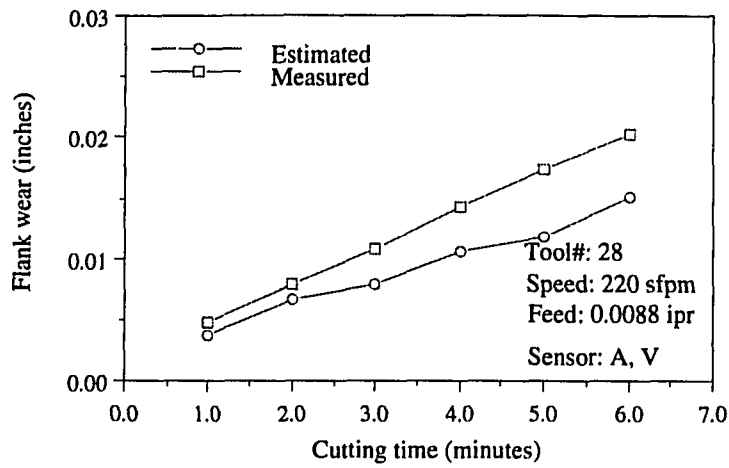
Figure 9.26: Continued

294
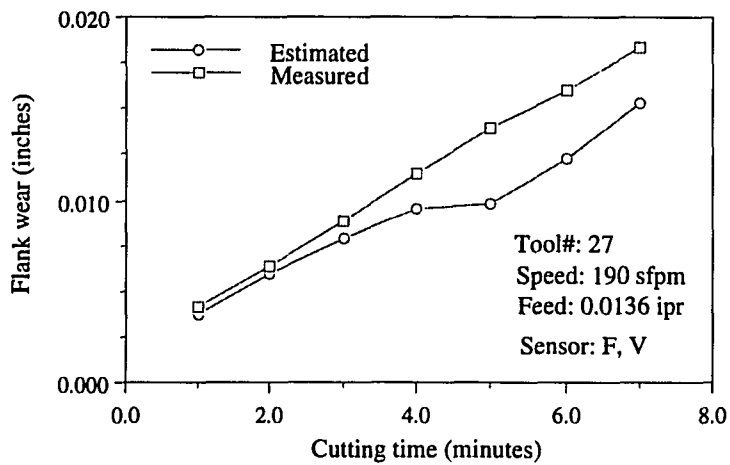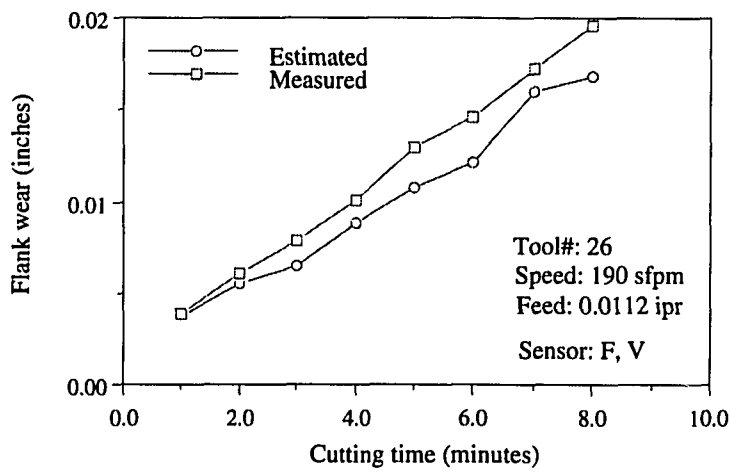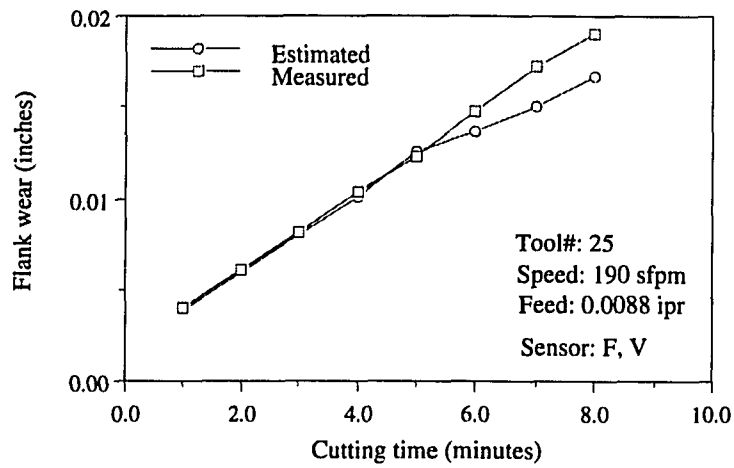
Figure 9.27: Flank wear estimation results using force and vibration measurements

(Figure Continues)

295
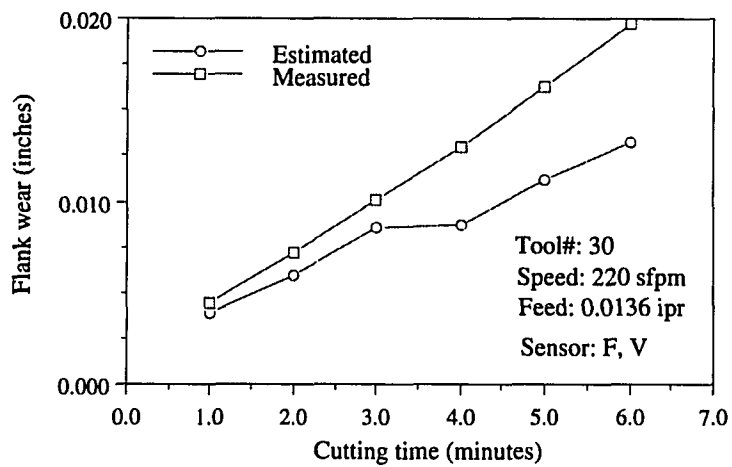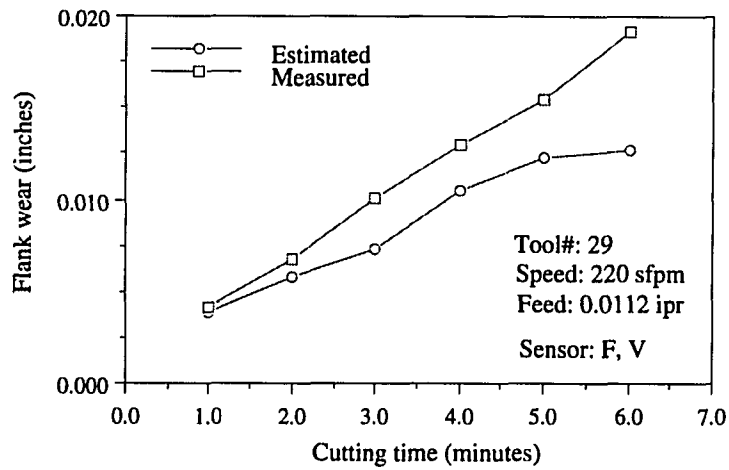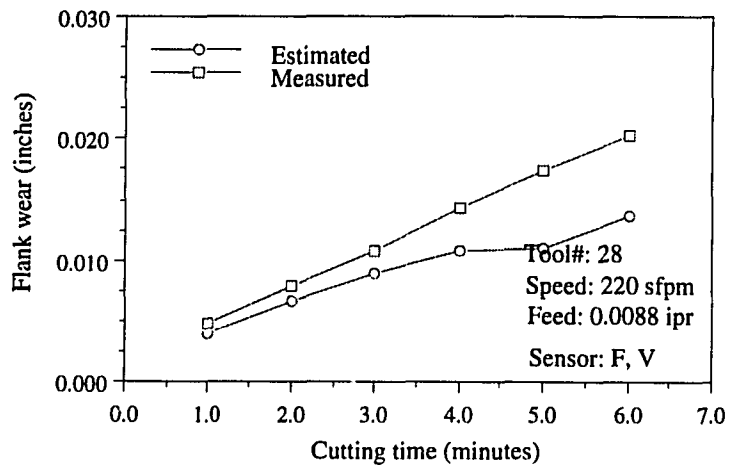
Figure 9.27: Continued

296
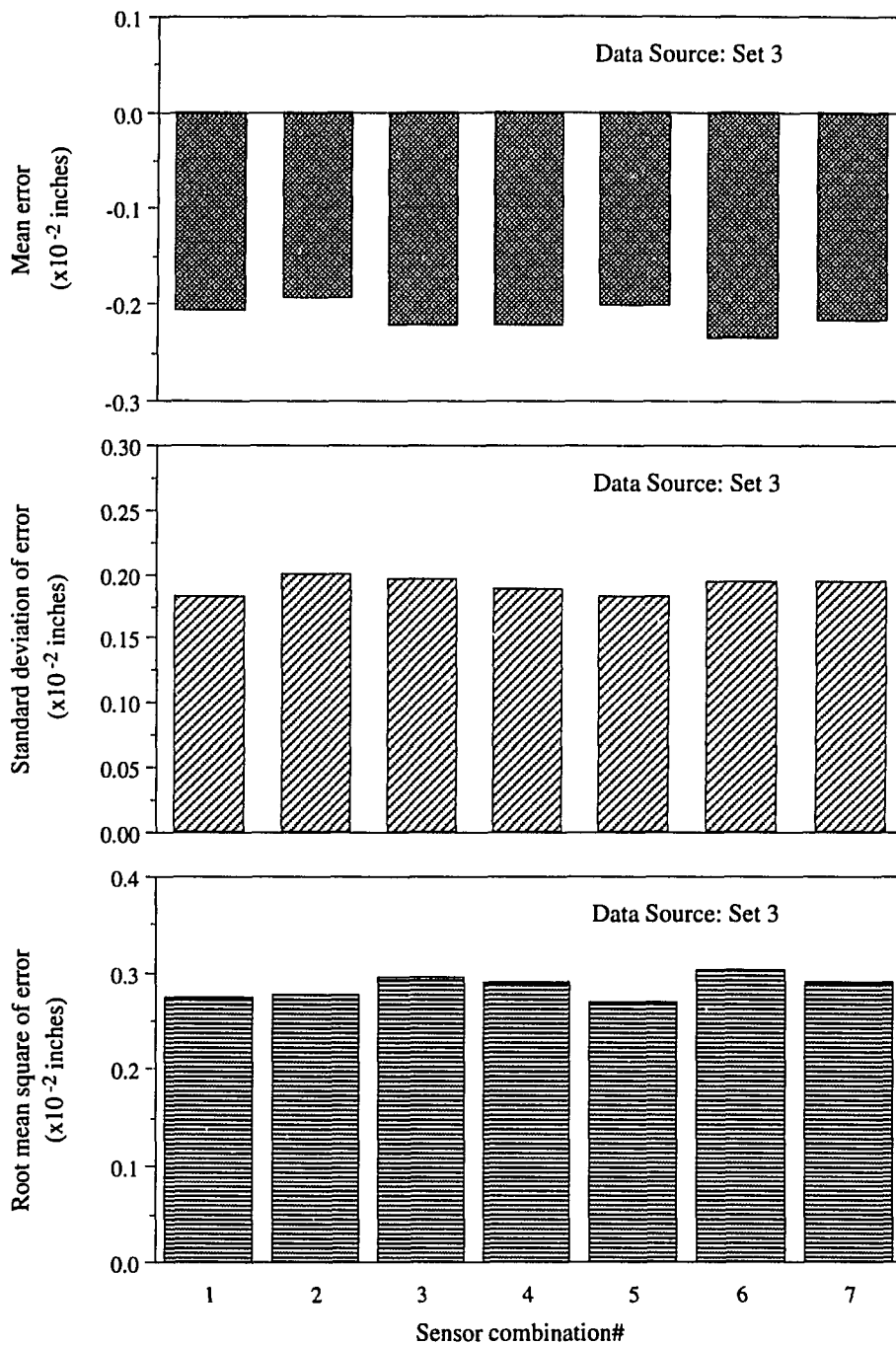
Figure 9.28: Mean and standard deviation, and RMS of estimation errors for data in Set 3

297
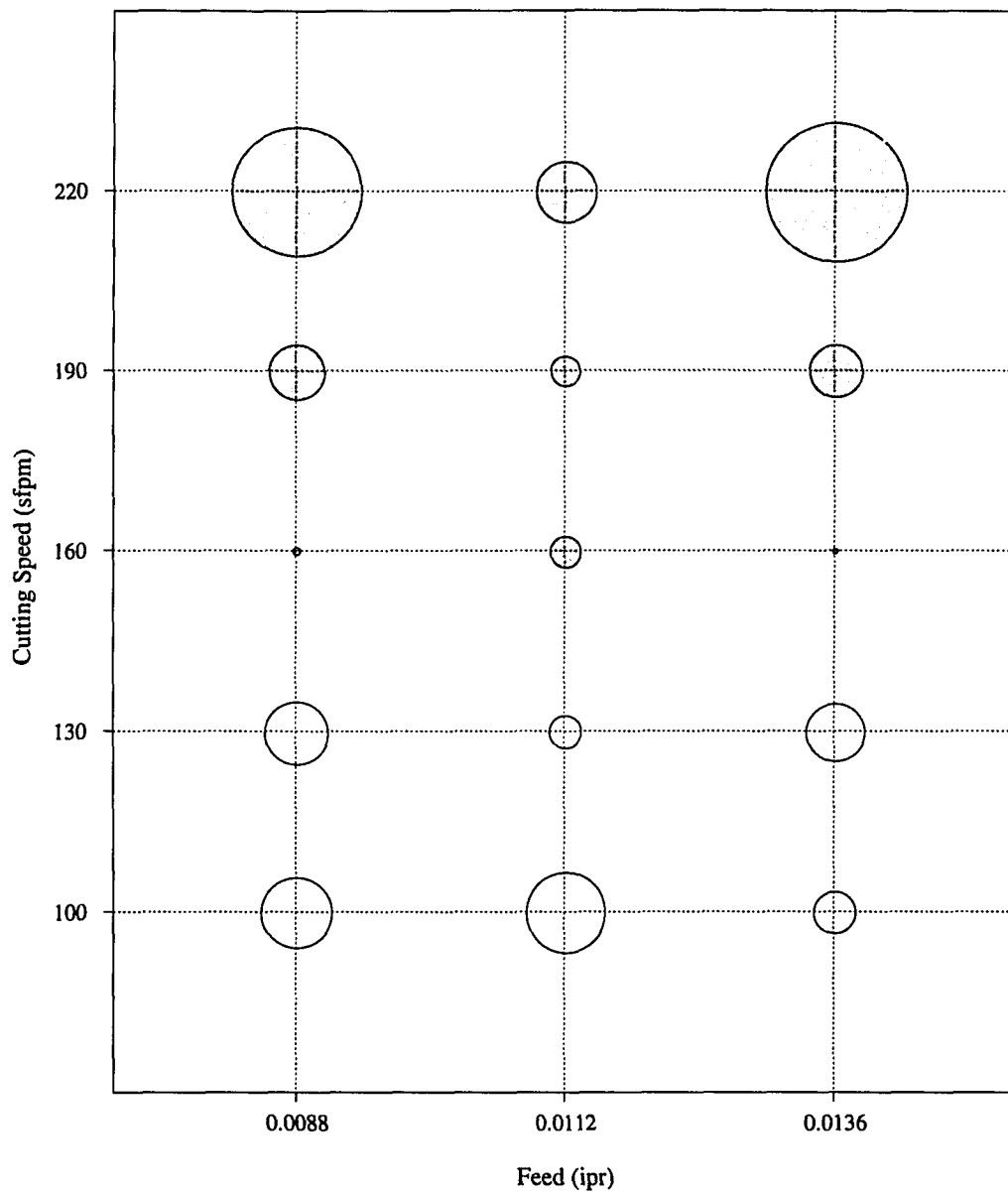
Figure 9.29: Variation of mean of estimation errors with respect to cutting speed and feed in the case of AE sensor. The diameter of a bubble is proportional to the mean error. A shaded bubble represents a negative value.

298
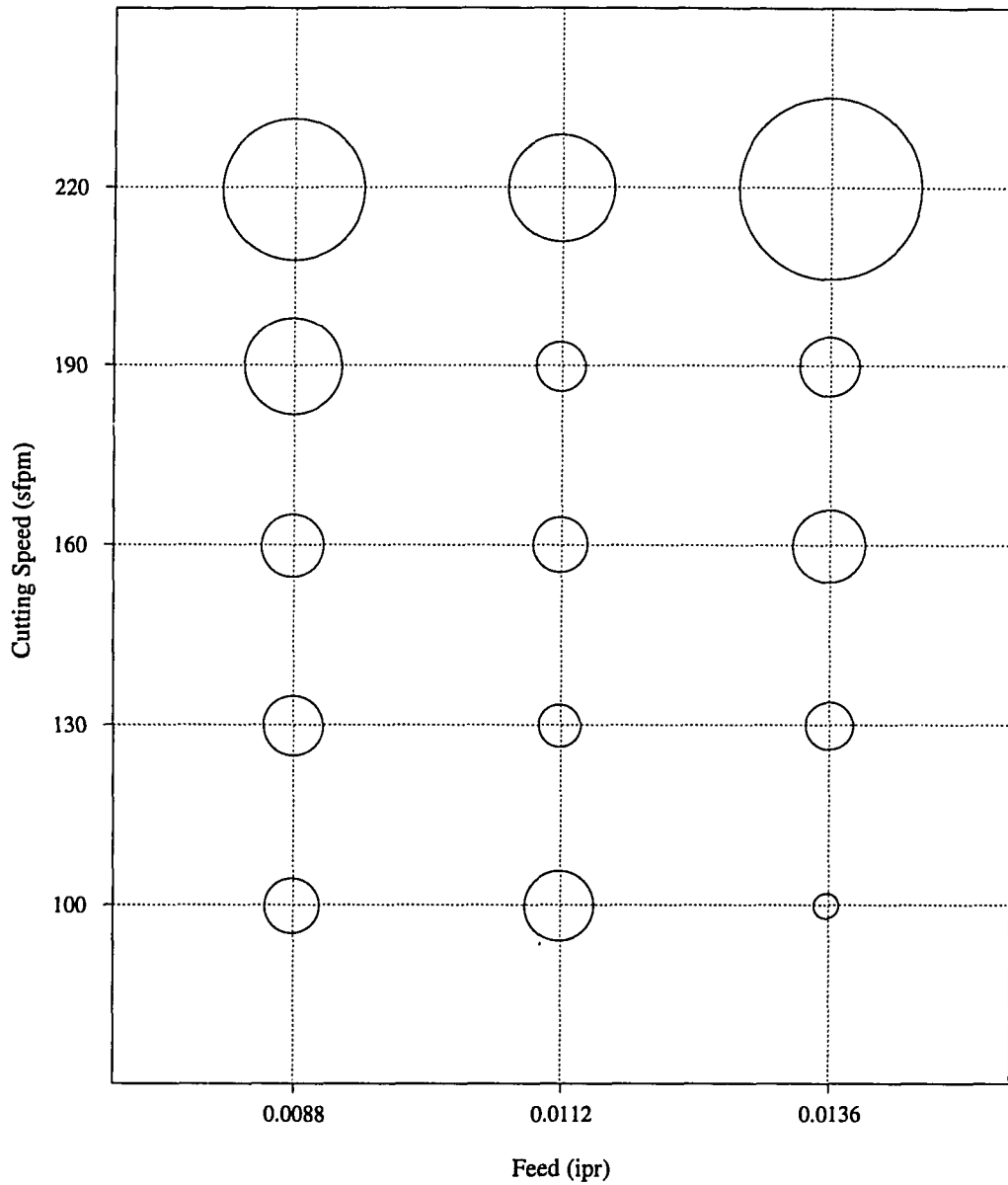
Figure 9.30: Variation of standard deviation of estimation errors with respect to cutting speed and feed in the case of AE sensor. The diameter of a bubble is proportional to the standard deviation of error. A shaded bubble represents a negative value.

299
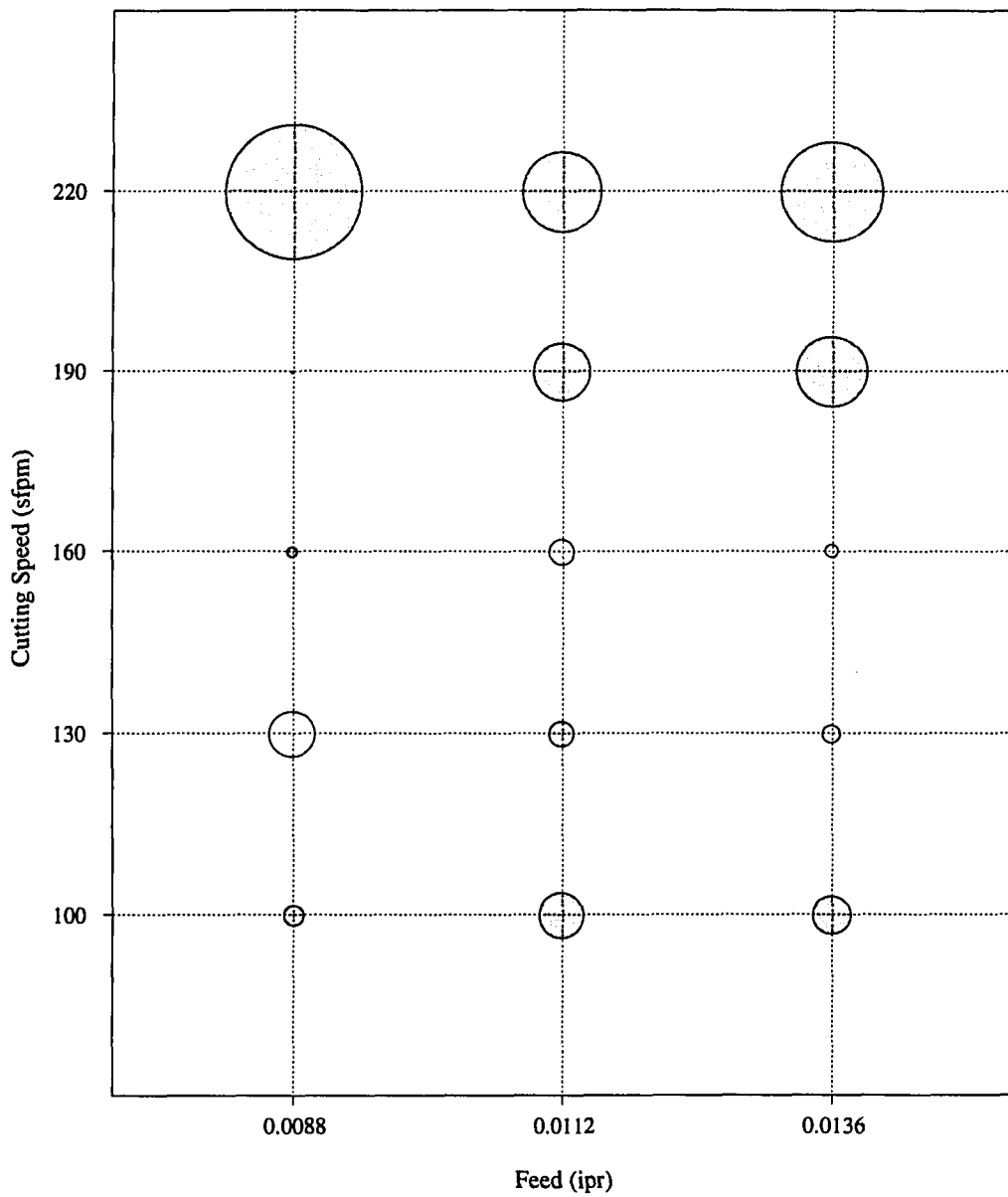
Figure 9.31: Variation of mean of estimation errors with respect to cutting speed and feed in the case of force sensors. The diameter of a bubble is proportional to the mean error. A shaded bubble represents a negative value.

300
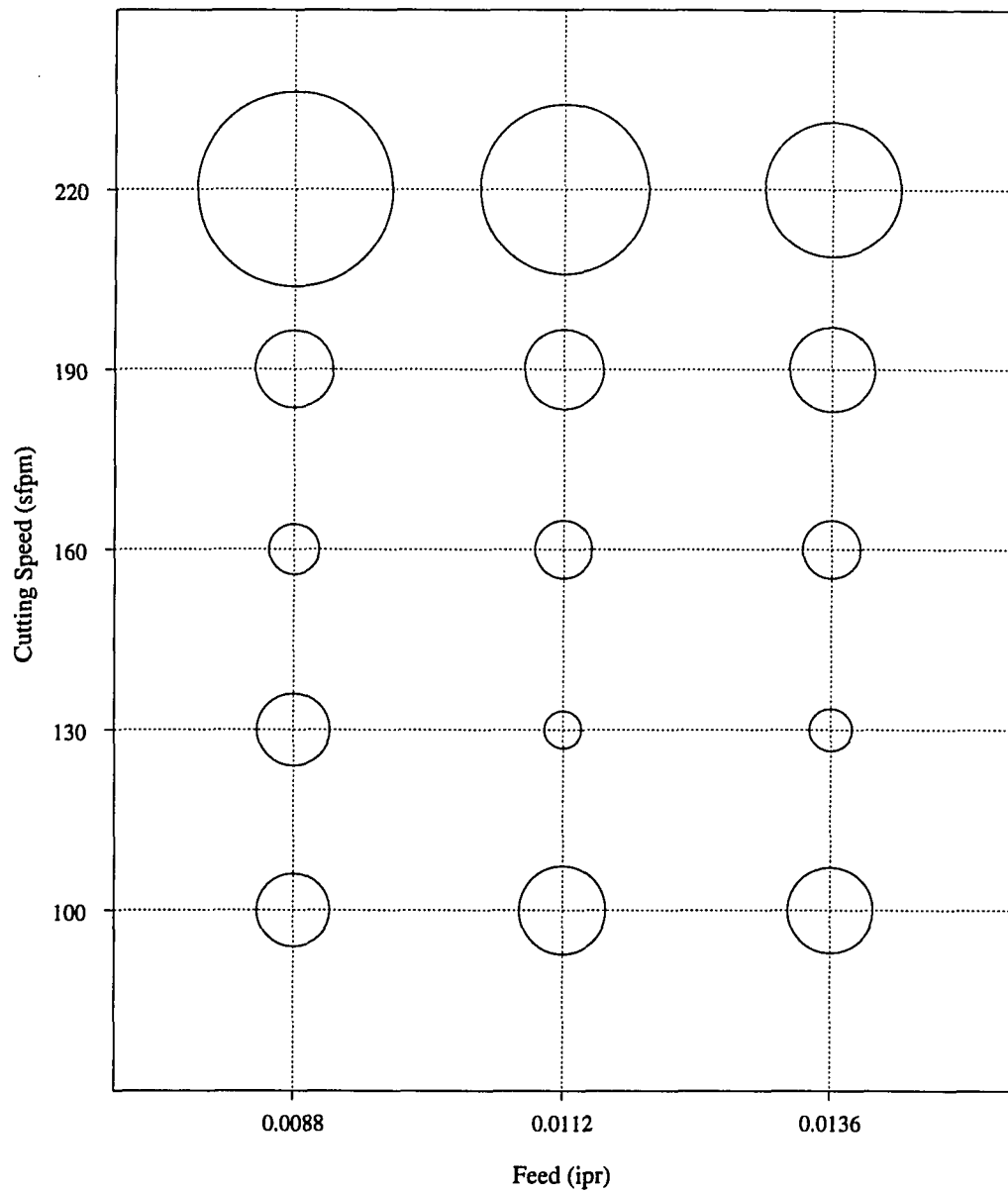
Figure 9.32: Variation of standard deviation of estimation errors with respect to cutting speed and feed in the case of force sensors. The diameter of a bubble is proportional to the standard deviation of error. A shaded bubble represents a negative value.

301
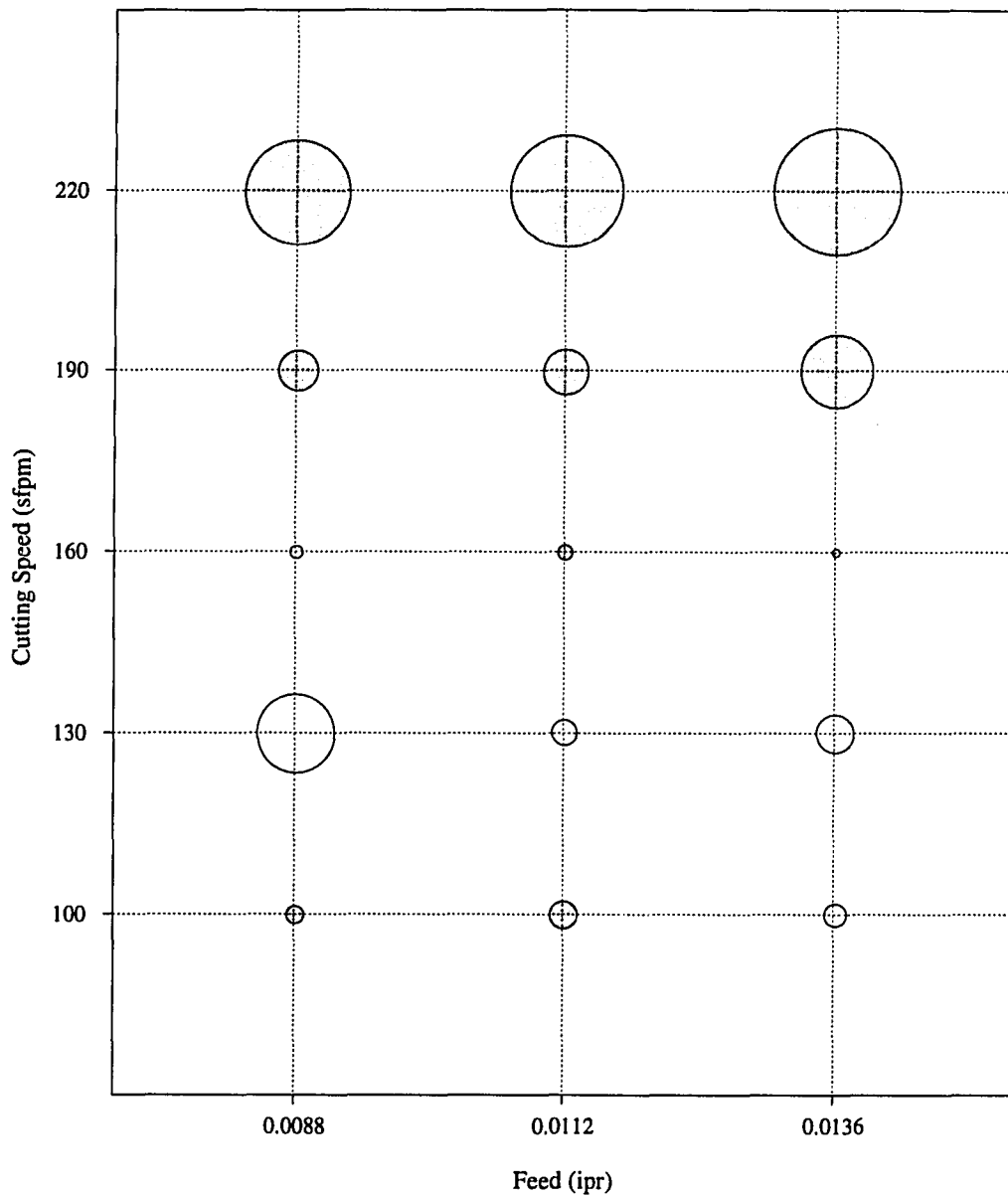
Figure 9.33: Variation of mean of estimation errors with respect to cutting speed and feed in the case of vibration sensors. The diameter of a bubble is proportional to the mean error. A shaded bubble represents a negative value.

302
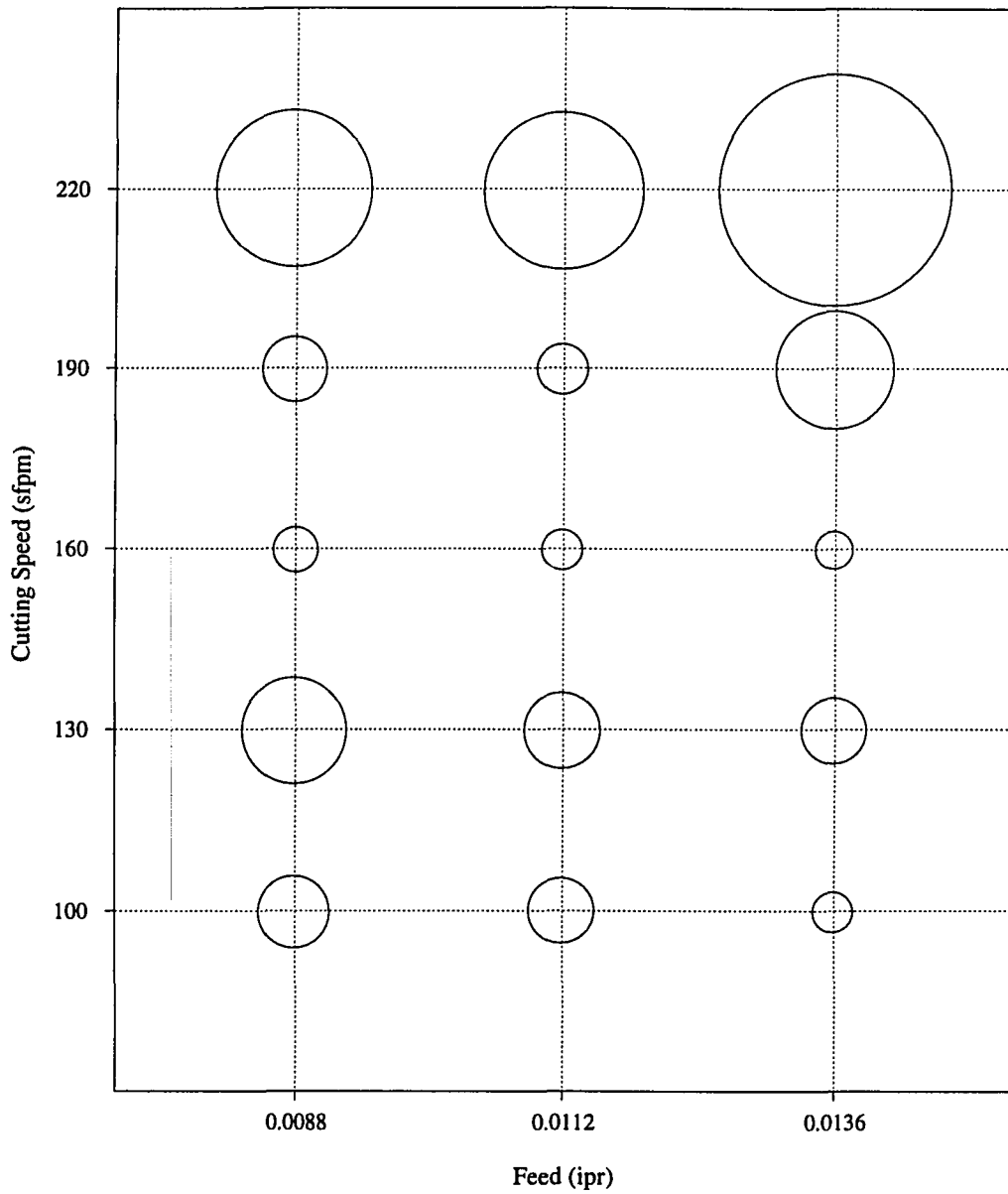
Figure 9.34: Variation of standard deviation of estimation errors with respect to cutting speed and feed in the case of vibration sensors. The diameter of a bubble is proportional to the standard deviation of error. A shaded bubble represents a negative value.

303

Chapter 10

CONCLUSIONS, UNIQUENESS, AND FUTURE WORK

10.1 Introduction

This chapter summarizes the research work in this thesis, and it also highlights the uniqueness of this work and indicates the direction for future work.

10.2 Summary of the Work

In this research, an new on-line method for flank wear estimation in turning processes is developed. The application of sensor data fusion and neural networks are the important features of this method. Multisensor data fusion provides the advantages of redundancy, complementariness, and timeliness of information. On the other hand, neural networks provide the correct flank wear estimates by relating sensor measurements with flank wear. On-line flank wear estimation in turning processes is primarily required for devising optimal tool change strategies and implementing adaptive control schemes.

Previous tool wear monitoring methods based on analytical models have not been effective because of insufficient understanding of the physics of the metal removal process. The present approach circumvents the difficulty associated with the identification of the correct structure and the parameters of analytical models by using neural networks which relate sensor measurements to flank wear. However, the proposed method is open to further improvement when correct analytical

304

models are available.

In this work, force, vibration, and AE signals from the turning process are measured by a three-axis dynamometer, two accelerometers, and an acoustic emission sensor respectively. The signals from the sensors are digitized using a battery of data acquisition systems. The digitized data is preprocessed using a set of bandpass filters to improve the signal to noise ratio. The sensor data from the force and vibration sensors are represented by univariate or multivariate ARMA models. The coefficients of the ARMA models are used as the features of force and vibration signals. The sensor data from the AE sensor is represented by discrete wavelet transforms. The coefficients of the discrete wavelet transforms are used as the features of the AE signals. The patterns created from coefficients of either ARMA models or discrete wavelet transforms are input to a specially designed neural network to compute flank wear estimates.

The neural network architecture for flank wear estimation is designed by combining Kohonen's feature maps, radial basis functions, and recurrent neural networks. The flank wear estimation neural network has three stages. In the first stage, a Kohonen's feature map finds the centers of the input space defined by the sensor measurements. In the second stage, a radial basis function utilizes these centers and approximates the function that maps the sensor measurements to flank wear. The inputs from each sensor are fed to a separate pair of Kohonen's feature map and radial basis function to obtain the flank wear estimates based on the sensor measurements. In the third stage, a recurrent neural network integrates the flank wear estimates from different sensors to provide more accurate and robust flank wear estimates. The flank wear estimation neural network that is developed in this research eliminates the difficulty in identifying the network topology, guar-

antees the network convergence, and reduces the demand on supervised training time and supervised data which is expensive to collect. This neural network is fault tolerant to metal-cutting fluctuations caused by non-homogeneities and hardspots in the workpiece material.

The performance of the flank wear estimation method is studied by conducting a set of experiments. The results indicate that the proposed method gives flank wear estimates accurate enough for most manufacturing applications.

10.3 Conclusions

- The force, vibration, and AE signals generated in a turning process contain information about gradually progressing flank wear. This information can be successfully extracted and used for on-line estimation of flank wear.

- The proposed flank wear estimation method provides accurate flank wear estimates within the range of operating conditions that were used during neural network training data generation.

- The onset of crater wear on a cutting tool adversely affects the accuracy of flank wear estimates.

- Among all the possible sensor combinations, AE-force sensors provide the most accurate and consistent flank wear estimates. In contrast, force vibration sensors provide the least accurate flank wear estimates.

- The discrete wavelet transforms are very effective in representing AE signals for the purpose of flank wear estimation.

306

- The univariate ARMA models are more appropriate than multivariate ARMA models for representing force and vibration signals for purpose of flank wear estimation.

- The RMS of estimation errors for all sensor combinations fall in range of 0.0010 to 0.0017 inches. This is a considerably low estimation error, making the proposed estimation method very attractive to the real-world applications.

10.4 Uniqueness and Contributions of the Work

- The previous neural network-based tool wear monitoring methods classify tools into two or three classes: fresh, medium, or worn-out. In contrast, the present method predicts continuous flank wear values. Continuous flank wear estimation is a problem of much higher dimension than the problem of classifying tools based on their tool wear level.

- A neural network with useful properties for flank wear estimation is created by combining Kohonen's feature maps, radial basis functions, and recurrent neural networks. This neural network architecture requires less training time compared to a feedforward neural network. This neural network architecture also offers a unique fault tolerance property.

- The effectiveness of univariate and multivariate ARMA models to represent force and vibration signals is investigated. In the existing method, the force and vibration signals are modeled by AR models but not by ARMA models. The present work has, for the first time, investigated the representation of these signals by ARMA models for flank wear estimation.

- The effectiveness of discrete wavelet transforms to represent AE signals is investigated. This is the first in the metal cutting area that the representation

307

of AE signals by discrete wavelet transforms has been investigated. In no other AE-based tool wear estimation method reported in literature have such accurate flank wear estimation results been achieved.

- Extensive experiments are conducted to verify and validate the performance of the flank estimation method proposed in this research. There are no comparable experimental results published in literature. Although some experimental results are available in literature, those results were obtained for different tool and workpiece materials and different experimental set up then those of the present research. These differences make it hard to compare the present flank wear estimation results with the existing once.

- The proposed flank wear estimation method is generic enough to extend its application to other sensor-based monitoring problems in manufacturing.

10.5 Future Work

Extension to the present work is possible in two directions: finding new applications to the present techniques and improving the techniques themselves.

The present work can be potentially applied to the following sensor-based monitoring problems in manufacturing:

1. Machine tool monitoring and diagnosis.

2. Breakdown prediction of mechanical components such as bearing and gears.

3. Real-time quality control of manufacturing processes.

4. Non-destructive inspection using AE signals.

The present technique can be improved with the following alternative signal representation schemes and computational methods:

308

1. Use wavelet pockets to represent AE signals.

2. Use discrete wavelet transforms or wavelet pockets to represent force and vibration signals.

3. Compute ARMA coefficients using recursive estimation methods based on the Kalman-Filter principle.

4. Adaptively change the RBF network weights with respect to cutting time to improve the estimation accuracy.

5. Develop a hybrid flank wear estimation method which uses both analytical models and neural networks.

309

REFERENCES

Ahalt, S. C., Krishnamurthy, A. K., Chen, P., and Melton, D. E., (1990). "Competitive Learning Algorithms for Vector Quantization," *Neural Networks*, Vol. 3, No. 3, pp. 277–291.

Arulalan, G. R., Ralson, P. A. S., and Ward, T. L., (1988). "Adaptive Observer for Optimal Control of Lathe Tool Wear," *Proceedings of the 1988 American Control Conference*, Vol. 3, pp. 1930–1937.

Balakrishnan, P., Trabelsi, H., Kannaty-Asibu, Jr. E., and Emel, E., (1989). "A Sensor Fusion Approach to Cutting Tool Monitoring," in *Advances in Manufacturing Systems Integration and Processes*, Dornfeld, D. A. (ed.), pp. 101–108.

Bath, M., and Sharp, R., (1968). "In-Process Control of Lathes Improves Accuracy and Productivity," *Proceedings of the 9th International Machine Tool Design and Research Conference*, pp. 1209–1221.

Billett, R. A., (1968). "Studies of Cutting Temperature Control Applied to a Lathe Spindle Speed," *Proceedings of the 9th International Machine Tool Design and Research Conference*, pp. 1273–1287.

Blackman, R. B., and Tukey, J. W., (1959). *The Measurement of Power Spectra From the Point of View of Communications Engineering*, New York: Dover Publications.

Boothroyd, G., Eagle, J. M., and Chisholm, W. J., (1967). "Effect of Tool Flank Wear on The Temperatures Generated During Metal Cutting," *Proceedings of the 8th International Machine Tool Design and Research Conference*, pp. 667–680.

Boothroyd, G., and Knight, W. A., (1989). *Fundamentals of Machining and Machine Tools*, New York: Marcel Dekker, Inc.

Box, G. E. P., and Jenkins, G. M., (1970). *Time Series Analysis–Forecasting and Control*, San Francisco: Holden-Day.

Box, G. E. P., Hunter, W. G., and Hunter, J. S., (1978). *Statistics for Experimenters: an Introduction to Design, Data Analysis, and Model Building*, New York: Wiley.

Box, G. E. P., and Tiao, G. C., (1977). "A Canonical Analysis of Multiple Time Series," *Biometrika*, Vol. 64, pp. 355–365.

Burke, L. I., (1993). "An Unsupervised Neural Network Approach to Tool Wear Identification," *IIE Transactions*, Vol. 25, No. 1. pp. 16–25.

Burke, L. I., and Rangwala, S., (1991). "Tool Condition Monitoring in Metal Cutting: A Neural Network Approach," *Journal of Intelligent Manufacturing*, Vol. 2, No. 5, pp. 269–280.

Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A., (1983). *Graphical Methods for Data Analysis*, Boston: Duxbury Press.

Carpenter, G. A., and Grossberg, S., (1987). "ART2: Self-Organization of Stable Category Recognition Codes for Analog Input Patterns," *Applied Optics*, Vol. 26, pp. 4919–4930.

Carpenter, G. A., and Grossberg, S., (1991). "ART2-A: An Adaptive Resonance Algorithm for Rapid Category Learning and Recognition," *Neural Networks*, Vol. 4, No. 4, pp. 493–504.

Centner, R., (1964). *Final Report on Development of Adaptive Control Techniques for a Numerically Controlled Milling Machine*, Technical Documentary Report ML-TDR-64-279.

Chen, S., Billings, A., Cowan, C. F. N., and Grant, P. M., (1990). "Practical Identification of NARMAX models using Radial Basis Functions," *International Journal of Control*, 1990, Vol. 52, No. 6, pp. 1327–1350.

Chen, S, Cowan, C. F. N., and Grant, P. M., (1991). "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks," *IEEE Transactions on Neural Networks*, Vol. 2, No. 2, pp. 302–309.

Chittayil, K., (1994) *Acoustic Emission Sensing for Tool Wear Monitoring and Control in Metal Cutting*, Unpublished Ph.D. dissertation, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Part, PA, 16802.

Choi, G. S., Wang, Z. X., Dornfeld, D. A., and Tsujino, K., (1990). "Development of an Intelligent On-Line Tool Wear Monitoring System for Turning Operations," *Proceedings of the 1990 Japan-USA Symposium on Flexible Automation*, pp. 683–690.

Chow, J. G. C., (1984). *Sensor Development for On-Line Monitoring and The Determination of Temperature Distributions in Machining*, Ph.D. Disserta-

tion, Department of Mechanical Engineering, Carnegie-Mellon University, Pittsburgh.

Chryssolouris, G., and Domroese, M., (1988). "Sensor Integration for Tool Wear Estimation in Machining," in *Sensors and Controls for Manufacturing*, Kennatey Jr., E., Koren, T., and Stein, J. L. (eds.), New York: ASME Press Vol. PED 33, pp. 115–123.

Chryssolouris, G., Domroese, M., Beaulieu, P., (1992). "Sensor Synthesis for Control of Manufacturing Processes," *Journal of Engineering for Industry*, Vol. 114, No. 2, pp. 158–174.

Chui, C. K., (1992). *An Introduction to Wavelets*, Boston: Academic Press, Inc.

Cody, M. A., (1994). "The Wavelet Packet Transform," *Dr. Dobb's Journal*, Vol. 20, No.4, pp. 44–54.

Colwell, L. V., (1975). "Cutting Temperature Versus Tool Wear," *Annals of the CIRP*, Vol. 24, No. 1, pp. 73–76.

Colwell, L. V., Mazur, J. C., (1979). "Real Time Computer Diagnostics (A Research Tool for Metal Cutting)," *Annals of the CIRP*, Vol. 28, No. 1, pp. 49–52.

Colding, B., and Novak, A., (1981). "Sensing of Workpiece Diameter, Vibration and Out-Off-Roundness by Laser—Way to Automate Quality Control," *Annals of the CIRP*, Vol. 30, No. 1, pp. 473–476.

Cooley, J. W., and Tukey, J. W., (1965). "An Algorithm for Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, Vol. 19, pp. 297–301.

Cybenko, G., (1989). "Approximation by Superpositions of a Sigmoidal Function," *Mathematics of Control, Signals, and Systems*, Vol. 2, No. 4, pp. 303–314.

Dan, L., and Mathew, J., (1990). "Tool Wear and Failure Monitoring Techniques for Turning—A Review," *International Journal of Machine Tools Manufacture*, Vol. 30, No. 4, pp. 579–598.

Danai, K., and Ulsoy, A. G., (1987). "A Dynamic State Model for On-Line Tool Wear Estimation in Turning," *Journal of Engineering for Industry*, Vol. 109, No. 4, pp. 396–399.

Danai, K., and Ulsoy, A. G., (1988a). "An Adaptive Observer for On-Line Tool Wear Estimation in Turning—Part I: Theory," *Proceedings of the 1988 American Control Conference*, Vol. 3, pp. 1930–1937.

Danai, K., and Ulsoy, A. G., (1988b). "An Adaptive Observer for On-Line Tool Wear Estimation in Turning—Part II: Result," *Proceedings of the 1988 American Control Conference,* Vol. 3, pp. 1938–1944.

Daubechies, I., (1988). "Orthonormal Bases of Compactly Supported Wavelets," *Communications in Pure and Applied Mathematics,* Vol. 41, pp. 909–996.

Daubechies, I., (1990). "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *Transactions on Information Theory,* Vol. 36, No. 5, pp. 961–1005.

Daubechies, I., (1992). *Ten Lectures on Wavelets,* Philadelphia: Society for Industrial and Applied Mathematics.

DeFilippi, A, and Ippolito, R., (1969). "Adaptive Control in Turning: Cutting Forces and Tool Wear Relationship for P10, P20, P30 Carbides," *Annals of the CIRP,* Vol. 17, pp. 377–385.

Devijver, P. A., and Kittler, J., (1982). *Pattern Recognition—A Statistical Approach,* Englewood Cliffs: Prentice Hall.

Diei, E. N., and Dornfeld, D. A., (1987). "A Model of Tool Fracture Generated Acoustic Emission During Machining," *Journal of Engineering for Industry,* Vol. 109, No. 3, pp. 227–233.

Dornfeld, D. A., (1984). "The Role of Acoustic Emission in Manufacturing Process Monitoring," *Proceedings of the Conference on Sensor Technology for Untended Manufacturing,* SME Technical Paper MS84-924, pp. 69–74.

Dornfeld, and D. DeVries, F., (1990). "Neural Network Sensor Fusion for Tool Condition Monitoring," *Annals of the CIRP,* Vol. 39, No. 1, pp. 101–105.

Dornfeld, D. A., and Lan, M. S., (1983). "Chip Form Detection Using Acoustic Emission," *Proceedings of the 11th North American Manufacturing Research Conference (NAMRC),* pp. 386–389.

Du, R., and Yan, D., (1991). "Time-Frequency Distribution of Acoustic Emission Signal for Tool Wear Detection in Turning," in *Proceedings of the 4th World Meeting on Acoustic Emission (AEWG-35) and 1st International Conference on Acoustic Emission in Manufacturing,* Vahaviolos, S. J. (ed.), Columbus: American Society for Nondestructive Testing, pp. 269–285.

Dunteman, G. H., (1989). *Principal Components Analysis,* London: Sage Publications.

313

Elanayar, S., Shin, Y. C., and Kumara, S. R. T., (1990). "Machining Condition Monitoring for Automation using Neural Networks," in *Monitoring and Control for Manufacturing Processes*, Liang, S. Y., and Tsao, T-C. (eds.), New York: ASME Press, pp. 85–95.

El Gomayel, J. I., and Bregger, K. D., (1986). "On-Line Tool Wear Sensing for Turning Operations," *Journal of Engineering for Industry*, Vol. 108, No. 1, pp. 44–47.

Eman, K. F., and Wu, S. M., (1987). "Present and Future Trends in Stochastic Analysis of Cutting and Structural Dynamics," *Proceedings of the 15th North American Manufacturing Research Conference (NAMRC)*, pp. 426–432.

Emel, E., (1991). "Tool Wear Detection By Neural Network Based Acoustic Emission Sensing," in *Control of Manufacturing Processes*, Malkin, S., and Danai, K. (eds.), New York: ASME Press, pp. 79–85.

Emel, E., and Kannatey-Asibu, Jr., E., (1986). "Characterization of Tool Wear and Breakage by Pattern Recognition Analysis of Acoustic Emission Signals," *Proceedings of the 14th North American Manufacturing Research Conference (NAMRC)*, pp. 266–272.

Emel, E., and Kannaty-Asibu, Jr., E., (1988). "Tool Failure Monitoring in Turning by Pattern Recognition Analysis of AE Signals," *Journal of Engineering Industry*, Vol. 110, No. 2, pp. 137–145.

Forsythe, G. E., Malcolm, M. A., and Moler, C. B., (1977). *Computer Methods for Mathematical Computations*, Englewood Cliffs: Prentic-Hall.

Geirsson, A., (1990). *Monitoring System for a Diamond Turning Machine*, Ph.D. Dissertation, Department of Mechanical Engineering, Stanford University.

Gelb, A., (ed.), (1974). *Applied Optimal Estimation*, Cambridge: The MIT Press.

Ghosh, J., and Chakravarthy, S. V., (1992). "The Rapid Kernel Classifier: A Link Between the Self-Organizing Feature Map and the Radial Basis Function Network," in *Intelligent Engineering Systems Through Artificial Neural Networks*, Dagli, C. H., Burke, L. I., and Shin, Y. C., (eds.), New York: ASME Press.

Goldburg, M., (1993). *Application of Wavelets to Quantization and Random Process Representations*, Ph.D. Dissertation, Stanford University, Stanford.

314

Groover, M. P., Karpovich, R. J., and Levy, E. K., (1977). "A Study of The Relationship Between Remote Thermocouple Temperatures and Tool Wear in Machining," *International Journal of Production Research*, Vol. 25, No. 2, pp. 129–141.

Hertz, J., Krogh, A., Palmer, R. G., (1992). *Introduction to the Theory of Neural Computation*, New York: Addision-Wesley Publishing Company.

Hillmer, S. C., and Tiao, G. C., (1979). "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," *Journal of the American Statistical Association*, Vol. 74, pp. 652–660.

Hinds, B. K., (1977). "Control of Tool Wear During Metal Cutting Using a Computer and On Line Measurements," *International Journal of Production Research*, Vol. 15, No. 3, pp. 291–301.

Holdway, R. M., (1989). "Enhancing Supervised Learning Algorithm Via Self-Organization," *Proceedings of IEEE INNS IJCNN*, Vol. 2, pp. 523–530.

Hornik, K. M., Stinchcombe, M., and White, H., (1989). "Multilayer Feedforward Networks Are Universal Approximaters," *Neural Networks*, Vol. 2, No. 5, pp. 359–366.

Huang, W. Y., and Lippmann, R. P., (1988). "Neural Net and Traditional Classifiers," in *Neural Information Processing Systems*, Anderson, D. Z. (ed.), New York: American Institute of Physics, pp. 387–396.

Inasaki, I., Aida, S., and Fukuoka, S., (1987). "Monitoring System for Cutting Tool Failure using and Acoustic Emission Sensor," *JSME International Journal*, Vol. 30, No. 261, pp. 523–528.

Inasaki, I., and Yonetsu, S., (1981). "In-Process Detection of Cutting Tool Damage by Acoustic Emission Measurement," *Proceedings of the 22th International Machine Tool Design and Research Conference*, pp. 261–268.

ISO, (1972). *Tool Life Testing with Single-Point Turning Tools*, ISO 5th Draft Proposal, ISO/TC 29/WGG22 (Secretariate 37), Vol. 91, March 1972.

Iwata, K., and Moriwaki, T., (1977). "An Application of Acoustic Emission Measurement to In-Process Sensing of Tool Wear," *Annals of the CIRP*, Vol. 26, No. 1, pp. 21–26.

Jackson, I. R. H., (1988). "Convergence Properties of Radial Basis Function," *Construction Approximation*, Vol. 4, pp. 243–264.

315

Jammu, V. B., Danai, K., and Malkin, S., (1993). "Unsupervised Neural Network for Tool Breakage Detection in Turning," *Annals of the CIRP*, Vol. 42, No. 1, pp. 67–70.

Jetly, S., (1984). "Measuring Cutting Tool Wear On-line: Some Practical Considerations," *Manufacturing Engineering*, Vol. 93, No. 1, pp. 55–60.

Jiang, C. Y., Zhang, Y. Z., and Xu, H. J., (1987). "In-Process Monitoring of Tool Wear Stage by the Frequency Band-Energy Method," *Annals of the CIRP*, Vol. 36, No. 1, pp. 45–48.

Jolliffe, I. T., (1986). *Principal Component Analysis*, New York: Springer Verlag.

Kakino, Y., (1980). "In-Process Detection of Tool Breakage of Monitoring Acoustic Emission," in *Proceedings of International Conference on Cutting Tool Materials*, Gorsler, F. W. (ed.), Metals Park, Ohio: The Society, pp. 29–43.

Kamarthi, S., Sankar, G. S., Cohen, P. H., and Kumara, S. R. T., (1991). "On-line Tool Wear Monitoring Using a Kohonen's Feature Map," in *Intelligent Engineering Systems through Artificial Neural Networks*, Dagli, C. H. Kumara, S. R. T., and Shin, Y. C. (eds.), New York: ASME Press, pp. 639–645.

Kannatey-Asibu, Jr., E., and Dornfeld, D. A., (1981). "Quantitative Relationships for Acoustic Emission from Orthogonal Metal Cutting," *Journal of Engineering for Industry*, Vol. 103, No. 3, pp. 330–340.

Kannatey-Asibu, Jr., E., and Dornfeld, D. A., (1982). "A Study of Tool Wear Using Statistical Analysis of Metal-Cutting Acoustic Emission," *Wear*, Vol. 76, No. 2, pp.247–261.

Kannatey-Asibu, Jr., E., and Emel, E., "Linear Discriminant Function Analysis of Acoustic Emission Signals for Cutting Tool Monitoring," *Mechanical Systems and Signal Processing*, Vol. 1, No. 4, pp. 333–347.

Kegg, R. L., (1984). "On-Line Machine and Process Diagnostics," *Annals of the CIRP,*, Vol. 32, No. 2, pp.469–473.

Koch, U., (1970). "Experimental and Theoretical Analysis of Lathe Tool Temperature in Oblique Cutting," *Proceedings of the 11th International Machine Tool Design and Research Conference*, pp. 533–540.

Kohonen, T., (1990). "The Self-Organizing Map," *Proceedings of the IEEE*, Vol. 78, No. 9, pp. 1464–1480.

Kong, S-G., and Kosko, B., (1991). "Differential Competitive Learning for Centroid Estimation and Phoneme Recognition," *IEEE Transactions on Neural Networks*, Vol. 2, No. 2, pp. 118–124.

Koopmans, L. H., (1974). *The Spectral Analysis of Time Series*, New York: Academic Press.

Koren, Y., and Lenz, E., (1972). "Mathematical Model for the Flank Wear While Turning Steel with Carbide Tools," *CIRP Proceedings on Manufacturing Systems*, Vol. 1., No. 2, pp. 127–139.

Koren, Y., (1978). "Flank Wear Model of Cutting Tools Using Control Theory," *Journal of Engineering for Industry*, Vol. 100, No. 1, pp. 103–109.

Koren, Y., (1989). "Adaptive Control Systems for Machining," *Manufacturing Review*, Vol. 2, No. 1, pp. 6–15.

Koren, Y., Ko, T.-R., Ulsoy, A. G., and Danai, K., (1991). "Flank Wear Estimation Under Varying Cutting Conditions," *Journal of Engineering for Industry*, Vol. 113, No. 2, pp. 300–307.

Kumara, S. R. T., Kamarthi, S. V., Bukkapatnam, S., and Lee, J., (1994). "Sensor Based Monitoring for Real-Time Quality Control in Manufacturing," *Proceedings of the AAAI 1994 Spring Symposium on Detecting and Resolving Errors in Manufacturing Systems*.

Lan, M. S., and Dornfeld, D. A., (1982). "Experimental Studies of Tool Wear via Acoustic Emission Analysis," *Proceedings of the 10th North American Manufacturing Research Conference*, pp. 305–311.

Lan, M. S., and Dornfeld, D. A., (1984). "In-Process Tool Fracture Detection," *Journal of Engineering Materials and Technology*, Vol. 106, No. 2, pp. 111–118.

Lan, M. S., and Dornfeld, D. A., (1986). "Acoustic Emission and Machining—Process Analysis and Control," *Advanced Manufacturing Processes*, Vol. 1, No. 1, pp. 1–21.

Lan, M. S., and Naerheim, Y., (1985). "Application of Acoustic Emission Monitoring in Machining," *Proceedings of the 13th North American Manufacturing Research Conference (NAMRC)*, pp. 310–313.

Leem, C. S., and Dreyfus, S. E., (1992). "A Neural Network with Automatic Feature Selection for Monitoring Tool Wear in Metal-Cutting," in *Intelligent*

317

Engineering Systems Through Artificial Neural Networks, Dagli, C. H., Burke, L. I., and Shin, Y. C. (eds.), New York: ASME Press, pp. 814–820.

Levy, E. K., Tsai, C. L., and Groover, M. P., (1976). "Analytical Investigation of The Effect of Tool Wear on The Temperature Variations in A Metal Cutting Tool," *Journal of Engineering for Industry*, Vol. 98, pp. 251–257.

Liang, S. Y., and Dornfeld, D. A., (1989). "Tool Wear Detection Using Time Series Analysis of Acoustic Emission," *Journal of Engineering for Industry*, Vol. 111, No. 3, pp. 199–205.

Liao, Y. S., (1986). "Development of A Monitoring Technique for Tool Change Purpose in Turning Operations," *Proceedings of the 26th International Machine Tool Design and Research Conference*, pp. 331–337.

Lindstrom, B., and Lindberg, B., (1983). "Measurement of Dynamic Cutting Forces in the Cutting Process, a New Sensor for In-Process Measurement," *Proceedings of the 24th International Machine Tool Design and Research Conference*, pp. 137–147.

Lipman, M. P., Nevis, B. E., and Kane, G. E., (1967). "A Remote Sensor Method for Determining Average Tool-Chip Interface Temperature in Metalcutting," *Journal of Engineering for Industry*, Vol. 89, pp. 333–338.

Lippmann, R. P., (1987). "Introduction To Computing With Neural Nets," *IEEE ASSP Magazine (Acoustics, Speech, and Signal Processing)*, Vol. 4, No. 2, pp. 4-22.

Lippmann, R. P., (1989). "Review of Neural Networks for Speech Recognition," *Neural Computation*, Vol. 1, No. 1, pp. 1–38.

Lister, P. M., and Barrow, G., (1986). "Tool Condition Monitoring Systems," *Proceedings of the 26th International Machine Tool Design and Research Conference*, pp. 271–288.

Mackinnon, R., Wilson, G. E., and Wilkinson, A. J., (1986). "Tool Condition Monitoring Using Multi-Component Force Measurement," *Proceedings of the 26th International Machine Tool Design and Research Conference*, pp. 317–324.

Mallat, S. G., (1989). "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions On Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674–693.

318

Martin, K. F., Brandon, J. A., Grosvenor, B. I., and Owen, A., (1986). "A Comparison of In-Process Tool Wear Measurement Methods in Turning," *Proceedings of the 26th International Machine Tool Design and Research Conference,* pp. 289-295.

Martin, P., Mutel, B., and Drapier, J. P., (1974). "Influence of Lathe Tool Wear on the Vibrations Sustained in Cutting," *Proceedings of the 15th International Machine Tool Design and Research Conference,* pp. 251-257.

Mathias, R. A., Boock, W., and Welch, A., (1980). "Adaptive Control: Monitoring and Control of Metal-Cutting Processes," *Technology of Machine Tools: A Survey of the State of the Art by the Machine Tool Task Force,* Vol. 4, pp. 7.13.1-7.13.8.

McAulay, A. D., and Li, J., (1992). "Wavelet Data Compression for Neural Network Preprocessing," *Proceedings of SPIE on Signal Processing, Sensor Fusion, and Target Recognition,* SPIE Vol. 1699, pp. 356-365.

Micheletti, G. F., (1968). "Tool Wear and Cutting Forces in Steel Turning," *Annals of the CIRP,* Vol. 16, pp. 353-360.

Micheletti, G. F., Koenig, W., and Victor, H. R., (1976). "In-Process Tool Wear Sensors for Cutting Operations," *Annals of the CIRP,* Vol. 25, No. 2, pp. 483-496.

Moody, J., and Darken, C. J., (1989). "Fast Learning in Networks of Locally-Tuned Processing Units," *Neural Computation,* Vol. 1, No. 2, pp. 281-294.

Moriwaki, T., (1983). "Application of Acoustic Emission Measurement to Sensing of Wear and Breakage of Cutting Tool," *Bulletin of the Japan Society of Precision Engineering,* Vol. 17, No. 3, pp. 154-160.

Moriwaki, T., and Tobito, M., (1990). "A New Approach to Automatic Detection of Life of Coated Tool Based on Acoustic Emission Measurement," *Journal of Engineering for Industry,* Vol. 112, No. 3, pp. 212-218.

Nason, G. P., and Silverman, B. S., (1993). *The Discrete Wavelet Transform in S,* Statistical Research Report 93:07, School of Mathematics Sciences, University of Bath, UK.

Nedeljković, V., (1993). "A Novel Multilayer Neural Networks Training Algorithm that Minimizes the Probability of Classification Error," *IEEE Transactions on Neural Networks,* Vol. 4, No. 4, pp. 650-659.

319

Neter, J., Wasserman, W., and Kutner, M. H., (1985). *Applied Linear Statistical Models*, Boston: Irwin.

Noori-Khajavi, A., and Komanduri, R., (1993). "On Multi-Sensor Approach to Drill Wear Monitoring," *Annals of the CIRP*, Vol. 42, No. 1, pp. 71–74.

Novak, A., and Ossbahr, G., (1986). "Reliability of the Cutting Force Monitoring in FMS-installations," *Proceedings of the 26th International Machine Tool Design and Research Conference*, pp. 325–329.

Ono, K., Stern, R., and Long, Jr., M., (1971). "Application of Correlation Analysis to Acoustic Emission," in *Acoustic Emission*, ASTM Special Technical Publication 505, Philadelphia: ASTM.

Oppenheim, A., and Schafer, R., (1975). *Digital Signal Processing*, New York: Printice-Hall.

Oraby, S. E., and Hayhurst, D. R., (1991). "Development of Models for Tool Wear Force Relationships in Metal cutting,"*International Journal of Machining Science*, Vol. 33, No. 2, pp. 125–138.

Pagano, M., (1974). "Estimation of Models of Autoregressive Signal Plus White Noise," *Annals of Statistics*, Vol. 2, pp. 99–108.

Pandit, S. M., and Kashou, S., (1982). "A Data Dependent Systems Strategy of On-Line Tool Wear Sensing," *Journal of Engineering for Industry*, Vol. 104, No. 3, pp. 217–223.

Park, J., and Sandberg, I. W., (1991). "Universal Approximation Using Radial Basis-Function Networks," *Neural Computation*, Vol. 3, No. 2, pp. 246–257.

Park, J.-J., and Ulsoy, A. G., (1989a). "On-Line Tool Wear Estimation Using Force Measurement and A Nonlinear Observer," in *Control Issues in Manufacturing Processes*, Stein, J. L., and Koren, Y., and Holmes, J. (eds.), New York: ASME, pp. 55–63.

Park, J.-J., and Ulsoy, A. G., (1989b). "Methods for Tool Wear Estimation From Force Measurement Under Varying Cutting Conditions," in *Control Issues in Manufacturing Processes*, Stein, J. L., and Koren, Y., and Holmes, J. (eds.), New York: ASME, pp. 45–53.

Park, J.-J., and Ulsoy, A. G., (1990). "On-Line Flank Wear Estimation Using Adaptive Observers," in *Automation of Manufacturing Processes*, Danai, K., and Malkin, S. (eds.), New York: ASME, pp. 13–22.

320

Park, J.-J., and Ulsoy, A. G., (1992). "On-Line Tool Wear Estimation Using Force Measurement and a Nonlinear Observer," in *Journal of Dynamic Systems, Measurement and Control,* Vol. 114, No. 4, pp 666–672.

Park, J.-J., and Ulsoy, A. G., (1993a). "On-Line Flank Wear Estimation Using an Adaptive Observer and Computer Vision. Part 1: Theory," *Journal of Engineering for Industry,* Vol. 115, No. 1, pp. 30–36.

Park, J.-J., and Ulsoy, A. G., (1993b). "On-Line Flank Wear Estimation Using an Adaptive Observer and Computer Vision. Part 2: Experiment," *Journal of Engineering for Industry,* Vol. 115, No. 1, pp. 37–43.

Peklenik, J., (1970). "Geometric Adaptive Control of Manufacturing Systems," *Annals of the CIRP,* Vol. 18, pp. 265–272.

Press, W. H., Flannery, B. P., Teukolosky, S. A., Vetterling, W. T., (1992). *Numerical Recipes in C, the Art of Scientific Computing,* New York: University Press.

Ralson, P. A. S., Ward, T. L., and Stottman, D. J. C., (1988). "Computer Observer for In-Process Measurement of Lathe Tool Wear," *Computers in Industrial Engineering,* Vol. 15, No. 1–4, pp. 217–222.

Raman, S., (1989). *Coupled Sensing for Tool Wear Modeling,* Ph.D. Dissertation, Department of Industrial and Management Systems Engineering, The Pennsylvania State University, University Park.

Rangwala, S., Dornfeld, D., (1990). "Sensor Integration Using Neural Networks for Intelligent Tool Condition Monitoring," *Journal of Engineering for Industry,* Vol. 112, No. 3, pp. 219–228.

Rao, B. S., (1986). "Tool Wear Monitoring Through the Dynamics of Stable Turning," *Journal of Engineering for Industry,* Vol. 108, No. 3, pp.183–190.

Renals, S., and Rohwer, R., (1989). "Phoneme Classification Experiment using Radial Basis Functions," *Proceedings of the IEEE INNS IJCNN,* Vol. 1, pp. 461–467.

Rioul, O., and Vetterli, M., (1991). "Wavelets and Signal Processing," *IEEE Signal Processing Magazine,* Vol. 8, No. 4, pp. 14–38.

Rumelhart, D. E., and McClelland, J. L., (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations* London: The MIT Press.

Salami, M. J. E., (1985). *Application of ARMA Models in Multicomponent Signal Analysis*, Ph.D. Dissertation, University of Calgary, Calgary, Alberta, Canada.

Sata, T., and Matsushima, K., (1974). "On-Line Control of the Cutting State by the Pattern Recognition Technique," *Annals of the CIRP*, Vol. 23, No. 1, pp. 151–152.

Schaffer, G., (1983). "Sensors: The Eyes and Ears of CIM," *American Machinist*, Vol. 127, No. 7., pp 109–124.

Shillam, N. F., (1971). "The On-line Control of Cutting Condition Using Direct Feedback," *Proceeding of the 12th International Machine Tool Design and Research Conference*, pp. 15–21.

Sokolowski, A., Liu, J. J., and Kosmol, J., (1992). "On the Correlation Between the Vibration Measurement and Tool Wear in Turning," *Proceeding of the 1992 JAPAN/USA Symposium on Flexible Automation*, Vol. 2, pp. 1075–1081.

Solaja, V., and Vukelja, D., (1973). "Identification of Tool Wear Rate by Temperature Variation of a Carbide Tip," *Annals of the CIRP*, Vol. 22, No. 1, pp. 5–6.

Spirgeon, D., and Slater, R. A. C., (1974). "In-Process Indication of Surface Roughness Using A Fiber-Optics Transducer," *Proceedings of the 15th International Machine Tool Design and Research Conference*, pp. 339–347.

Suzuki, H., and Weinamann, K. J., (1985). "An On-Line Tool Wear Sensor for Straight Turning Operations," *Journal of Engineering for Industry*, Vol. 107, No. 4, pp.397–399.

Szu, H. H., Telfer, B., and Kadambe, S., (1992). "Neural Network Adaptive Wavelets for Signal Representation and Classification," *Optical Engineering*, Vol. 31, No. 9, pp. 1907–1916.

Taglia, A. D., Portunato, S., and Toni, P., (1976). "An Approach to On-Line Measurement of Tool Wear by Spectrum Analysis," *Proceedings of the 17th International Machine Tool Design and Research Conference*, pp. 141–148.

Takata, S., and Sata, T., (1986). "Model Referenced Monitoring and Diagnosis–Application to Manufacturing System," *Computers in Industry*, Vol. 7, No. 1, pp. 31–43.

322

Takeyama, H., Doi, Y., Mitsoka, T., and Sekiguchi, H., (1967). "Sensors of Tool Life for Optimization of Machining," *Proceedings of the 8th International Machine Tool Design and Research Conference*, pp. 191–208.

Takeyama, H., Sekiguchi, H., Murata, R., and Matsuzaki, H., (1976). "In-process Detection of Surface Roughness in Machining," Annals of the CIRP, Vol. 25. No. 1, pp. 467–471.

Takeyama H., Sekiguchi, H, and Takada, K., (1970). "One Approach for Optimizing Control in Metal Cutting," *Annals of the CIRP*, Vol. 18, pp. 345–351.

Tansel, I. N., and McLaughlin, C., (1993). "Detection of Tool Breakage in Milling Operations-II. The Neural Network Approach," *International Journal of Machine Tools and Manufacturing*, Vol. 33, No. 4, pp. 545–558.

Tansel, I. N., Mekdeci, C., Rodriguez, O., and Uragun, B., (1993). "Monitoring Drill Conditions with Wavelet Based Encoding and Neural Networks," *International Journal of Machine Tools and Manufacturing*, Vol. 33, No. 4, pp. 559–575.

Taraman, K., Swando, R., and Yamauchi, W., (1974). "Relationship Between Tool Forces and Flank Wear," *SME Technical Paper*, No. MR74-704.

Teti, R., and Dornfeld, D.,(1989). "Modeling and Experimental Analysis of Acoustic Emission from Metal Cutting," *Journal of Engineering for Industry*, Vol. 111, No. 3, pp. 229–237.

Teti, R., and Micheletti, G. F., (1989). "Tool Wear Monitoring Through Acoustic Emission," *Annals of The CIRP*, Vol. 38, No. 1, pp. 99–102.

Tiao, G. C., and Box, G. E. P., (1981). "Modeling Multiple Time Series with Applications," *Journal of the American Statistical Association*, Vol. 76, pp. 802–816.

Tlusty, J., and Andrews, G. C., (1983). "A Critical Review of Sensors for Unmanned Machining," Annals of the CIRP, Vol. 32. No. 2, pp.563–572.

Turkovich, B. F., and Kramer, B. M., (1986). "A Comprehensive Tool Wear Model," *Annals of the CIRP*, Vol. 35, No. 1, pp. 67–70.

Uehara, K., Kiyosawa, F., and Takeshita, H., (1979). "Automatic Tool Wear Monitoring in NC Turning," *Annals of the CIRP*, Vol. 28, No. 1, pp. 39–42.

Usui, E., Shirakashi, T., and Kitagawa, T., (1984). "Analytical Prediction of Cutting Tool Wear," *Wear*, Vol. 100, pp. 129–151.

323

Wang, Z., and Dornfeld, D. A., (1992). "In-Process Tool Wear Monitoring Using Neural Networks," *Proceedings of 1992 Japan-USA Symposium on Flexible Automation*, Vol. 1, pp. 263-270.

Weiier, E. J., Schrier, H. M., and Weichbrot, B., (1969). "What Sound can Be Expected from a Worn Tool?" *Journal of Engineering for Industry*, Vol. 91, No. 3, pp. 525-534.

Whitehouse, D. J., Hingle, H. T., and Rakels, J. H., (1983). "The Practical Application of Diffraction Technique to Assess Surface Finish of Diamond Turned Parts," *Annals of the CIRP*, Vol. 32, No. 1, pp. 499-501.

Whitney, A., (1971). "A Direct Method of Non-parametric Measurement Selection," *IEEE Transactions on Computers*, Vol. 20, pp. 1100-1103.

Wilson, G. T., (1973). "The Estimation of Parameters in Multivariate Time Series Models," *Journal of the Royal Statistical Society*, Series B, Vol. 35, pp. 76-85.

Wold, H., (1954). *A study in the Analysis of Stationary Time Series*, Stockholm: Almqvist & Wiksell.

Xu, L., Krzyzak, A, and Oja, E., (1993). "Rival Penalized Competitive Learning for Cluster Analysis, RBF Net, and Curve Detection," *IEEE Transactions on Neural Networks*, Vol. 4, No. 4, pp. 636-649.

Yao, Y., and Fang, X. D., (1992). "Modeling of Multivariate Timer Series for Tool Wear Estimation in Finish-Turning," *International Journal of Machine Tools and Manufacturing*, Vol. 32, No. 4, pp. 495-508.

Zhang, Q., and Benveniste, A., (1991). "Approximation by Nonlinear Wavelet Networks," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 3417-3420.

Zhang, Q., and Benveniste, A., (1992). "Wavelet Networks," *IEEE Transaction on Neural Networks*, Vol. 3, No. 6, pp. 889-898.

324

Appendix A

CODE FOR PRINCIPAL COMPONENT ANALYSIS

```
#include "mar.h"


#define NORMALIZE    0     /* Set 1, if inputs need to be normalized */
#define NSERIES      3     /* Number of series to be modeled         */
#define NPOINTS      512   /* Number of points in series             */
#define NPAT         10    /* Number of patterns                     */
#define AR_ORDER     20    /* Order of the MAR model of the series   */

#define ROWS         NPOINTS-AR_ORDER
#define COLS         NSERIES*AR_ORDER

double **Series;
double **Res_Series;
double **AR_Coff;
double **Cov_Matrix;
double **Design_Matrix;
double **Measurement_Matrix;
double **Matrix_U;
double **Matrix_V;
double *Vector_W;
double **Matrix_UU;
double **Matrix_VV;
double *Vector_WW;
double *Sd_Dev;
double *Series_Eval;
double **Series_Evec;
double *Res_Ser_Eval;
double **Res_Ser_Evec;

typedef struct {
    double  Data[NSERIES][NPOINTS];
}REC;

/*------------------ Principal Component Analysis ------------------*/

main ()
{
int rec,      /* current record number                      */
    scn,      /* current series component number            */
    i;        /* Index for elements of the input patterns   */

double Prediction_Error[NSERIES];

FILE *Series_fp;
```

325

```
FILE *Eval_Ser_fp;
FILE *Eval_Res_Ser_fp;
FILE *Res_Ser_fp;
FILE *Coff_fp;
FILE *Pred_Err_fp;

Initialize_Series();
Initialize_Res_Series();
Initialize_AR_Coff();
Initialize_Cov_Matrix();
Initialize_Design_Matrix();
Initialize_Measurement_Matrix();
Initialize_Matrix_U();
Initialize_Matrix_V();
Initialize_Vector_W();
Initialize_Matrix_UU();
Initialize_Matrix_VV();
Initialize_Vector_WW();
Initialize_Sd_Dev();
Initialize_Series_Eigenvalues();
Initialize_Series_Eigenvectors();
Initialize_Res_Ser_Eigenvalues();
Initialize_Res_Ser_Eigenvectors();

Series_fp = Open_Input_Series_File ();
Eval_Ser_fp = Open_Series_Eval_File_To_Write();
Eval_Res_Ser_fp = Open_Res_Ser_Eval_File_To_Write();
Coff_fp = Open_AR_Coff_File_To_Write();
Pred_Err_fp = Open_Prediction_Error_File_To_Write();

Res_Ser_fp = Open_Res_Series_File_To_Write();

for(rec=0; rec< NPAT; rec++) {
    Read_Record_From_Series_File(rec, Series, Series_fp);
    Shift_Series_To_Zero_Mean(Series,  NPOINTS);
    Build_Design_Matrix(Design_Matrix);
    Build_Measurement_Matrix(Measurement_Matrix);
    for(scn=0; scn< NSERIES; scn++) {
        Find_AR_Coefficients (Design_Matrix, Measurement_Matrix[scn],
        AR_Coff[scn], ROWS, COLS, Matrix_U, Matrix_V, Vector_W, Sd_Dev);
    }
    Compute_Residue_Series(Series, Res_Series, Prediction_Error);
    Shift_Series_To_Zero_Mean(Res_Series, ROWS);

    Find_Eigenvalues_Of_Covariance_Matrix(Series, Series_Eval,
                                            Series_Evec, NPOINTS);
    Find_Eigenvalues_Of_Covariance_Matrix(Res_Series, Res_Ser_Eval,
                                            Res_Ser_Evec,ROWS);

    Compute_Covariance_Eigenvalues(Series, Series_Eval,
                        Series_Evec, NPOINTS);
```

326

```
        Compute_Covariance_Eigenvalues(Res_Series, Res_Ser_Eval,
                                Res_Ser_Evec, ROWS);
        Write_Series_Eigenvalues(rec, Series_Eval, Eval_Ser_fp);
        Write_Res_Ser_Eigenvalues(rec, Res_Ser_Eval, Eval_Res_Ser_fp);
        Write_Res_Series(rec, Series, Res_Ser_fp);

        Write_AR_Coff(rec, AR_Coff, Coff_fp);
        Write_Prediction_Error(rec, Prediction_Error, Pred_Err_fp);
}
Write_Cpu_Time("mar.c", "cpu_mar.out");

fclose (Series_fp);
fclose (Eval_Ser_fp);
fclose (Eval_Res_Ser_fp);
fclose (Res_Ser_fp);
fclose (Coff_fp);

}
```

327

Appendix B

SIG CODE TO BANDPASS FILTER FEED FORCE DATA

```
!-------------------- filter for ff0.dat --------------------!
!
tsread 1 ff0.dat (e12.6)    0.0003333333 0
demean 1 2
bpbessel 2 3 3 1000 7
tswrite 3 ff0.dat (e12.6)
!
!
                            * * *
!
!
!-------------------- filter for ffn.dat --------------------!
!
tsread 1 ff1.dat (e12.6)    0.0003333333 0
demean 1 2
bpbessel 2 3 3 1000 7
tswrite 3 ff1.dat (e12.6)
```

328

Appendix C

SIG CODE TO BANDPASS FILTER FEED VIBRATION DATA

```
!--------------------- filter for fv0.dat ---------------------!
!
tsread 1 fv0.dat (e12.6)   0.0000384615 0
demean 1 2
bpbessel 2 3 1000 10000 7
tswrite 3 fv0.dat (e12.6)
!
!
                          * * *
!
!
!--------------------- filter for fvn.dat ---------------------!
!
tsread 1 fv1.dat (e12.6)   0.0000384615 0
demean 1 2
bpbessel 2 3 1000 10000 7
tswrite 3 fv1.dat (e12.6)
```

329

Appendix D

SIG CODE TO BANDPASS FILTER AE DATA

```
!------------------- filter for ae0.dat -------------------!
!
tsread 1 ae0.dat (e12.6)   0.0000010000 0
demean 1 2
bpbessel 2 3 50000 300000 7
tswrite 3 ae0.dat (e12.6)
!
!
                            * * *
!
!
!------------------- filter for aen.dat -------------------!
!
tsread 1 ae1.dat (e12.6)   0.0000010000 0
demean 1 2
bpbessel 2 3 50000 300000 7
tswrite 3 ae1.dat (e12.6)
```

330

Appendix E

SCA CODE FOR DETERMINING ORDER OF UNIVARIATE ARMA MODELS FOR CUTTING, FEED, AND RADIAL FORCE DATA

```
==SCANPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'for0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--

                            * * *

--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'forn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDFOR,RADIALFOR,MAINFOR.              @
FILE IS &FILENAME.                                         @
PRECISIONS ARE DOUBLE, DOUBLE, DOUBLE.                     @
                         NCOLS IS 1, 1, 1. NROW IS &NUMROWS.
--
SCAN VARIABLE IS FEEDFOR. MAXLAG IS AR(10), MA(15).--
--
SCAN VARIABLE IS RADIALFOR. MAXLAG IS AR(10), MA(15).--
--
SCAN VARIABLE IS MAINFOR. MAXLAG IS AR(10), MA(15).--
--
RESTART--
--
RETURN--
//--
```

331

SCA CODE FOR DETERMINING ORDER OF MULTIVARIATE ARMA MODELS FOR CUTTING, FEED, AND RADIAL FORCE DATA

```
==MSCANPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'for0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
                            * * *
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'forn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDFOR,RADIALFOR,MAINFOR.                 @
FILE IS &FILENAME.                                            @
PRECISIONS ARE DOUBLE, DOUBLE, DOUBLE.                        @
                            NCOLS IS 1, 1, 1. NROW IS &NUMROWS.
--
SCAN VARIABLE ARE FEEDFOR,RADIALFOR,MAINFOR.                  @
                            MAXLAG IS AR(10), MA(15).
--
RESTART--
--
RETURN--
//--
```

SCA CODE FOR DETERMINING ORDER OF UNIVARIATE ARMA MODELS FOR CUTTING AND FEED VIBRATION DATA

```
==SCANPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vib0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
                         * * *
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vibn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDVIB, MAINVIB. FILE IS &FILENAME.        @
PRECISIONS ARE DOUBLE, DOUBLE.                                  @
                    NCOLS IS  1, 1. NROW IS &NUMROWS.
--
SCAN VARIABLE IS FEEDVIB. MAXLAG IS AR(10), MA(15).--
--
SCAN VARIABLE IS MAINVIB. MAXLAG IS AR(10), MA(15).--
--
RESTART--
--
RETURN--
//--
```

Appendix H

SCA CODE FOR DETERMINING ORDER OF MULTIVARIATE ARMA MODELS FOR CUTTING AND FEED VIBRATION DATA

```
==MSCANPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vib0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--

                          * * *
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vibn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDVIB, MAINVIB. FILE IS &FILENAME.        @
PRECISIONS ARE DOUBLE, DOUBLE.                                  @
                        NCOLS IS  1, 1. NROW IS &NUMROWS.
--
SCAN VARIABLE ARE FEEDVIB, MAINVIB.                            @
                        MAXLAG IS AR(10), MA(15).
--
RESTART--
--
RETURN--
//--
```

Appendix I

SCA CODE FOR UNIVARIATE ARMA MODELING OF CUTTING, FEED, AND RADIAL FORCE DATA

```
==ARMAPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'for0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--

                        * * *        .

--            ~.
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'forn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDFOR,RADIALFOR,MAINFOR. FILE IS &FILENAME.@
PRECISIONS ARE DOUBLE, DOUBLE, DOUBLE.                          @
                        NCOLS IS  1, 1, 1. NROW IS &NUMROWS.
--
TSMODEL NAME IS MF_SENSOR.                                      @
MODEL IS (1-P1*B-P2*B**2-P3*B**3)MAINFOR =                      @
                        (1-T1*B-T2*B**2-T3*B**3)NOISE.
--
TSMODEL NAME IS FF_SENSOR.                                      @
MODEL IS (1-P1*B-P2*B**2-P3*B**3)FEEDFOR =                      @
                        (1-T1*B-T2*B**2-T3*B**3)NOISE.
--
TSMODEL NAME IS TF_SENSOR.                                      @
MODEL IS (1-P1*B-P2*B**2-P3*B**3)RADIALFOR =                    @
                        (1-T1*B-T2*B**2-T3*B**3)NOISE.
--
ESTIM MODEL IS MF_SENSOR. OUTPUT IS LEVEL(BRIEF).--
ESTIM MODEL IS FF_SENSOR. OUTPUT IS LEVEL(BRIEF).--
ESTIM MODEL IS TF_SENSOR. OUTPUT IS LEVEL(BRIEF).--
--
RESTART--
--
RETURN--
//--
```

335

SCA CODE FOR MULTIVARIATE ARMA MODELING OF CUTTING, FEED, AND RADIAL FORCE DATA

```
==MARMAPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'for0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--

                        * * *
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'forn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDFOR,RADIALFOR,MAINFOR. FILE IS &FILENAME.@
PRECISIONS ARE DOUBLE, DOUBLE, DOUBLE.                          @
                        NCOLS IS  1, 1, 1. NROW IS &NUMROWS.
--
MTSMODEL NAME IS FORSENSOR. SERIES ARE FEEDFOR,RADIALFOR,MAINFOR.@
MODEL IS (I-P1*B-P2*B**2-P3*B**3)SERIES =                       @
                        (I-T1*B-T2*B**2-T3*B**3)NOISE.
--
MESTIM MODEL IS FORSENSOR. OUTPUT IS LEVEL(BRIEF).--
--
RESTART--
--
RETURN--
//--
```

336

SCA CODE FOR UNIVARIATE ARMA MODELING OF CUTTING AND FEED VIBRATION DATA

```
==ARMAPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vib0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
                         * * *
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vibn dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDVIB, MAINVIB. FILE IS &FILENAME.        @
PRECISIONS ARE DOUBLE, DOUBLE.                                  @
                         NCOLS IS 1, 1. NROW IS &NUMROWS.
--
TSMODEL NAME IS M_VIBSENSOR.                                    @
MODEL IS (1-P1*B-P2*B**2-P3*B**3-P4*B**4-P5*B**5-P6*B**6)MAINVIB @
                    (1-T1*B-T2*B**2-T3*B**3-T4*B**4)NOISE.
--
TSMODEL NAME IS F_VIBSENSOR.                                    @
MODEL IS (1-P1*B-P2*B**2-P3*B**3-P4*B**4-P5*B**5-P6*B**6)FEEDVIB @
                    (1-T1*B-T2*B**2-T3*B**3-T4*B**4)NOISE.
--
ESTIM MODEL IS M_VIBSENSOR. OUTPUT IS LEVEL(BRIEF).--
-
ESTIM MODEL IS F_VIBSENSOR. OUTPUT IS LEVEL(BRIEF).--
--
RESTART--
--
RETURN--
//--
```

SCA CODE FOR MULTIVARIATE ARMA MODELING OF CUTTING AND FEED VIBRATION DATA

```
==MARMAPROC
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vib0 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--

                        * * *

--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
ASSIGN FILE IS 20. EXTERNAL-NAME IS 'vib1 dat b'.--
CALL ARMA_ONE. SYMBOLIC-VALUES ARE NUMROWS(512), FILENAME(20).--
--
RETURN--


==ARMA_ONE
--
PROFILE REVIEW. STYLE IS BATCH. PRECISION IS DOUBLE.--
PARAMETER SYMBOLIC-VARIABLES ARE NUMROWS(512), FILENAME(20)--
--
INPUT VARIABLES ARE FEEDVIB, MAINVIB. FILE IS &FILENAME.        @
PRECISIONS ARE DOUBLE, DOUBLE.                                  @
                        NCOLS IS 1, 1. NROW IS &NUMROWS.
--
MTSMODEL NAME IS VIBSENSOR. SERIES ARE FEEDVIB, MAINVIB.        @
MODEL IS (I-P1*B-P2*B**2-P3*B**3-P4*B**4-P5*B**5-P5*B**5)SERIES= @
                        (I-T1*B-T2*B**2-T3*B**3-T4*B**4)NOISE.
--
MESTIM MODEL IS VIBSENSOR. OUTPUT IS LEVEL(BRIEF).--
--
RESTART--
--
RETURN--
//--
```

338

CODE FOR DISCRETE WAVELET TRANSFORM

```
#include "wl.h"

#define NPOINTS        256
#define NPAT           41

#define DC_START       192          /*  DC_START >= NPOINTS/2       */
#define DC_END         224          /*  DC_END <= NPOINTS-1         */

#define ALPHA_DEG      45.0         /* 0 <= ALPHA_DEG < 360         */
#define BETA_DEG       90.0         /* 0 <= BETA_DEG < 360          */
#define DIS_COLS       7

#define LEVELS         ((int) log2((double) NPOINTS))

#define ALPHA          (ALPHA_DEG*0.0174533)  /* -pi <= ALPHA < pi  */
#define BETA           (BETA_DEG*0.0174533)   /* -pi <= BETA < pi   */

typedef enum {DECOMP, RECON} TRANSFORM_TYPE;
typedef enum {HAAR, CODY, DEB2, DEB3, DEB4,
                   DEB5, DEB6, DEB7, DEB8, DEB9, DEB10} WL_TYPE;

int    NCOEFF;
double *Signal;
double *Rec_Signal;
double *Wave_Coeff;
double *LP_Filter;
double *HP_Filter;
double *Detailed_Coeff;
double **Expansion_Coeff;

/*------------------- Discrete Wavelet Transform -------------------*/
main ()
{

TRANSFORM_TYPE Transform=DECOMP;
WL_TYPE Wavelet_Type=DEB7;
FILE    *Filelist_fp;
FILE    *DCoeff_fp;
char    Filename[80];
int     np;
int     Filter_Len;

Filelist_fp = Open_Filelist_File_To_Read();
DCoeff_fp = Open_Detailed_Coeff_File_To_Write();
Get_Ncoeff(Wavelet_Type);
```

339

```
Initialize_Signal((NPOINTS+NCOEFF));
Initialize_Rec_Signal((NPOINTS+NCOEFF));
Initialize_Wave_Coeff(NCOEFF);
Initialize_LP_Filter(NCOEFF);
Initialize_HP_Filter(NCOEFF);
Initialize_Expansion_Coeff_Tree(NPOINTS, LEVELS);
Initialize_Detailed_Coeff(NPOINTS);
Generate_Wavelet_Coefficients(Wavelet_Type, Wave_Coeff);
Filter_Len = Make_Wavelet_Filters(Wave_Coeff, LP_Filter,
                                          HP_Filter, Transform);
for(np=0; np< NPAT; np++) {

    Get_Filename(Filename, Filelist_fp, np);

    Read_AE_Signal(Signal, Filename, NPOINTS);

    Tree_Zero(Expansion_Coeff, NPOINTS, LEVELS);

    Wavelet_Decomposition(Signal, NPOINTS, LP_Filter, HP_Filter,
                                  Filter_Len, LEVELS, Expansion_Coeff);

    Copy_Detailed_Coeff(Expansion_Coeff, Detailed_Coeff, NPOINTS, LEVELS);

/*
 * Threshold_Detailed_Coeff(Detailed_Coeff, DC_START, DC_END);
 */

    Write_Detailed_Coeff(Detailed_Coeff, np, DC_START,
                                            DC_END, DCoeff_fp);

/*
 * Wavelet_Reconstruction(Expansion_Coeff, NPOINTS, LP_Filter,
 *                        HP_Filter, Filter_Len, LEVELS, Rec_Signal);
 *
 * Write_Signal_And_Reconstructed_Signal(Signal, Rec_Signal, NPOINTS);
 */

}
fclose(Filelist_fp);
fclose(DCoeff_fp);
}
```

340

CODE FOR KOHONEN'S FEATURE MAP

```c
#include "rbf.h"


#define RAND_MAX        2147483648.0
#define ONODES          196     /* Number of output nodes             */
#define HEIGHT          14      /* Number of rows in output map       */
#define LENGTH          14      /* Number of columns in output map    */
#define START_RADIUS    7       /* Initial neighborhood radius        */
#define NIT             5000    /* Number of iterations               */
#define FIRSTPHASE      2000    /* No. of iterations in 1st phase      */
#define DISPLAY_FRQ     200     /* Frequency of recording map         */
#define NORMALIZE       0       /* Set 1, if inputs & wts. normalized */
#define ALL_CENTERS     0       /* Set 1, if all output node be used  */
#define KFM_COV         0       /* Set 1, if cov of centers be modified*/
#define WT_READ         0       /* Set 1, if wts read from a file     */
#define FRESH_TOOL      0.010   /* Wear threshold for fresh tool      */
#define WORN_TOOL       0.015   /* Wear thr. for worn out tool        */
#define STD_DEV         0.5     /* SDV of initial random weights      */
#define MEAN            0.0     /* Mean of initial random weights     */

int INODES;                     /* Number of input nodes              */
int NPAT;                       /* Number of training patterns        */
int NRBFC;                      /* Number of Rbf centers              */
double **Input_Pattern;
double *Wear;
double **Node_Weights;
double **Rbf_Center;
double ***Covariance;
double **Rbfc_Cov_Diag;
double *Activation;
char *Display;
FILE *map_fp;

int Bins[ONODES][3];

/*--------------------- Kohonen's Feature Map ------------------------*/
main ()
{

double Curr_Gain;       /* Gain of the learning equations            */

int Curr_It,            /* Loop variable for number of iterations    */
    Curr_Radius,        /* Radius of neighborhood of a winning node  */
    Winning_Node,       /* Winning node which receives max. output   */
    np;                 /* Index variable for number of patterns     */
```

```
map_fp = fopen("map.log","w");

Read_Input_Patterns();
Initialize_Output_Activation();
if(KFM_COV) {
    Initialize_Covariance_Matrix();
}
Initialize_Display_Array();
if (WT_READ) {
    Read_Node_Weights();
}
else {
    Initialize_Weights();
}
if (NORMALIZE) {
    Normalize_Inputs();
    Normalize_Weights();
}
Curr_It=0;

do {
    ++Curr_It;
    Set_Current_Gain(Curr_It, &Curr_Gain);
    Set_Current_Radius(Curr_It, &Curr_Radius);
    for(np=0;np<NPAT;np++) {
        Find_Winning_Node_MD(np, &Winning_Node);
        Modify_Wts_Around_Winning_Node(np,Winning_Node,Curr_Radius,
                                       Curr_Gain);
        Label_Winning_Node_Display(np,Winning_Node);
    }
    Display_Feature_Map(Curr_It);
} while (Curr_It<NIT);

fclose(map_fp);
Find_Rbf_Centers();
Write_Cpu_Time("rbf_kfm.c", "cpu_kfmrbf.out");
Write_Node_Weights();
Write_Rbf_Centers();
Write_Rbfc_Cov_Diag();
}
```

CODE FOR TRAINING RADIAL BASIS FUNCTION NETWORK

```
#include "rbf.h"
#include "bp.h"


#define NPAT        168        /* Number of training patterns     */
#define INODES      32         /* Number of input nodes           */
#define ONODES      1          /* Number of nodes in output layer */
#define OSCALE      50.0       /* Scale factor for network output */
#define SLOP        1.0        /* It defines slope of sigmoidal fn. */
#define GAIN        0.90       /* Gain for modifying wts and thresh. */
#define NIT         5000       /* Number of iterations            */
#define DIS_FRQ     10         /* Frequency of writing Error      */
#define NORMALIZE   1          /* Set 1, if inputs be normalized  */
#define USE_KFM_COV 0          /* Set 1, if kfm cov values be used */
#define USE_COMP_COV 1         /* Set 1, if computed cov be used  */
#define COV_RADIUS  0.5        /* used if,USE_KFM_COV=USE_COMP_COV=0 */
#define WT_READ     0          /* Set 1, if wts read from a file  */
#define SATURATION  100.0      /* Saturation limit of net input   */

#define RAND_MAX    2147483648.0
#define MEAN        0.0        /* Mean of initial random weights  */
#define STD_DEV     0.5        /* SDV of initial random weights   */

int NRBFC;                     /* Number of Rbf centers           */
int NNRBFC;                    /* Number of Rbf centers plus 1    */
int PITER;

double **Input_Pattern;
double **Target_Pattern;
double **Rbf_Center;
double **Rbfc_Activation;
double **Rbfc_Cov_Diag;
double **Rbfn_Weights;
double *Rbfn_Outputs;
double *Rbfn_Onode_Bpe;
double *Wear;

extern void    Initialize_Rbfn_Weights();
extern void    Read_Rbfn_Weights();
extern void    Initialize_Rbfn_Outputs();
extern void    Read_Rbf_Centers();
extern void    Normalize_Input_Patterns();
extern void    Normalize_Rbf_Centers();
extern void    Compute_Rbfc_Activation_Matrix();
extern void    Read_Rbfc_Cov_Diag();
```

343

```
extern void    Compute_Rbfc_Cov_Diag();
extern void    Compute_Rbfn_Output_Activation();
extern void    Modify_Rbfn_Weights();
extern double  Rbf_Center_Activation();


/*-------------- Training Radial Basis Function Network -------------*/

main ()
{

double Error;

int Curr_It,          /* Loop variable for number of iterations    */
    np;               /* Index variable for number of patterns     */

FILE *Error_fp;

Read_Input_Patterns();

Read_Rbf_Centers();

if (NORMALIZE) {
    Normalize_Input_Patterns();
    Normalize_Rbf_Centers();
}

Read_Target_Patterns();

if (USE_KFM_COV) {
    Read_Rbfc_Cov_Diag();
}
else if (USE_COMP_COV) {
    Compute_Rbfc_Cov_Diag();
}

Compute_Rbfc_Activation_Matrix();

if (WT_READ) {
    Read_Rbfn_Weights();
    Read_Previous_Iteration();
}
else {
    Initialize_Rbfn_Weights();
    PITER = 0;
}

Initialize_Rbfn_Outputs();
Initialize_Backpropagation_Error();
Error_fp = Open_Error_File_To_Write();

Curr_It=0;
```

344

```
do {
    ++Curr_It;
    Error = 0.0;
    for(np=0; np< NPAT; ++np) {
        Compute_Rbfn_Output_Activation(np);
        Compute_Error_At_Output_Layer(np, &Error);
        Compute_Output_Layer_Error_Gradient(np);
        Modify_Rbfn_Weights(np, GAIN);
    }
    Write_Error(Curr_It, DIS_FRQ, Error, Error_fp);
} while (Curr_It< NIT);

fclose(Error_fp);

Write_Cpu_Time("rbft.c","cpu_rbft.out");

Write_Rbfn_Weights();

Write_Current_Iteration();
}
```

345

Appendix P

CODE FOR FLANK WEAR ESTIMATION
USING RADIAL BASIS FUNCTION NETWORK

```c
#include "rbf.h"
#include "bp.h"


#define NPAT        168         /* Number of training patterns   */
#define INODES      64          /* Number of input nodes         */
#define ONODES      1           /* Number of nodes in output layer */
#define OSCALE      50.0        /* Scale factor for network output */
#define SLOP        1.0         /* It defines slope of sigmoidal fn. */
#define NORMALIZE   1           /* Set 1, if inputs be normalized */
#define USE_KFM_COV 0           /* Set 1, if kfm cov values be used */
#define USE_COMP_COV 1          /* Set 1, if computed cov be used */
#define COV_RADIUS  0.5         /* used if,USE_KFM_COV=USE_COMP_COV=0 */
#define SATURATION  100.0       /* Saturation limit of net input */


#define RAND_MAX    2147483648.0
#define MEAN        0.0         /* Mean of initial random weights */
#define STD_DEV     0.5         /* SDV of initial random weights */


int NRBFC;                      /* Number of Rbf centers         */
int NNRBFC;                     /* Number of Rbf centers plus 1  */
double **Input_Pattern;
double **Target_Pattern;
double **Rbf_Center;
double **Rbfc_Activation;
double **Rbfc_Cov_Diag;
double **Rbfn_Weights;
double *Rbfn_Outputs;
double *Wear;


extern void    Initialize_Rbfn_Weights();
extern void    Read_Rbfn_Weights();
extern void    Initialize_Rbfn_Outputs();
extern void    Read_Rbf_Centers();
extern void    Normalize_Input_Patterns();
extern void    Normalize_Rbf_Centers();
extern void    Compute_Rbfc_Activation_Matrix();
extern double  Rbf_Center_Activation();
extern void    Read_Rbfc_Cov_Diag();
extern void    Compute_Rbfc_Cov_Diag();
extern void    Compute_Rbfn_Output_Activation();
extern FILE    *Open_Wear_Estimation_File_To_Write();
extern void    Write_Asci_Wear_Estimates();
```

346

```
/*---------- Estimation using Radial Basis Function Network ---------*/

main ()
{

int    np;              /* Index variable for number of patterns      */

FILE *Est_fp;

Read_Input_Patterns();

Read_Rbf_Centers();

if (NORMALIZE) {
    Normalize_Input_Patterns();
    Normalize_Rbf_Centers();
}

Read_Target_Patterns();

if (USE_KFM_COV) {
    Read_Rbfc_Cov_Diag();
}
else if (USE_COMP_COV) {
    Compute_Rbfc_Cov_Diag();
}

Compute_Rbfc_Activation_Matrix();

Read_Rbfn_Weights();

Initialize_Rbfn_Outputs();

Est_fp = Open_Wear_Estimation_File_To_Write();

for(np=0; np< NPAT; ++np) {
    Compute_Rbfn_Output_Activation(np);
    Write_Asci_Wear_Estimates(np, Est_fp);
}

fclose(Est_fp);
Write_Cpu_Time("rbfe.c","cpu_rbfe.out");
}
```

347

Appendix Q

CODE FOR TRAINING RECURRENT NEURAL NETWORK

```c
#include "gbp.h"


#define NPAT         168         /* Number of training patterns      */
#define INODES       3           /* Number of input nodes            */
#define ONODES       1           /* Number of nodes in output layer  */
#define HNODES       10          /* Number of nodes in hidden layer  */
#define NIT          5000        /* Number of iterations             */
#define OSCALE       40.0        /* Scale factor for network output  */
#define SLOP         0.25        /* It defines slope of sigmoidal fn. */
#define DIS_FRQ      10          /* Frequency of writing Error       */

#define WT_READ      0           /* Set 1, if wts  read from a file  */
#define NORMALIZE    0           /* Set 1, if inputs be normalized   */
#define GAIN         0.90        /* Gain for modifying wts and thrsh. */
#define ALPHA        0.25        /* Gain for modifying wts and thrsh. */
#define RUMELHART    0           /* Set 1,if Rumelhart learn. be used */
#define SEJNOWSKI    0           /* Set 1,if Sejnowski learn. be used */

#define SATURATION   100.0       /* Saturation limit of net input    */
#define MEAN         0.0         /* Mean of initial random weights   */
#define STD_DEV      0.5         /* SDV of initial random weights    */
#define RAND_MAX     2147483648.0

double **Input_Pattern;
double **Target_Pattern;
double **IH_Wts;
double **HO_Wts;
double **IH_LWts;
double **HO_LWts;
double *Onode_Thr;
double *Hnode_Thr;
double *Onode_LThr;
double *Hnode_LThr;
double *Onode_Act;
double *Hnode_Act;
double *Onode_Bpe;
double *Hnode_Bpe;
double *Wear;

int  PITER;

/*------------------ Training Recurrent Neural Network ---------------*/

main ()
```

348

```
{

double Error;

int curr_it,          /* Loop variable for number of iterations      */
    np;               /* Index variable for number of patterns       */

FILE *Error_fp;

Read_Input_Patterns();
if (NORMALIZE) {
    Normalize_Inputs_Patterns();
}
Read_Target_Patterns();
if (WT_READ) {
    Read_Network_Weights();
    Read_Node_Thresholds();
    Read_Previous_Iteration();
}
else {
    Initialize_Network_Weights();
    Initialize_Node_Thresholds();
    Initialize_Network_Lag_Weights();
    Initialize_Node_Lag_Thresholds();
    PITER = 0;
}
Initialize_Network_Activations();
Initialize_Backpropagation_Error();
Error_fp = Open_Error_File_To_Write();
curr_it=0;
do {
    ++curr_it;
    Error = 0.0;
    for(np=0; np< NPAT; ++np) {
        Compute_Hidden_Layer_Activation(np);
        Compute_Output_Layer_Activation();

        Compute_Error_At_Output_Layer(np, &Error);

        Compute_Output_Layer_Error_Gradient(np);
        Compute_Hidden_Layer_Error_Gradient();

        Modify_Input_Hidden_Layer_Weights(np, GAIN, ALPHA);
        Modify_Hidden_Output_Layer_Weights(GAIN, ALPHA);

        Modify_Hidden_Layer_Thresholds(GAIN, ALPHA);
        Modify_Output_Layer_Thresholds(GAIN, ALPHA);
    }
    Write_Error(curr_it, DIS_FRQ, Error, Error_fp);
} while (curr_it< NIT);
fclose(Error_fp);
```

349

```
Write_Cpu_Time("gbpt.c","cpu_gbpt.out");
Write_Network_Weights();
Write_Node_Thresholds();
Write_Current_Iteration();
}
```

350

CODE FOR FLANK WEAR ESTIMATION
USING RECURRENT NEURAL NETWORK

```c
#include "gbp.h"

#define NPAT        95      /* Number of training patterns       */
#define INODES      3       /* Number of input nodes             */
#define ONODES      1       /* Number of nodes in output layer   */
#define HNODES      10      /* Number of nodes in hidden layer   */
#define GAIN        0.90    /* Gain for modifying wts and thrsh. */
#define OSCALE      40.0    /* Scale factor for network output   */
#define SLOP        0.25    /* It defines slope of sigmoidal fn. */
#define NORMALIZE   0       /* Set 1, if inputs be normalized    */
#define SATURATION  100.0   /* Saturation limit of net input     */

#define INCREMENT   0.004   /* Saturation limit of net input     */

#define RAND_MAX    2147483648.0

double **Input_Pattern;
double **Target_Pattern;
double **IH_Wts;
double **HO_Wts;
double *Onode_Thr;
double *Hnode_Thr;
double *Onode_Act;
double *Hnode_Act;
double *Wear;

/*------------ Estimation using Recurrent Neural Network -----------*/

main ()
{

double Error;

int curr_it,        /* Loop variable for number of iterations    */
    np,             /* Index variable for number of patterns     */
    i;

FILE *Wear_fp;

Read_Input_Patterns();
Read_Target_Patterns();
if (NORMALIZE) {
    Normalize_Inputs_Patterns();
}
```

```
Read_Network_Weights();
Read_Node_Thresholds();
Initialize_Network_Activations();
Wear_fp = Open_Wear_Estimation_File_To_Write();

for(np=0; np< NPAT; ++np) {

    if(Input_Pattern[np][0] == 0.0) {
        Onode_Act[0] = 0.0;
    }
    else {
        Input_Pattern[np][0] = Onode_Act[0];
    }

    for(i=1; i< INODES; i++) {
        if(Input_Pattern[np][i] < Onode_Act[0]) {
            Input_Pattern[np][i] = Onode_Act[0];
        }
        else if(Input_Pattern[np][i] > (Onode_Act[0]+INCREMENT)) {
            Input_Pattern[np][i] = Onode_Act[0]+INCREMENT;
        }
    }

    Compute_Hidden_Layer_Activation(np);
    Compute_Output_Layer_Activation();
    Write_Asci_Wear_Estimates(np, Wear_fp);
}

fclose(Wear_fp);
Write_Cpu_Time("gbpe.c","cpu_gbpe.out");
}
```

VITA

EXPERTISE SUMMARY
Neural networks and knowledge based systems in design and manufacturing, process monitoring and diagnosis, intelligent sensors/sensor integration, pattern classification and group technology, metal cutting theory, time series analysis, and wavelet transforms.

CURRENT POSITION
Assistant Professor of Industrial Engineering and Information Systems, Northeastern University, Boston, MA, 02115.

EDUCATION

Ph.D. in Industrial Engineering
- The Pennsylvania State University, University Park, PA. Dec. 1994.
- Dissertation: "On-Line Tool Wear Estimation in Turning Through Sensor Data Fusion and Neural Networks."
- Co-Advisors: **Dr. Soundar R. T. Kumara, Dr. Paul H. Cohen**

M.S. in Industrial Engineering
- The Pennsylvania State University, University Park, PA. June 1990.
- Thesis: "An Investigation into the Application of Neural Networks for Component Design Data Retrieval."
- Co-Advisors: **Dr. Soundar R. T. Kumara, Dr. Victor E. Sanvido**

B.S. in Chemical Engineering
- Sri Venkateswara University, Tirupati, AP, India., June 1983.
- Project: "Manufacture of Furfural from Sugar Cane Waste."

RESEARCH INTERESTS
- Neural networks and knowledge based systems in design and manufacturing
- Distributed and cooperative artificial intelligence applied to systems integration
- Monitoring, diagnosis and control of machining processes
- Intelligent sensors/sensor integration
- Uncertainty in intelligent systems
- Design theory
- Wavelet transforms

TEACHING INTERESTS
- Neural networks and knowledge based systems in design and manufacturing
- Monitoring and diagnosis of machining processes
- Uncertainty in intelligent systems

AWARDS
- Won the **First Prize** in the Fourth Annual Graduate Research Exhibition (1989) for "Optical Neural Networks and Their Applications in Manufacturing Systems," at The Penn State University, University Park, PA.

- Won the **First Prize** in the National Student's Design Competition (1982) for "Manufacture of Formaldehyde," conducted by Indian Council for Science Museums, Bangalore, India.

PUBLICATIONS
1 book chapter, 6 journal papers, and 7 conference papers.

EXPERIENCE
Worked on several projects sponsored by NSF, industry, government organization, national laboratory, and the College of Engineering and the Department of Industrial Engineering at the Penn State University.