DEVELOPMENT OF A WEAR MONITORING SYSTEM FOR TURNING TOOLS USING ARTIFICIAL INTELLIGENCE

CORNELIUS SCHEFFER

Submitted in partial fulfilment of the requirements for the degree *Philosophiae Doctor* (Ph.D.) in the Faculty of Engineering, Built-Environment and Information Technology, University of Pretoria, Pretoria, South Africa.

2002

Supervisor: Prof. P. S. Heyns
Co-supervisor: Prof. Z. Katz

Development of a wear monitoring system for turning tools using artificial intelligence

Cornelius Scheffer
Supervisor: Prof. P. S. Heyns
Co-supervisor: Prof. Z. Katz
Department: Mechanical and Aeronautical Engineering
Degree: Ph.D.

SUMMARY

Tool wear during metal cutting operations is a continuous source of economic loss for the manufacturing industry. Complete tool failure after a certain degree of wear has occurred, can also have catastrophic consquences. One problem with monitoring tool wear is that the rate of wear and its geometric growth is always unique. This in effect means that a statistical approach for optimising the use of tool inserts cannot be realised, because too much losses will still occur together with the possibility of catastrophic failure. An alternative approach to achieve optimal tool use and prevent costly failures is on-line sensor-based monitoring of the tool inserts. In this research, a method is proposed for cost-effective on-line monitoring for turning tool inserts, based on Artificial Intelligence (AI) modelling, with the focus on turning operations.

To establish Tool Condition Monitoring (TCM) methods in industry, a generic method is required that can be applied to different types of operations. In this study, hard turning and interrupted cutting of Aluminium are investigated. A method is proposed that can effectively monitor the wear of the tool inserts used in these operations. It was shown that more than one wear mode can monitored in this way. Cutting conditions (*e.g.* speed, feed rate and depth of cut) can be included to ensure that the accuracy of the system is not affected if these conditions vary or change. A sensor-integrated tool holder was developed during the course of this work and it was shown how this tool can be used to reconstruct the cutting forces of a machining operation. A calibration procedure for the sensor-integrated tool was also developed.

The specific AI methods used in this research is Neural Networks (NNs). It was shown that using a novel fomulation of NNs, accurate monitoring can be achieved under shop floor conditions. This is achieved by training a combination of static and dynamic NNs. The new method is compared with other formulations and methods for further improvement are also investigated. Furthermore, an innovative training algorithm for on-line training of the NNs is also presented, after investigating many conventional and new optimisation algorithms. Achieving reliable and accurate wear monitoring under shop floor conditions is very significant since this has never been achieved to satisfaction up to date. In this case, using cost-effective custom developed hardware, advanced signal processing techniques and a novel formulation of NNs, the degree and rate of wear can be predicted at any given time. Comprehensive reviews of metal cutting, tool wear, modelling techniques, TCM and commercial TCM systems form part of this study as relevant background information.

Keywords: *condition monitoring, wear, neural networks, machine tool, vibration, machining, lathe, artificial intelligence.*

Ontwikkeling van 'n slytasiemoniteringstelsel vir draaibeitels deur gebruik te maak van kunsmatige intelligensie

Cornelius Scheffer
Promotor: Prof. P. S. Heyns
Mede-promotor: Prof. Z. Katz
Departement: Meganiese en Lugvaartkundige Ingenieurswese
Graad: Ph.D.

OPSOMMING

Slytasie van beitels wat gebruik word vir metaal sny operasies is 'n voortuderende bron van ekonomiese verliese in die vervaardigingsbedryf. 'n Katastrofiese faling van die beitel as gevolg van slytasie het ook ernstige gevolge. Die feit dat die geometrie en die tempo van die slytasie altyd uniek is, maak 'n statistiese benadering om die gebruik van die beitels te optimeer problematies. So 'n metode sal nie die moontlikheid van 'n katastrofiese faling uitskakel nie en sal steeds ekonomiese verliese meebring. 'n Alternatiewe metode om optimale gebruik van snybeitels aan te help en falings te voorkom is sensorgebaseerde monitering. Hierdie navorsing stel 'n koste-effektiewe metode voor vir intydse monitering van draaibeitels deur die gebruik van kunsmatige intelligensie. Om Masjiengereedskap ToestandsMonitering (MTM) metodes in die bedryf te vestig, word 'n generiese metode benodig wat op verskillende operasies toegepas kan word. In hierdie werk is die draai van verharde staal en die onderbroke draai van Aluminium ondersoek. 'n Metode word voorgestel wat die slytasie van beitels gedurende hierdie operasies betroubaar kan monitor. Daar word ook aangetoon hoe meer as een slytasiemodus op 'n beitel gemonitor kan word. Die masjineringsparameters kan ook in ag geneem word om te verseker dat die stelsel nie deur snytoestand variases beïnvloed word nie.

'n Sensor-geïntegreerde beitelhouer is ook ontwikkel en daar word aangetoon hoe die kragkomponente van 'n draai operasie daarmee bepaal kan word. Kalibrasieprosedures vir die sensor-geïntegreerde beitelhouer is ook ontwikkel. Die spesifieke kunsmatige intelligensie metode wat in hierdie navorsing aangewend word, is Neurale Netwerke (NNe). Daar word aangetoon hoe akkurate monitering van beitelslytasie op die fabrieksvloer moontlik is deur middel van 'n unieke formulering van NNe. 'n Kombinasie van statiese en dinamise NNe word opgelei om die slytasie intyds te voorspel. Die nuwe metode word vergelyk met ander formulerings, en metodes vir verdere verbetering word ook ondersoek. 'n Innoverende opleidingsalgoritme vir intydse opleiding van die dinamiese NNe word ook voorgestel na ondersoek van verskeie konvensionele en nuwe optimeringsalgoritmes. Die feit dat betroubare en akkurate slytasiemonitering op die fabrieksvloer berwerkstellig is, is 'n belangrike bydrae van hierdie werk omdat so iets nie voorheen geïmplementeer is nie. In hierdie werk is koste-effektiewe hardeware ontwikkel en gebruik saam met gevorderde seinprosessering en 'n unieke NN formulering om die waarde en tempo van die slytasie op engige gegewe tydstip te voorspel. Omvattende studies in metaalsny, beitelslytasie, modelleringstegnieke, MTM en kommersiële MTM stelsels word as relevante agtergrondmateriaal vir hierdie navorsing ingesluit.

Sleutelwoorde: *toestandsmonitering, slytasie, neurale netwerke, masjiengereedskap, vibrasie, masjinering, draaibank, kunsmatige intelligensie.*

ACKNOWLEDGEMENTS

The author wishes to extend his gratitude towards the following people for their valuable contribution to this research:

- Prof. Stephan Heyns at the University of Pretoria (UP) for his advice, support and encouragement. His personal interest in the work was the driving force for a successful research collaboration and it was an honour to work closely with such a dedicated researcher.
- Prof. Z. Katz from the Rand Afrikaans University (RAU) in Johannesburg for his special interest in the work, his advice, support and technical supervision that contributed to the success of the project.
- Mr. Hubert Kratz from the Werkzeugmaschinenlabor (WZL) at RWTH-Aachen, Germany for his all his inputs and assistance during the various visits to Aachen.
- Friends, staff and colleagues at the UP including Marietjie Calder, Frans Windell, Francois du Plooy, Corné Stander, Willie Smit, Derik van Staden, Bernhard Meister, Leon Staphorst, Jan Gutter, Prof. Albert Groenwold & Prof. Jan Snyman for their encouragement support, help and advice.
- Dr. D.E. Dimla from the DeMontfort University in Leicester for his interest in the work and advice during visits to Leicester, England.
- Dr. Bernhard Sick from the University of Passau, Germany for all his inputs through e-mail and also for a very fruitful visit to Passau.
- Various other colleagues and friends at WZL in the group of Prof. Klocke including Willi Weiss, Andrey da Silva, Ludwig Minten, Drazen Veselovac, Karsten Vormann, Sebatian Mader, Janli Ji, Martin Reuber & Rene Schmitz.
- Prof. Jun Ni from the University of Michigan, Ann Arbor, Michigan, U.S.A. for a very interesting visit to his lab and his technical inputs.
- Prof. I.S. Jawahir for his special personal interest in the research and technical inputs, and also for a very fruitful two-day visit to his labs in Lexington, Kentucky, U.S.A.
- Prof. R.E. DeVor and Prof. S. Kapoor from the University of Illinois at Urbana-Champaign, Illinois, U.S.A. for their inputs and advice on mechanistic modelling.
- Mr. Bill Kline from Montronix Inc., Ann Arbor, Michigan, U.S.A. for the very interesting visit and practical inputs.
- Mr. Mike Yaremchuck from Artis Systems in Detroit, Michigan, U.S.A. for his advice and very interesting visits to the Ford manufacturing and development facilities in Detroit.
- Mr. Mario Gioia from UV Water Systems in Pretoria for his encouragement and understanding.
- Mr. Ty Dawson and Prof. Tom Kurfess from the Georgia Institute of Technology, Atlanta, Georgia, U.S.A. for their assistance and various discussions in solving the modelling problem in hard turning.
- Gerrie Pieterse, Chris van der Merwe, Justin Moyes, Karl Wenk and various others at KOLBENCO (PTY) LTD for their assistance in setting up a TCMS in their plant.
- Mr. Johan van Staden from Inher SA for arranging a visit to De Beers industrial diamonds and

also Dr. Martin Suckling for a very interesting visit to their research facilities at De Beers.

- Prof. Ahmad Ghasempoor from the Ryerson Polytechnic University, Ontario, Canada, for his inputs on the formulation of NNs.
- Mr. V. Sansoni and Mr. K. F. Von Langermann at Tranconsa SA for their assistance in identifying an industrial need.
- Prof. Marco Santochi from the University of Pisa, Italy for advice on developing a sensor-integrated tool.
- Prof. Dong-Woo Cho from the Pohang University of Science and Technology, South Korea for his interesting inputs.
- Prof. John W. Sutherland from the Michigan Technological University, Houghton, Michigan, U.S.A. for his helpful advice.
- Prof. Nidal Abu-Zahra from the University of Wisconsin Milwaukee for his comments on the ultrasonic approach to TCM.

The author would like to thank the following institutions for their financial contributions:

- National Research Foundation of South Africa
- University of Pretoria
- Mellon Foundation
- Deutscher Akademischer Austauschdienst (DAAD)

On a personal note, I would also like to thank my friends and family for their encouragement, and the most important person, my wife Zonia, for her everlasting understanding and support in being married to an eternal student: Her encouragement played a vital role in completing this huge project that was a central part of our life for so many years.

TABLE OF CONTENTS

NOMENCLATURE

UPPERCASE

A	constant (SOM), approximations (wavelet analysis), pulse magnitude (strain gauges)
$\boldsymbol{A}$	calibration matrix
A_c	area of chip load (mechanistic model)
B	constant (SOM), pulse magnitude (strain gauges)
BL	length of groove backwall wear
BW	width of groove backwall wear
C	constant (Taylor equations)
C_o	wave velocity (strain gauges)
C_x	covariance
D	workpiece diameter , details (wavelet analysis), NN initialisation space
E	modulus of elasticity , signal entropy (wavelet analysis)
F_{cw}	cutting force due to tool flank wear (mechanistic model)
F_f	frictional cutting force (mechanistic model)
F_f	feed force (orthogonal and oblique cutting)
F_r	radial force (oblique cutting)
F_n	normal cutting force (mechanistic model)
F_s	sampling rate
F_t	tangential force (orthogonal and oblique cutting)
F_{tw}	thrust force due to tool flank wear (mechanistic model)
Fx,y,z	directional static forces (turning)
Fxd	natural frequency of tool holder feature (hard turning)
G_L	gauge length (strain gauges)
I	second moment of area
$\boldsymbol{I}$	identity matrix
K	kurtosis
K_D	depth of the crater wear (hard turning)
K_f	specific feed pressure constant (mechanistic models)
K_f	frictional cutting pressure (mechanistic model)
K_L	approximate length of the crater wear (hard turning)
K_M	crater wear centre from edge
K_n	normal cutting pressure (mechanistic model)
K_t	specific cutting pressure constant (mechanistic models)
K_T	crater wear depth
KT	depth of groove backwall wear
K_V	volume of crater wear (hard turning)

K_{Vfr}	full range of crater wear volume
K_{Vmd}	maximum deviation from true crater wear volume
K_{Vrms}	rms deviation from true crater wear volume
K_{VTCMS}	crater wear volume predicted by the monitoring system
K_{Vtrue}	measured crater wear volume
K_W	width of the crater wear (hard turning)
L	length of cut along shaft (turning), surface profile sample length, gauge length (strain
M	number of files to compress
N	nose wear
NL_1	notch wear length on main cutting edge
NL_2	notch wear length on secondary cutting edge
NW_1	notch wear width on main cutting edge
NW_2	notch wear width on secondary cutting edge
P_L	pulse length (strain gauges)
R_a	roughness average
P_{xx}	power spectral density of X
P_{xy}	cross spectral density of X and Y
R_t	theoretical surface roughness (turning)
R_z	rms surface roughness
S	signal, skewness
SD	depth of secondary face wear
SW	width of secondary face wear
T	tool life (taylor equation), pulse period (strain gauges), time interval, sampling time
T_R	reference tool-life
V	cutting speed (orthogonal and oblique cutting)
V_c	chip speed (orthogonal and oblique cutting)
V_0	Wheatstone bridge excitation voltage
V_R	reference cutting speed
VB	flank wear
$VB(i)$	on-line wear estimation at step i
VB_A	area of flank wear (hard turning)
VB_{Afr}	full range of flank wear area
VB_{Amd}	maximum deviation from the true flank wear area
VB_{ATCMS}	flank wear area predicted by the monitoring system
VB_{Atrue}	measured flank wear area
VB_{Arms}	rms deviation from the true flank wear area
VB_{avg}	average flank wear (hard turning)
VB_L	approximate length of the flank wear (hard turning)

VB_{max}	maximum flank wear
VB_{notch}	notch wear
VB_p	width of the plastic flow region (mechanistic model)
W_C	coating effect factor
W_g	chip-groove effect factor
X_{rms}	rms of X

LOWERCASE

a	neuron output, wavelet scale function
$a_1, a_2 \ldots a_p$	AR parameters
b	chip width (orthogonal and oblique cutting), length of flank wear, neuron bias
$b_1, b_2 \ldots b_q$	MA parameters
c	index of BMU, trigger channel
c_0, c_1	linear coefficients (mechanistic model)
doc	depth of cut
e_i	eigenvectors (PCA)
f	feed rate, neuron activation function
f	frequency of interest (strain gauges)
fh	upper frequency boundary
fl	lower frequency boundary
f_o	natural frequency (dynamometer)
f_r	resonant frequency (dynamometer)
h	depth of cut (orthogonal and oblique cutting)
h	height of surface profile
$h_{c(x),i}$	neighbourhood function (SOM)
i	cutting edge inclination angle (oblique cutting), neuron index (SOM), monitoring outer step, trigger threshold
i_e	effective inclination angle (mechanistic model)
k_{ij}	calibration matrix coefficients
l	tool overhang length
lm	cutting length per minute (turning)
m	input vectors (SOM)
$\mathbf{m}$	weight vector (SOM)
n	rotational speed of spindle (turning), constant (Taylor equation), dimension of input vectors (SOM), number of weight and bias values in the dynamic NN, total number of tool wear measurements, number of trigger events, index of ARMA models, mode of vibration
p	neuron input, order of AR model
q	order of MA model

r	tool nose radius
r_i	locality index (SOM)
r_n	tool nose radius (mechanistic model)
s	signal
t	index of learning step (SOM)
t_c	is the uncut chip thickness(mechanistic model)
tol	convergence tolerance
u	sequence for MA model
v	voltage (calibration)
w	neuron weight, width of cut (mechanistic model)
x	distance from tool tip (mechanistic model)
$\boldsymbol{x}$	weight vector of dynamic NN, sample vector (SOM)
Δw	width of element (mechanistic model)
$\bar{x}$	mean value

GREEK SYMBOLS

$\alpha(t)$	learning rate function (SOM)
α_{en}	effective normal rake angle (mechanistic model)
α_n	normal rake angle (mechanistic model)
α_r	rake angle (oblique cutting)
β	clearance angle
β_n	end condition parameter
ε_{out}	reported peak strain (strain gauges)
ε_R	peak strain (strain gauges)
ϕ_c	shear angle (oblique cutting)
γ^2	coherence function
γ_{eL}	effective lead angle (mechanistic model)
η	average accuracy with respect to the full range of wear, chip-flow angle (oblique cutting)
η_c	chip-flow angle (mechanistic model)
λ_i	eigenvalues (PCA)
μ_x	mean of vector population (PCA)
ρ	mass density of tool holder material, correlation coefficient
σ	minimum accuracy with respect to the full range of wear
σ^2	variance
σ_{eff}	effective stress (mechanistic model)
σ_o	tool tip normal stress (mechanistic model)

σ_w	normal tool flank stress (mechanistic model)
τ_o	tool tip shear stress (mechanistic model)
τ_w	shearing tool flank stress (mechanistic model)
ω	frequency
$\Delta\xi$	sections on tool (mechanistic model)
ψ	wavelet function
ψ_s^2	frequency band energy from PSD function

ABBREVIATIONS

3-D	Three Dimensional
A/D	Analogue to Digital
AC	Adaptive Control
ACC	Adaptive Control Constraint
ACO	Adaptive Control Optimisation
ADELINE	Adaptive Linear Neuron Network
AE	Acoustic Emission
AI	Artificial Intelligence
ANOVA	Analysis of Variance
AR	Auto-Regressive
ARD	Automatic Relevance Determination
ARMA	Auto-Regressive Moving Average
ART	Adaptive Resonance Theory
ASPS	Automated Sensors and Signal Processing Selection
BHN	Brinell Hardness Number
BMU	Best-Matching Unit (SOM)
BW	Bandwidth
CAD	Computer Aided Design
CBN	Cubic Boron Nitride
CF	Crest Factor
CNC	Computer Numerical Control
CWT	Continuous Wavelet Transform
DN	Dynamic Network
DWT	Discrete Wavelet Transform
ECT	Equivalent Chip Thickness
ETOP	Energy Trajectory Optimisation Program
EUR	Euro
EWMA	Exponentially Weighted Moving Average
FEM	Finite Element Method
FF	Feedforward
FFT	Fast Fourier Transform

FRF	Frequency Response Function
GA	Genetic Algorithm
GEOMC	Global Economical Optimal Machining Conditions
IDD	Independent Identically Distributed
IDWT	Inverse Discrete Wavelet Transform
KBES	Knowledge-Based Expert Systems
KFM	Kohonen Feature Map
LFOP	Leap-Frog Optimisation Program
MLP	Multilayer Perceptron
MRR	Metal Removal Rate
NC	Numerical Control
NFS-S	Supervised Neuro-Fuzzy System
NN	Neural Network
OMC	Optimal Machining Conditions
OR	Operations Research
PC	Personal Computer
PCA	Principal Component Analysis
PLS	Partial Least Squares
PSD	Power Spectral Density
PSOA	Particle Swarming Optimisation Algorithm
RAMV	Ratio of Absolute Mean Value
RAN	Resource Allocation Network
RBF	Radial Basis Function
RMC	Recommended Machining Conditions
RNN	Recurrent Neural Network
SEM	Scanning Electron Microscope
SeTAC	Sequoia Triaxial Acceleration Computer
SLP	Single Layer Perceptron
SN	Static Network
SOF	Statistical Overlap Factor
SOM	Self-Organising Map
SPC	Statistical Process Control
SQSD	Spherical-Quadratic Steepest Descent
SS	Stainless Steel
SURE	Stein's Unbiased Risk Estimate
TCM	Tool Condition Monitoring
TCMS	Tool Condition Monitoring System
TDL	Tapped Delay Line
TDNN	Time Delay Neural Network
TMC	Technical Machining Conditions

1. Introduction

1.1 Preface

In the modern age millions of products are manufactured on a daily basis by a variety of processes. Many products or components of products are manufactured from metal and metal alloys. One of the most basic methods to form bulk metal into a desired final shape is through the process of metal cutting, also referred to as machining. Metal cutting is essentially the removal of excess material from a workpiece by moving a working tool over the surface of the workpiece. Through this, a certain shape is attained together with a desired surface quality of the final product. Machining is usually the final step in the manufacturing process of a metal component, following other bulk deformation processes such as casting, forging and rolling. Conventional machining operations are turning, milling and drilling. A turning operation is pictured in Figure 1.1.

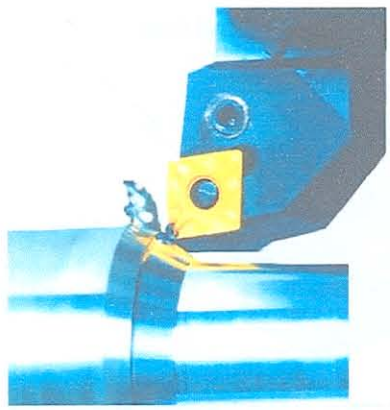


Figure 1.1: Turning operation [1]

Due to the increasing demands for faster and more accurate machining, Computer Numerically Controlled (CNC) machines are commonly used for the above-mentioned processes. These machines are often unmanned and components are moved to and from the machines with component feed devices. A modern CNC machine for turning operations (CNC lathe) is shown in Figure 1.2. Despite the high level of technology built into every aspect of machining, there is still one factor present that hampers the reliability and complete automation of the processes. This factor is the tool wear. Tool wear is the loss of material on the edges of the cutting tool. Although tool wear can be minimised by selecting proper machining conditions, it cannot be completely eliminated. Unfortunately, even a small quantity of tool wear may cause a defect in a machined component. Furthermore, secondary damage due to tool wear can be extreme and even catastrophic [2]. For this reason, many approaches to Tool Condition Monitoring (TCM) have been developed through the years. However, none of the methods developed up to date seem to fulfil the requirements for TCM in industry.

Figure 1.2: Modern CNC lathe

The aim of this thesis is to develop a system that can predict the severity of wear on the edge of turning tool inserts without a direct measurement on the tool tip. A typical turning tool holder with an insert is shown in Figure 1.3. A photo of a worn tool insert is shown in Figure 1.4.

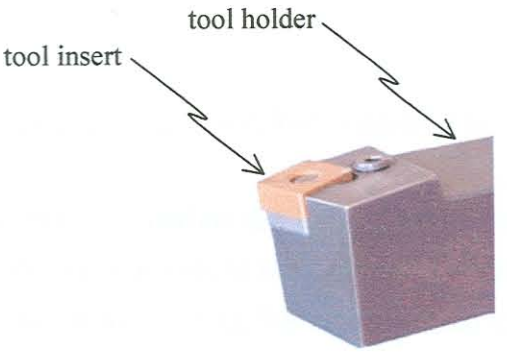


Figure 1.3: Turning tool holder with insert

In this thesis, indirect measurement methods are used in conjunction with Artificial Intelligence (AI) schemes to assist in an accurate estimation of the tool wear. Different turning operations are investigated to determine if the AI approach can be treated as a generic approach for TCM in turning. Due to the complexity of the TCM problem, many aspects of machining, signal measurement, signal analysis and AI modelling are covered in this text. The general theme of the thesis can best be described as a condition monitoring methodology with application in the field of machining.

Figure 1.4: Worn turning tool insert

1.2 Industrial need

The use of fast and accurate manufacturing equipment has gained more ground in recent years due to the high demands of the fast growing manufacturing industry. In order to justify the investment associated with the purchase of such equipment, it is necessary to achieve the maximum possible utilisation of each machine. Monitoring of the manufacturing process plays a very important role to avoid down time of the machine, or to prevent unwanted conditions such as chatter vibration, excessive tool wear or tool breakage. This is also very important in the unmanned machining environment, where the machine must be able to operate non-stop without an operator checking for errors.

Tool wear is one important factor that should be monitored if reliable machining is required. The manufacturing industry requires tool wear monitoring systems that are reliable, accurate and cost-effective. Systems that are currently available do not fulfil these requirements. Generally, wear monitoring systems is required for milling, drilling and turning operations. In the South African industry, the requirement is more focused on turning operations. Consequently it was decided to start the development of the monitoring system with turning operations. However, it is suggested that the proposed methodology can later be extended to more complex machining operations.

1.3 Economic aspects

Tool wear has considerable economic impacts. Therefore, intensive research are carried out in the following areas:

- development of sophisticated tool materials and tool coatings to minimise tool wear
- development of adaptive control strategies to minimise tool wear during machining
- development of Tool Condition Monitoring Systems (TCMS) to predict the tool life and optimise the use of machine tools
- optimisation strategies are used to optimise the machining parameters, mostly to maximise metal removal with minimum tool wear
- development of wear mechanism maps to optimise the use of tools and to compare different tool materials
- development of mathematical models (theoretical, numerical and empirical) to model cutting forces, tool wear, chip formation, surface finish *etc.*

The focus of this work is on TCM, which plays a significant role in the complete economic optimisation of production. The exact economic losses due to tool wear occurs due to scrapping of expensive parts, production down time and the non-optimal use of cutting inserts. Without TCM, a conservative approach is taken and the insert is recycled long before it should have been. This is done because the wear rate is very unpredictable. Sometimes the tool will wear quickly and sometimes slower, even when used with the same machining parameters. The economic impact of cutting tools on total production costs is reported in various studies, such as [3,4]. It is stated that a considerable amount of money can be saved through better use of cutting inserts themselves. Shop floor managers agree that a reliable TCMS will bring about significant cost savings and will thus justify its capital cost. Compared to the

overall cost of CNC machines, installing a TCMS will only cause a slight price increase. The possibility should also exist to retrofit old machines with a TCMS at low cost. Typical savings for a manufacturing plant with effective TCM could be more than 50% less scrap, tool inserts and down time. It should be mentioned here that production facilities could also be economically optimised through research in the other areas listed above, and that savings could be more than that achieved with TCM. Nevertheless, effective TCM should be part of the production process if a manufacturer wishes to stay competitive in the global economic environment.

The automotive industry is the principal user of TCMSs. A survey of more than 1000 TCMSs installed in industry is described in [5], and shows that the emphasis is mainly on turning and drilling, as depicted in Figure 1.5.

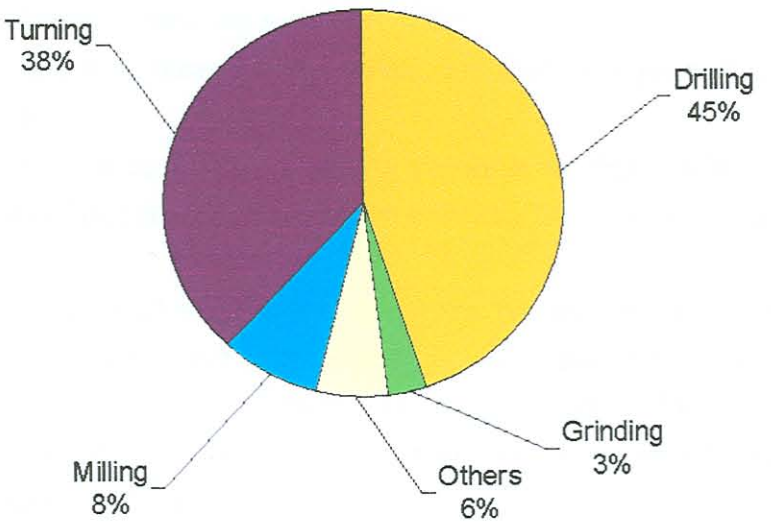


Figure 1.5: Survey of TCMSs by area of application [5]

1.4 Engineering aspects

A number of specialist engineering disciplines are combined in this work. The main focus areas are:

- Condition monitoring
- Artificial intelligence
- Signal processing
- Production / manufacturing
- Structural dynamics
- Data acquisition
- Mathematical optimisation

It will be shown that combining state-of-the-art technology from the different engineering disciplines is vital to achieve success in Tool Condition Monitoring (TCM). Many previous attempts at implementing TCM in industry failed due to a lack of knowledge in one of the key disciplines listed above.

1.5 Scope of research

1.5.1 Contribution

Despite the industrial need and many research papers a TCMS using AI has never successfully been implemented in a real production environment. The following list of quotes from recent literature suffice as proof that this has not yet been achieved [6]:

- "Although many unique characteristics of neural networks appeal to the researchers..., the physical realisation of a neural network monitoring system has not been seen in industries yet. The gap between academic enthusiasm for the neural network approach and industrial needs for a practical and reliable monitoring system tends to be caused by the simple adoption of existing neural network methodology without consideration of some critical issues in implementation of the tool-wear monitoring systems" – [7]
- "A reliable on-line wear measurement system does not exist yet and research in this area is continuing." – [8]
- "The process of metal cutting is a complex phenomenon that has been researched for many years but the aim of practical tool condition monitoring has yet to be achieved." – [9]

This list could easily be continued with remarks from other researchers active in the area of TCM. The proposed contribution of this research is to achieve practical tool wear monitoring using AI. This should be achieved with an AI methodology that is unique in terms of formulation and application. This research aims to make a significant contribution towards more reliable, accurate and cost-effective tool wear monitoring for turning.

1.5.2 Summary of scope

The scope of the research can be summarised as:

- To investigate the use of sensor signals to develop a wear monitoring system for turning tools.
- The wear monitoring system should utilise the advantages given by AI modelling.
- Sophisticated signal processing and AI methods must be investigated to improve the accuracy and reliability of the system.
- The proposed methods must be shown to apply to two different types of turning operations, and also be able to monitor more than one wear mode.
- An industrial implementation of the complete AI monitoring system must be achieved, and must be unique in terms of its formulation and application. It should be the first implementation of AI for TCM in a production environment.
- The research will provide significant new knowledge to the research community of how to approach the problem of TCM.
- The system must be a robust and cost-effective solution to TCM on a shop floor, providing a useful product to industry.

These are the main aspects of the research. There are also a number of detail factors that must be in-

vestigated, and these are divided into three subsections, listed below:

1.5.3 Signal measurement

- Investigate force, acoustic emission and acceleration measurements for TCM.
- Interpret the signals with respect to the particular machining operation.
- Identify the best signal collection approach for robust and cost-effective TCM.
- Install the selected sensor approach on a shop floor for TCM.
- Characterise the new sensor approach appropriately.
- Collect data under normal shop floor conditions with an automated data logger.
- Show that the proposed technique can be used for continuous wear estimation under shop floor conditions.

1.5.4 Signal processing

- Investigate various signal processing techniques with respect to the machining operation.
- Time, frequency, time-frequency as well as statistical analyses must be investigated for feasibility towards TCM.
- Conclude towards signal feature sensitivity and machining conditions.
- Compare wavelet analysis with digital filtering.
- Attempt to identify and filter the effect of disturbances from signals.
- Investigate and conclude towards the best feature selection and reduction methods for TCM.
- Suggest appropriate rules for feature selection for practical TCM.

1.5.5 Modelling and monitoring

- Investigate the feasibility of analytical, empirical and numerical modelling with respect to TCM.
- Show that an AI approach for TCM is the most advantageous for practical TCM.
- Conclude towards the feasibility of sensorless approaches (tool life equations).
- Suggest an AI approach for TCM that can estimate tool wear on-line using dynamic NNs.
- The AI method should be applicable to more than one turning operation, insensitive to noise, must be able to handle the effects of machining parameters, monitoring more than one wear mode and must be able to follow any geometrical development of tool wear.
- Compare the proposed method with other conventional NN paradigms.
- Attempt to improve the method and investigate the repeatability of the method.
- Investigate proper training algorithms for on-line NN training and implement the best algorithm.

1.5.6 Research steps

To achieve these objectives, the following studies / steps were undertaken:

- Review of machining (kinematics, processes *etc.*)
- Exhaustive review of commercially available TCMSs and related patents

- Exhaustive review of literature in the area of machining and TCM
- Various industrial visits to identify industrial and economic requirements
- Various academic visits to identify the state-of-the-art in relevant technology and identify limitations of previous work
- Review of methods for modelling machining kinematics and tool wear
- Review, design and development of custom hardware
- Planning and conduction of experiments and processing results
- Design, analysis and proposal of monitoring strategy
- Comparison with other approaches
- Proving and improving monitoring strategy
- Documentation

1.6 Structure

Taking the research objectives into account, two turning operations were investigated:

- Hard turning (in a research laboratory)
- Aluminium turning (on a real shop floor)

The reasons for these particular choices are discussed in Chapters 4 and 5 respectively. The hard turning experiments were conducted at the Laboratory of Machine Tools and Manufacture (WZL) at the Aachen University of Technology, Germany. The Aluminium turning experiments were conducted at Kolbenco Pty (Ltd) in Alrode, South Africa. In Chapter 6, there is a discussion with respect to the results of both sets of experiments. A review of relevant literature, and a short review of the mechanics of metal cutting and tool wear are documented in Chapters 2 and 3. The Conclusion and Future Outlook are included in Chapter 7. Many of the technical details are described in Appendices A – I.

2. Background to machining and tool wear

2.1 Introduction

In this chapter, the basics of metal cutting and tool failure are explored. Tool failure refers to the wear *and / or* the catastrophic breakage of tools. Selecting proper operating conditions can prevent catastrophic tool failure. With present technology, tool wear cannot be prevented. At some stage the tool will wear out and will require replacement. For the purpose of this study, the failure of single point turning tools are discussed, although the failure of tools used for processes like milling and drilling are similar. These tool types are called defined cutting edges because the exact geometry of the tool is known. A grinding wheel is an example of a non-defined cutting edge, because the geometry of the cutting edge was randomly generated. The mechanics of cutting with defined cutting edges are often similar. It is believed that many of the concepts proposed in this research can possibly be extended to other machining processes with defined cutting edges, even if these are not single point tools.

Machine tool failure is described in terms of failure mechanisms and modes. The failure mechanism is the underlying cause of the tool failure, whereas the failure modes are used to describe the nature or the appearance of the failure. The wear mechanisms and wear modes are covered in detail in this chapter. Some other relevant aspects like tool wear maps and mathematical models for tool life are also discussed as relevant background material.

2.2 Mechanics of the cutting process

2.2.1 Introduction

This section will cover an introduction to the mechanics of turning, starting with the basic principles of orthogonal and oblique cutting (refer to the textbook by Altintas [14] from which some information presented in this section was sourced). Although most common cutting operations are three-dimensional, the simple two-dimensional case of orthogonal cutting is useful to introduce the mechanics of metal cutting. The mechanics of more complex three-dimensional cutting operations are usually determined by applying a transformation model to orthogonal or oblique models.

2.2.2 Orthogonal cutting

A schematic representation of orthogonal cutting is depicted in Figure 2.1 (adopted from [14]). Orthogonal cutting is a process where metal is removed with a straight tool with the cutting edge perpendicular to the cutting velocity V. A metal chip with width b and depth h is sheared away from the workpiece with speed V_c. The cutting is assumed to be uniform and therefore it is modelled as a plane strain deformation process. In the case of orthogonal cutting the forces are exerted only in the direction of velocity and uncut chip thickness, called the tangential F_t and the feed force F_f.

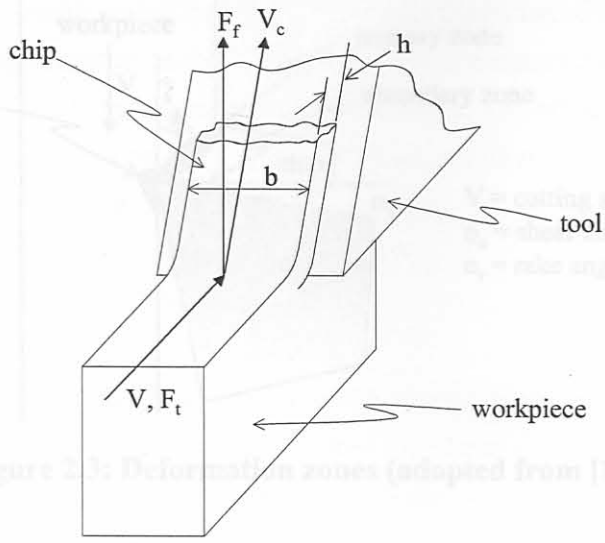


Figure 2.1: Orthogonal cutting geometry (adopted from [14])

2.2.3 Oblique cutting

In oblique cutting the cutting edge is orientated with an inclination angle and an additional third force, the radial force F_r, act in the radial direction. The geometry of oblique cutting is depicted in Figure 2.2 (adopted from [14]). During oblique cutting, the chip is sheared away from the workpiece with an angle η, called the chip-flow angle. The angle i between the workpiece and the cutting edge, is referred to as the cutting edge inclination angle.

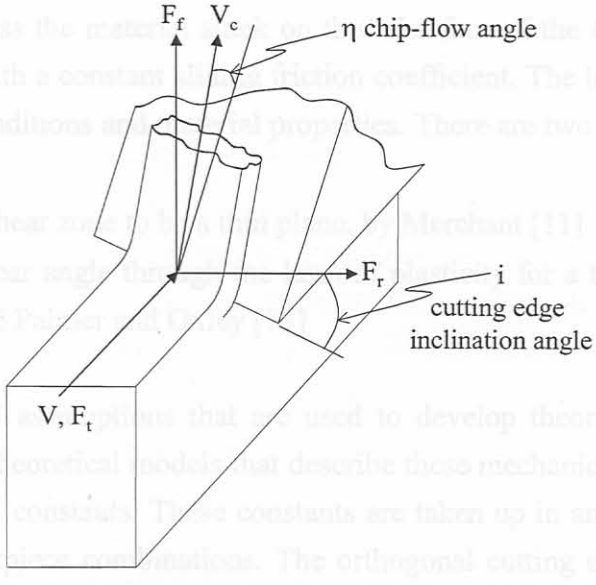


Figure 2.2: Oblique cutting geometry (adopted from [14])

There are three basic deformation zones formed during metal cutting, as shown in the cross-sectional view in Figure 2.3 (adopted from [14]). Stabler [10] found that the orthogonal angle is the same as the chip flow angle for orthogonal cutting (hence $\eta = i$). This is called Stabler's empirical chip flow rule and applies to many machining operations.

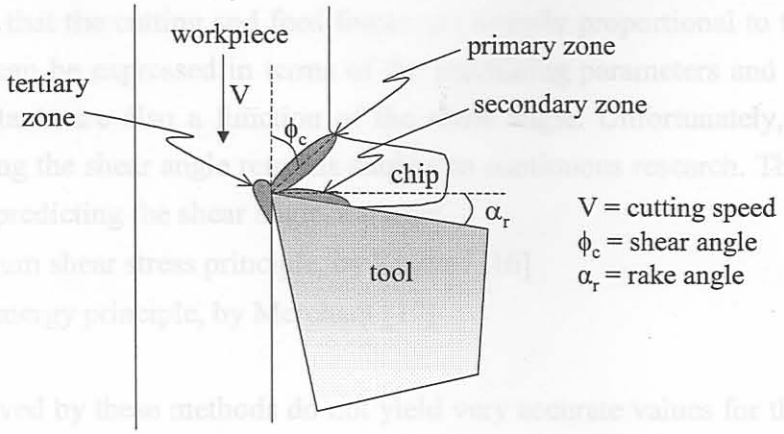


Figure 2.3: Deformation zones (adopted from [14])

When the tool tip penetrates the workpiece, workpiece material is sheared over the primary shear zone to form the chip. The sheared material partially deforms as the chip moves over the rake face of the tool in the secondary shear zone. The area where the flank face of the tool rubs along the newly machined surface is the tertiary shear zone. This is also referred to as the *friction zone* because the flank of the tool rubs against the newly machined surface. The chip initially sticks to the rake face of the tool in the *sticking zone*. A number of theoretical equations exist that describe the mechanics in the three deformation zones. See also for example the descriptions in reference [14].

The friction stress is approximately equal to the yield shear stress of the material at the sticking zone where the chip moves across the material stuck on the rake face of the tool. The chip will then stop sticking and start sliding with a constant sliding friction coefficient. The length of the contact zone depends on the machining conditions and material properties. There are two approaches for analysing the primary shear zone:

- Assuming the shear zone to be a thin plane, by Merchant [11]
- Predicting a shear angle through the laws of plasticity for a thick shear zone, by Lee and Shaffer [12] and Palmer and Oxley [13]

These are the fundamental assumptions that are used to develop theoretical models of the cutting forces. In many cases, the theoretical models that describe these mechanics still require experiments to determine certain empirical constants. These constants are taken up in an orthogonal cutting database for different tool and workpiece combinations. The orthogonal cutting database can then be used in extended models that apply to oblique machining without the need for further experiments. A description of these methods and equations is not within the scope of this text, but can be found in references such as [14,15]. The main problem with this approach is the infinite number of tool and workpiece combinations, and also the somewhat complex nature of the formulations. As a result, the methods are not often implemented to verify cutting force measurements and are of limited to use to industry.

Another approach is mechanistic modelling, which can assist to determine the 3-D cutting forces for any practical tool and workpiece combination. With this method, a specific cutting pressure constant (K_t) and a feed pressure constant (K_f) must be determined experimentally. The underlying assumption

of this approach is that the cutting and feed forces are directly proportional to these constants. Hence, the cutting forces can be expressed in terms of the machining parameters and the pressure constants. The pressure constants are also a function of the shear angle. Unfortunately, an accurate analytical method of predicting the shear angle remains subject to continuous research. There are two fundamental approaches for predicting the shear angle, namely:

- The maximum shear stress principle, by Krystof [16]
- Minimum energy principle, by Merchant [17]

The equations derived by these methods do not yield very accurate values for the shear angle, but provide a meaningful insight into the relationship between the shear angle, rake angle (refer Figure 2.3) and the friction coefficient of a tool and workpiece combination. The aim of tool design would be to keep the shear angle as small as possible to keep power consumption and cutting forces low. The rake angle must be at a maximum and a cutting fluid must be used to decrease the friction coefficient. By applying one of the assumptions stated above, the oblique cutting parameters can be solved and will provide a model to determine cutting forces.

There are also a number of empirical models, and the one proposed by Armarego is probably the most famous [14]. This approach assumes that the shear velocity is collinear with the shear force, and that the chip length ratio in oblique cutting is the same as for orthogonal cutting. If the Stabler rule is applied, the chip flow, shear and the normal friction angle can be used in an empirical or mechanistic model to solve cutting forces.

2.3 Turning

Turning is one of the oldest and simplest machining processes, and is used to machine cylindrical parts. During turning, a workpiece is clamped in a chuck that is fixed to a spindle. Long workpieces are held in the chuck and the centre of a tailstock. A single point tool is clamped on a tool post, and the tool post can move between the spindle and the tailstock. Conventional lathes have only one motor at a constant speed, and the speed is transmitted to the spindle and feed drive gearboxes with belts. The feed and speed can be changed with marked shift handles that are connected to the respective gearboxes. In modern CNC lathes the speed and feed can programmed numerically, because these have stepless computer controlled drives.

CNC lathes generally have a turret containing multiple tools, and the turret can often move along two axes. If the tool moves along the axis of the spindle, it will reduce the diameter of the workpiece and this is referred to as turning. If it moves perpendicular to the main axis, it will remove material along the flat face of the workpiece, in what is called a facing or parting operation. These are external turning operations. When metal is removed from the inside of a cylindrical workpiece, it is called internal turning or boring. With this combination of movements, a lathe can be used to machine many complex cylindrical parts, and can also be used to produce a screw thread. Often, roughing and finishing operations are performed to achieve a certain surface quality. A sketch of a typical turning operation is shown in Figure 2.4. The most important machining parameters are the:

- cutting speed [V] usually expressed in m/min
- feed [f] rate usually expressed in mm/rev
- depth of cut [doc] usually expressed in mm

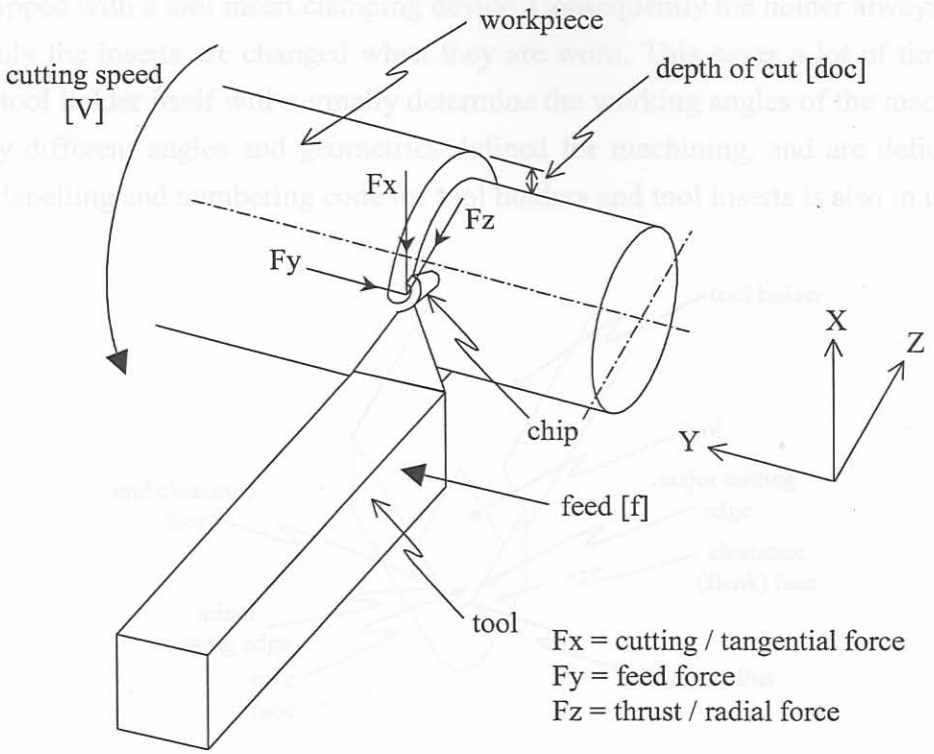


Figure 2.4: Diagrammatical turning operation

The cutting speed is defined by:

$$V = \pi D n \text{ [m/min]} \tag{2.1}$$

The feed rate is defined by:

$$f = \frac{lm}{n} \text{ [mm/rev]} \tag{2.2}$$

A total cutting time can also be determined, expressed as:

$$T = \frac{L}{n \times f} \text{ [min]} \tag{2.3}$$

where:

D = workpiece diameter [mm]
n = rotational speed of spindle [rpm]
lm = cutting length per minute [mm/min]
L = length of cut along shaft [mm]

The force response on the tool tip as a result of the turning operation is the three component cutting force, as shown in Figure 2.4. Turning is an oblique machining operation and hence these forces can be predicted by transforming the orthogonal cutting parameters to an oblique turning geometry using a conversion of the operational angles. Turning causes varying chip thickness and thus the angle transformation is applied separately to a region of uniform thickness and a region of varying thickness. The

mechanistic method discussed before is a frequently used to model the cutting forces for turning. It should be kept in mind that such a model will only apply to the static cutting forces.

Basic terminologies for turning tools are shown in Figure 2.5. Turning tools often consist of a standard tool holder equipped with a tool insert clamping device. Consequently the holder always remains in the machine and only the inserts are changed when they are worn. This saves a lot of time in setting up machines. The tool holder itself will normally determine the working angles of the machining process. There are many different angles and geometries defined for machining, and are defined by the ISO 3002/1 [18]. A labelling and numbering code for tool holders and tool inserts is also in use worldwide.

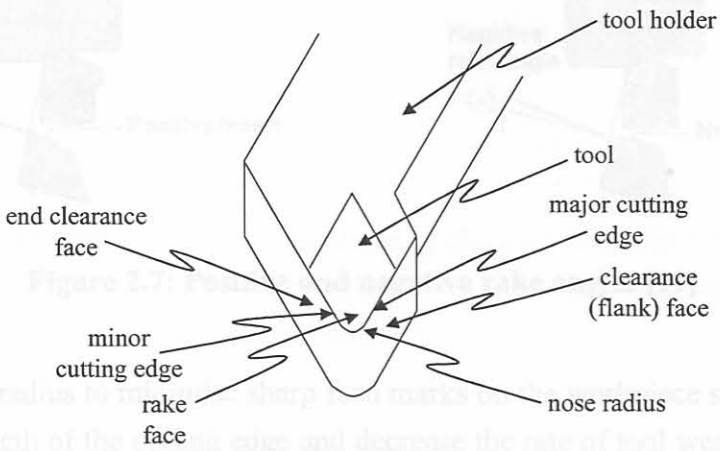


Figure 2.5: Turning tool terminologies

The basic geometries for a turning tool are depicted in Figure 2.6. The important parameters are the tool nose radius, rake and side cutting edge angles. It was noted that different terms are sometimes used for the same angles in the literature. In this document, the terms defined in Figure 2.6 will be used, which also corresponds to ISO 3002/1.

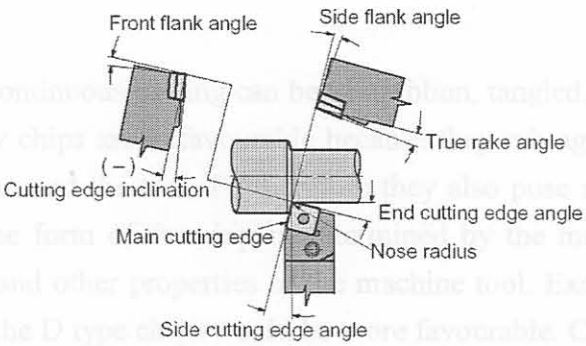


Figure 2.6: Basic angles for turning [19]

The ISO also separates between working angles and tool angles, the working angles being those that are used when the tool is in operation, and the tool angles being those of the tool entity itself. It is important to understand the influences of the main geometries on the turning operation with relevance to tool wear.

The chip lands and slides on the rake face of the tool, and the rake angles will determine the direction of the chip flow. Instead of the normal rake angle, a definition of a side and a back rake angle is often used. In orthogonal cutting, there is no back rake angle, and only a side rake angle is considered. Depending on the rake angles, tools are called positive, neutral or negative. Positive inserts produce higher shear angles and reduce cutting forces. Negative inserts produce higher forces but is useful in interrupted cutting due to their higher shock resistance, since the initial contact with the workpiece material is away from the weak cutting edge. The chip flow will be more towards the workpiece. This is depicted in Figure 2.7.

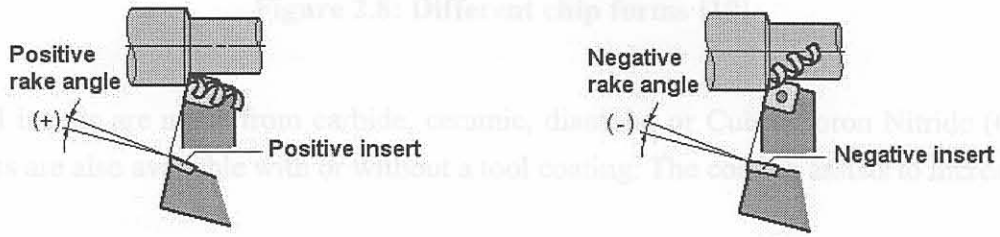


Figure 2.7: Positive and negative rake angles [19]

Tool tips have a small radius to minimise sharp feed marks on the workpiece surface. Increased radius also increases the strength of the cutting edge and decrease the rate of tool wear. However, a too large nose radius is also not advisable because it makes the tool susceptible to self-excited vibrations and also causes poor chip control. The theoretical surface roughness (R_t) is expressed as (also refer to Appendix I):

$$R_t = \frac{f^2}{8 \times r} \times 1000 \ [\mu m] \tag{2.4}$$

where:

f = feed rate [mm/rev]

r = nose radius [mm]

The chips produced during continuous turning can be of a ribbon, tangled, corkscrew, spiral or comma type. Tangled and corkscrew chips are unfavourable because they rub against the finished workpiece and can become entangled around the tool. Furthermore they also pose a danger to the operator and can cause tool breakage. The form of the chip is determined by the machining parameters, cutting fluid, material combination and other properties of the machine tool. Examples of different chips are shown in Figure 2.8, where the D type chips would be more favourable. Chip breakers can be clamped on tool holders or formed on tool inserts to assist in breaking long chips. They force the chip to curl toward the workpiece or tool, thus creating a large tensile stress that leads to breakage of the chip.

Jawahir *et al.* [20] investigated the effects of chip flow on tool wear with chip breaker type tools. It was found that the failure of these tools is generally due to improper groove utilisation, and hence the tool, machine and work material must be taken into account to evaluate the performance of these tools. A knowledge-based approach for designing chip breakers is suggested by Jawahir and Fang [21], tak-

ing the above-mentioned considerations into account. Other studies related to modelling the chip formation with analytical and numerical methods can be found in [22-29].

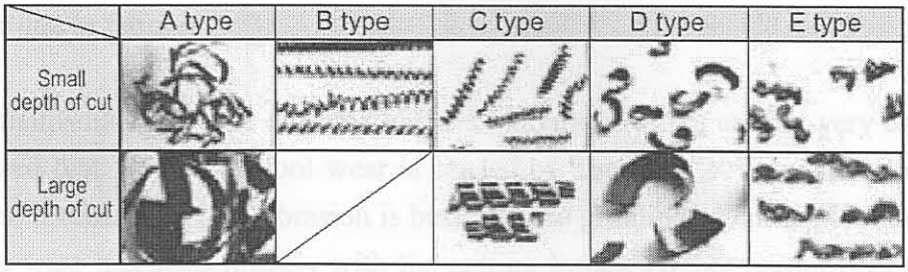


Figure 2.8: Different chip forms [19]

Turning tool inserts are made from carbide, ceramic, diamond or Cubic Boron Nitride (CBN). Many other variants are also available with or without a tool coating. The coating assists to increase tool life.

Another important phenomenon to turning operations is chatter vibrations. Chatter is self-excited vibration resulting from the generation of different chip thickness during machining. Initially, cutting forces excites a structural mode of the machine-workpiece system. This will leave a wavy surface finish on the workpiece. During the next revolution another wavy surface will be made in the same way. Depending on the phase shift between these two waves, the maximum chip thickness can grow and oscillate at a particular frequency that is close to a structural mode. This is called the regenerative chatter frequency [14]. Chatter cause a poor surface finish and can also lead to tool breakage. The analysis and prediction of chatter has been the subject of research for many years. Morimoto *et al.* [30] developed a piezoelectric shaker / actuator to regenerate the vibrations of the cutting process. In this way, unwanted vibrations such as chatter can be attenuated. The system is also helpful to determine the dynamic properties of the machine tool. Koizumi *et al.* [31] used a very interesting approach called the correlation integral in the time domain to identify chatter onset. Lägo *et al.* [32] designed a sensor and actuator integrated tool for turning and boring to control chatter. The tool holder shank vibrations are supplied to the actuator via a digital controller. An adaptive feedback control system is used to perform broadband vibration attenuation up to 40dB at different frequencies simultaneously.

2.3 Tool wear

2.3.1 Tool failure mechanisms

It is important to identify tool failure mechanisms in order to select appropriate machining parameters, and also for interpretation of the sensor signals during wear monitoring. If the underlying mechanisms are understood, phenomena in the sensor data can be attributed to certain tool failure mechanisms and modes. Mechanical loads, thermal loads, chemical reactions and abrasive loads, cause tool wear. The cutting conditions and the tool and workpiece materials influence these loads. The different loads can cause certain wear mechanisms, and depending on the loads, they may occur in combination. These mechanisms have either a physical or chemical characteristic that cause loss or deformation of tool material. Tool wear mechanisms can be classified into several types, summarised as follows [33]:

- abrasive wear resulting from hard particles cutting action
- adhesive wear associated with shear plane deformation
- diffusion wear occurring at high temperatures
- fracture wear due to fatigue

Other wear mechanisms are plastic deformation and oxidation, which are not very common in industry. It is estimated that 50% of all tool wear is caused by abrasion, 20% by adhesion, and the other 10% by other the mechanisms [3]. Abrasion is basically the grinding of cutting tool material. The volume of abrasive wear increases linearly with the cutting forces. Higher hardness of the tool material can reduce abrasive wear. During adhesion the high pressures and temperatures on the roughness peaks on the tool and the workpiece cause welding. These welding points are broken many times every second and cause removal of the tool material. Diffusion wear occurs at even higher cutting speeds where very high temperatures are present (especially when using hard metal tools).

2.3.2 Tool failure modes

Tool wear will generally occur as a combination of a number of wear modes, with one mode predominant. The dominant mode will depend on the dominant wear mechanism. For a given tool and workpiece combination, the dominant wear mode can be determined at different cutting speeds using the product of the cutting speed and the undeformed chip thickness [34]. The common wear modes are:

- flank wear
- crater wear
- chipping
- breakage
- nose wear
- plastic deformation
- cracking
- notch wear

The basic interpretations of causes, mechanisms, types and consequences of tool wear are summarised in Figure 2.9 (adapted from [34]). The consequences of tool wear are deviations in shape and roughness of the machined part, which cause the part to be discarded because it is out of the allowable tolerance. Figure 2.10 is a graphical representation of the different tool failure modes. Although they are shown separately in the figure, they can also occur in combinations, *e.g.* flank wear and notching. The most widely researched tool failure modes for turning with single point tools are flank wear, breakage and crater wear. Flank and crater wear are generally accepted as the normal tool failure modes, because the other failure modes can be avoided by selecting the proper machining parameters. In fact, it has also been shown that crater wear can also be avoided by selecting sufficiently low feed rate and cutting speed [35]. The growth of flank and crater wear is directly related to the cutting time (or length of cut), unlike some of the other failure modes, which can occur unexpectedly, even with a new tool. It is already well established that flank wear generally has the greatest influence on the workpiece dimensions and surface quality [36]. For this reason ways to predict flank wear has been the pursuit of researchers for many years. However, hard metal tools such (*e.g.* synthetic diamond) sometimes exhibit other dominant failure modes that affect workpiece quality [37], and it is thus of importance to comprehend the mechanics of the different failure modes.

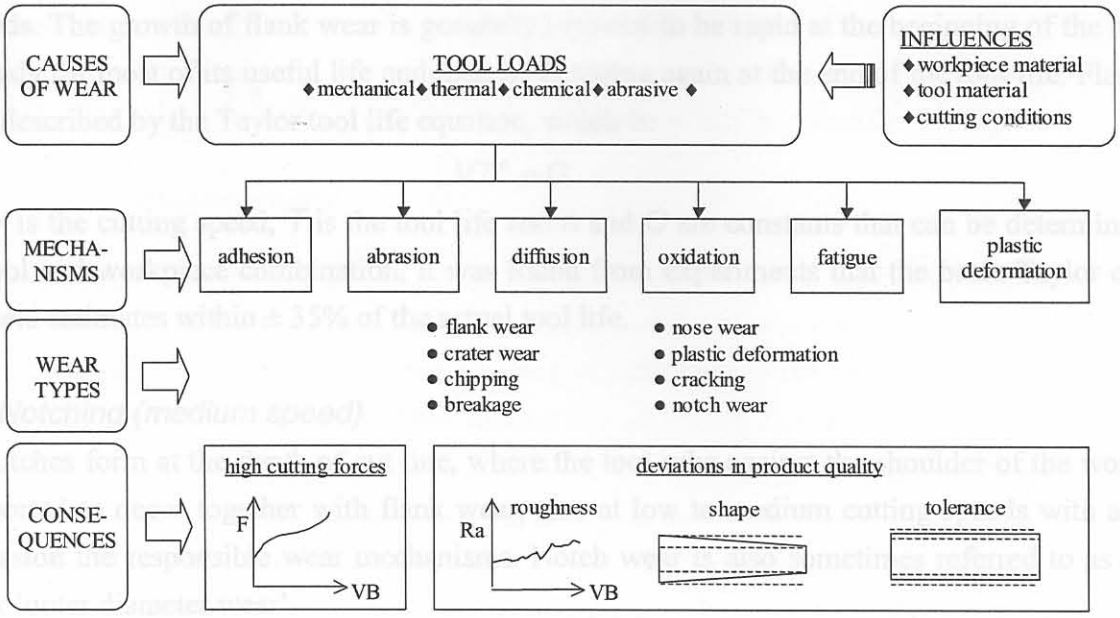


Figure 2.9: Causes, mechanisms, types and consequences of tool wear (adapted from [34])

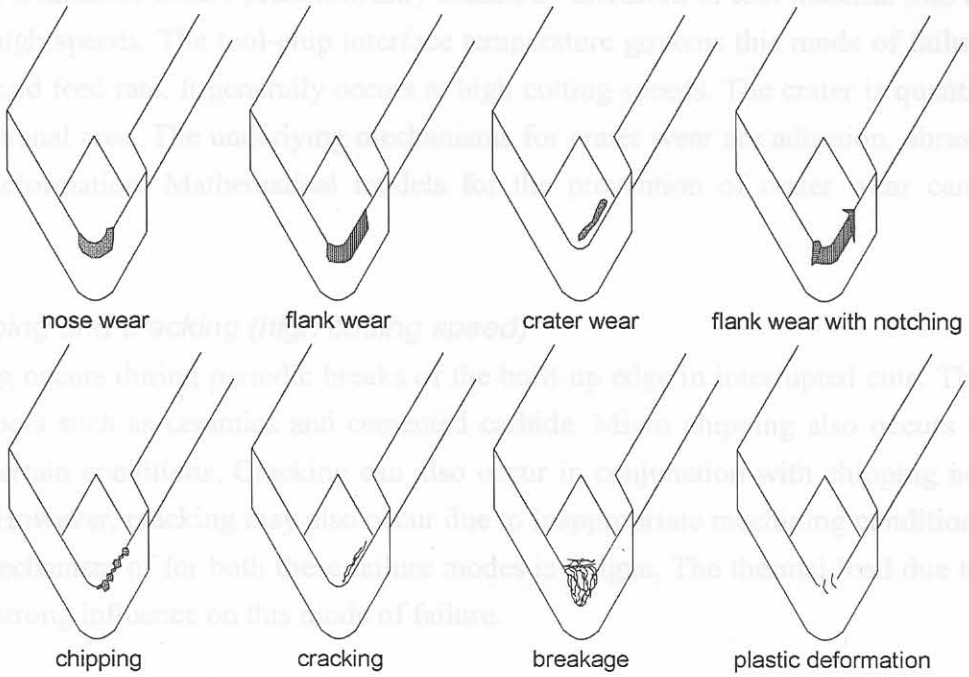


Figure 2.10: Tool failure modes

A. Nose wear (low speed)

Nose wear or edge rounding occurs through the abrasion wear mechanism on the major edges of the tool. Nose wear is caused by the selection of inappropriate cutting conditions and occurs on the tool tip at low cutting speeds.

B. Flank wear (medium speed)

Flank wear is the volumetric loss at the top of the tool edge, and is mainly caused by abrasion. Some authors affirm that the flank wear in coated tools first occurs due to abrasion, and at a later stage of tool life adhesion and diffusion also occurs [38]. Flank wear normally occurs at medium to low operat-

ing speeds. The growth of flank wear is generally reported to be rapid at the beginning of the tool life, then steady for most of its useful life and then accelerating again at the end of the tool life. Flank wear is often described by the Taylor tool life equation, which is:

$$VT^n = C \tag{2.5}$$

where V is the cutting speed, T is the tool life and n and C are constants that can be determined for a given tool and workpiece combination. It was found from experiments that the basic Taylor equation could yield estimates within ± 35% of the actual tool life.

C. Notching (medium speed)

Wear notches form at the depth of cut line, where the tool rubs against the shoulder of the workpiece. It is reported to occur together with flank wear, also at low to medium cutting speeds with adhesion and abrasion the responsible wear mechanisms. Notch wear is also sometimes referred to as 'groove wear' or 'outer diameter wear'.

D. Crater wear (high speed)

Crater wear is a mode of failure predominantly caused by diffusion of tool material into the chip when operating at high speeds. The tool-chip interface temperature governs this mode of failure, influenced by the speed and feed rate. It generally occurs at high cutting speeds. The crater is quantified by depth and cross-sectional area. The underlying mechanisms for crater wear are adhesion, abrasion, diffusion and plastic deformation. Mathematical models for the prevention of crater wear can be found in [39,40].

E. Chipping and cracking (high cutting speed)

Edge chipping occurs during periodic breaks of the built-up edge in interrupted cuts. This is common with brittle tools such as ceramics and cemented carbide. Micro chipping also occurs with diamond tools under certain conditions. Cracking can also occur in conjunction with chipping near the end of the tool life. However, cracking may also occur due to inappropriate machining conditions. The underlying wear mechanism of for both these failure modes is fatigue. The thermal load due to high cutting speeds has a strong influence on this mode of failure.

F. Plastic deformation (very high speed)

Plastic deformation is both a failure mechanism and failure mode. Plastic deformation starts when the temperature of the tool tip reaches a certain value. This implies that the tool yield strength is lowered below the existent normal stress. Further plastic deformation results in a temperature increase that causes complete failure. Mathematical models for the prevention of plastic deformation wear can be found in [39,40].

G. Breakage (very high speed)

Tool breakage or fracture is a mode of failure characterised by breakaway of material on the tool tip. Breakage occurs when the feed-rate is too high, or when a tool is used with too low fracture strength. Either plastic deformation and / or fatigue can be responsible for this mode of failure. It may also occur if a considerable degree of nose wear or severe depths of crater wear is present. Breakage is com-

mon with brittle tools such as ceramics and CBN. Breakage will normally occur at very high cutting speeds, but may also occur at lower speeds if an inappropriate tool material is selected for a certain task. Mathematical models for the prevention of tool breakage can be found in [39,40].

2.3.3 Tool wear measurement

The quantification of tool wear can be a very subjective matter due to human interpretation of a microscopic picture. In order to be consistent, it is suggested by machine tool developers that the same person perform all the wear measurements during experiments. Wear measurement of tools is done through the implementation of ISO 3685, summarised in Figure 2.11 for turning (after Dimla [41]). Flank wear is quantified in terms of VB, which is the mean of the wear height on the tool flank. The length of flank wear is also measured in terms of b. The maximum flank wear is VB_{max}. The notch wear can also be measured (if notch wear is present), in terms of VB_{notch}. Crater wear is quantified in terms of the crater depth K_T, and sometimes also the distance between the cutting edge and crater centre, quantified in terms of K_M.

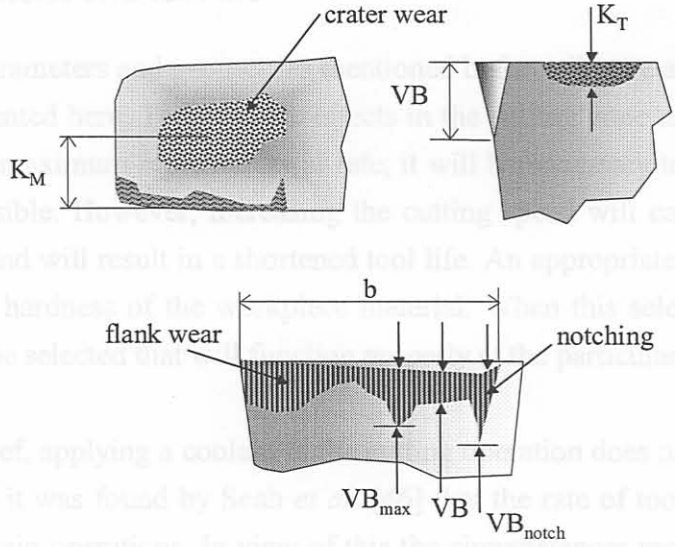
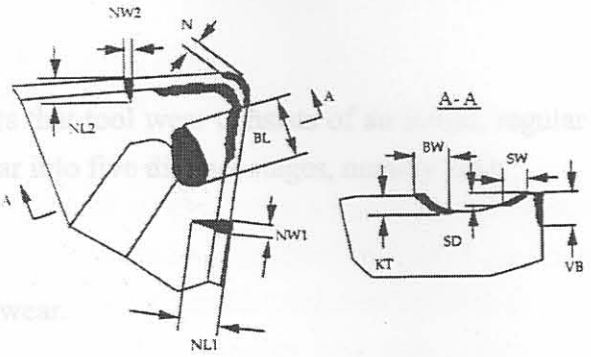


Figure 2.11: Convention for tool wear quantification (ISO 3685 after [41])

Sometimes other wear parameters are used. Recently, a method for measuring the wear for grooved tools was suggested by Jawahir *et al.* [42,43]. The suggested method has become standard practise in research for assessing wear with coated grooved tools. These parameters are shown in Figure 2.12. It should also be mentioned here that Jawahir *et al.* [44] extended their work with grooved tools in developing an equivalent toolface model. In essence, an equivalent flat toolface can be determined for grooved tools. The equivalent toolface can then be used to predict certain wear and failure modes of the grooved tools. Parakkal *et al.* [45] also proposed a mechanistic model for modelling the cutting forces during turning with grooved tools.



VB	flank wear	*N*	nose wear
BW	width of groove backwall wear	NL_1	notch wear length on main cutting edge
BL	length of groove backwall wear	NW_1	notch wear width on main cutting edge
KT	depth of groove backwall wear	NL_2	notch wear length on secondary cutting edge
SW	width of secondary face wear	NW_2	notch wear width on secondary cutting edge
SD	depth of secondary face wear		

Figure 2.12: Tool wear parameters for grooved tools [42]

2.3.4 Machining conditions and tool life

The various machining parameters and geometries mentioned before all have an effect on tool life, and a short discussion is presented here. Temperature effects in the cutting zone and in the tool itself govern tool life. To enable a maximum metal removal rate, it will be necessary to use the highest cutting speed and feed rates possible. However, increasing the cutting speed will cause increasing temperatures in the cutting zone and will result in a shortened tool life. An appropriate cutting speed should be selected according to the hardness of the workpiece material. When this selection is made, a certain grade of tool insert must be selected that will function properly at the particular cutting speed.

In contrast to popular belief, applying a coolant to the cutting operation does not necessarily reduce the rate of tool wear. In fact, it was found by Seah *et al.* [46] that the rate of tool wear increases when a coolant is applied for certain operations. In view of this the circumstances must be evaluated properly before a coolant is applied by default. Feed rate does not have a large influence on tool life, but there is a certain feed rate for optimal tool life for each operation.

Depth of cut also does not have such a large influence on the tool life compared to cutting speed. The depth of cut must be determined according to the required stock removal, shape of the workpiece and also rigidity of the machine and workpiece. A very small depth of cut causes friction and will shorten tool life. The rake angle also influences tool life. Positive rake angles can increase tool life but can only be used in certain applications. The side cutting edge angle (lead angle) can lower the impact on the tool and has an effect on the cutting forces. Increasing the side cutting edge angle increases chip contact length and decreases chip thickness. The result in that the cutting force is dispersed over a larger area and the tool life is prolonged. However, chip control and breakage is more difficult to achieve with increasing side cutting edge angle. Furthermore, it will increase the thrust forces and as a result cannot be used with long or slender workpieces.

2.3.5 Tool wear stages

It is assumed by most authors that tool wear consists of an initial, regular and fast wear stage [47,48]. Some authors divide tool wear into five distinct stages, namely [38]:

1. Initial stage of wear.
2. Regular stage of wear.
3. Micro-breakage stage of wear.
4. Fast wear stage.
5. Tool breakage.

It has been established by various researchers that the initial and fast (before tool breakage) stages wear occur more rapidly than the regular stage. A reason for this behaviour is very seldom given. Bonifacio and Diniz [38] explained that during the fast wear stage with coated carbide tools, the tool loses its coating and the tool substrate (which has less resistance) begins to perform the cut. During the initial stage, the tool edge loses its radius quickly and after which the process stabilise for a given amount of time. Lim [49] found that with a tungsten carbide tool the initial wear rate is also faster due to the breakage of the sharp cutting edge after which a finite wear land forms. Flank wear in relation to time or length of cut will typically appear as depicted in Figure 2.13. From the literature it is unclear if failure modes other than flank wear display this kind of behaviour.

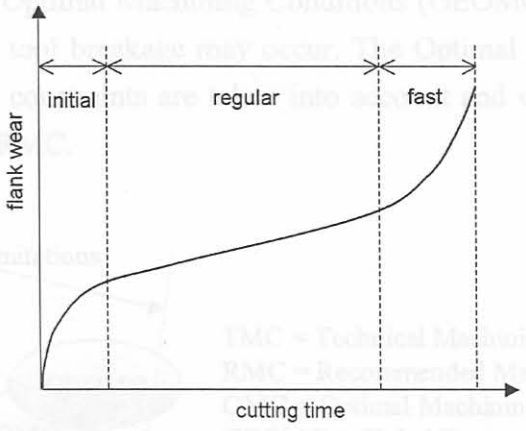


Figure 2.13: Flank wear in relation to cutting time

The wear rate is in fact a function of the wear mechanisms, and therefore any increase or decrease in the wear rate must be accounted for by investigating the wear mechanisms. The abrasion and adhesion mechanisms cause flank wear, and the cutting temperature influences the mechanisms. Increasing tool wear also cause an increase in the cutting temperature. This in its turn can cause increased activity of the abrasive and adhesive wear mechanisms, and could be the reason for increased wear rates near the end of tool life.

2.4 Tool wear mapping

2.4.1 Introduction

In order to optimise a metal removal process with a tool wear constraint, engineers must have ready access to information pertaining to the wear process of interest. User-friendly databases must be established to provide appropriate information for the choice of optimal operating conditions for a given set of materials. The best approach to present complex wear data is through a wear map. Such a map can provide a multi-dimensional graphical presentation of wear data. Different types of wear maps exist, such as wear-mode, wear-transition, wear-regime and wear-mechanism maps. The wear-mode map is the most common, where regions of the dominant wear mode are given for a range of operating conditions. Of course, wear-mode maps are not only useful for TCM. Wear maps are commonly used for the optimisation of a machining process and during the development of adaptive control systems for machine tools. This section will discuss a few examples of tool wear mapping.

An operating conditions map given by Lundholm [50] is reconstructed in Figure 2.14. The figure maps the different regions of operating conditions for ranges of feed rate and speed. The Technical Machining Conditions (TMC) is bounded only by the technical performance of the machine tool. The supplier of the tool will also state the Recommended Machining Conditions (RMC), which is in a conservative safety zone. If the process is economically optimised without taking the technical constraints into account, the Global Economical Optimal Machining Conditions (GEOMC) are found. However, in this region, excessive tool wear or tool breakage may occur. The Optimal Machining Conditions (OMC) can be computed when all the constraints are taken into account and will be safer than the GEOMC and more economical than the RMC.

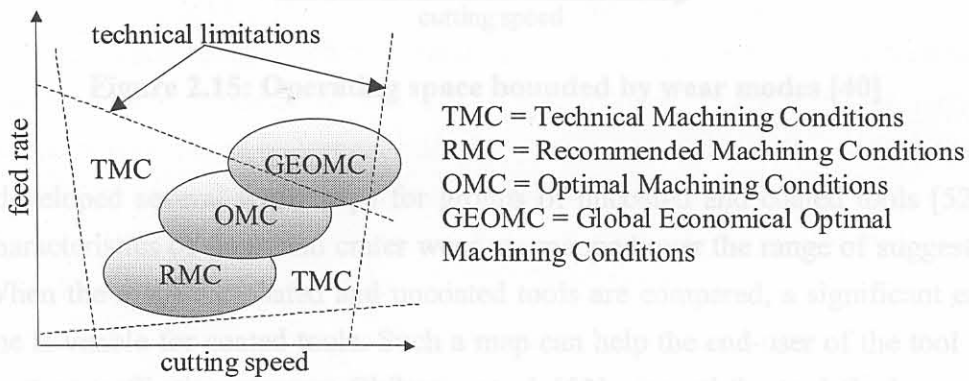


Figure 2.14: Classification of machining conditions [50]

The operating conditions map and wear mode maps are related, because both establish 'safety zones' where no excessive tool wear or catastrophic tool failure will occur. Furthermore, the dominant failure mechanisms and modes for a given range of operating conditions can also be determined. This information is very useful for:

- development of TCMS
- optimising of the cutting process

- adaptive control
- prevention of certain failure modes

However, to construct such a map demands a lot of experimental work, and the outcome of the map is very dependant on the tool and workpiece material combination. Especially due to the rapid development of new metals and advancements in machine tool technology, there will probably never be wear-mechanism maps for each and every tool and workpiece combination.

2.4.2 Wear map examples

The originator of graphical representation of tool wear data can be traced to Trent who produced a series of machining charts in the late 1950s [51]. The concept of these diagrams did not capture any further attention until Yen and Wright [40] proposed a map for turning tools. Yen and Wright constructed wear maps based on mathematical models. A reconstructed example of a tool wear map by them is shown in Figure 2.15. Lever *et al.* [39] produced very similar graphs to determine operational zones for the development of a machine learning system. The same types of analytical equations were used to construct the operational zones.

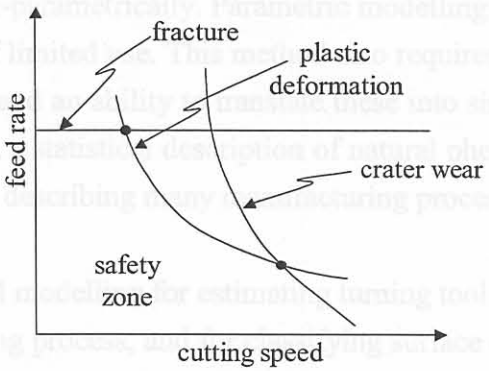


Figure 2.15: Operating space bounded by wear modes [40]

Researchers developed several wear maps for groups of uncoated and coated tools [52]. In both instances the characteristics of flank and crater wear are mapped over the range of suggested machining conditions. When the maps for coated and uncoated tools are compared, a significant enlargement of the safety zone is visible for coated tools. Such a map can help the end-user of the tool to employ the tool in the most cost-effective manner. Obikawa *et al.* [53] mapped the tool flank wear of a carbide tool in a 3-D graph to estimate the optimum cutting conditions and monitor the tool wear. In this case, experimental data were used to construct the map. Da *et al.* [54] also constructed tool wear maps for machining process optimisation, using analytical equations for tool wear and process constraints. The models were proved with experimental data. Recent work Li *et al.* [55] proposed a predictive mapping system for tool wear based on a modified tool wear model. The mapping system can predict wear rate maps accurately with cutting speed, feed rate and flank wear as parameters. The limitation of the system is the fact that the diffusion and adhesion constants for the two materials must be known. In order to establish the accuracy of the model, more work on a wider range of materials will be required.

2.5 Mathematical modelling

Mathematical models are very useful to study tool wear. Most models attempt to predict variables such as the cutting forces, temperatures, pressures, chip flow angles *etc*. Another mathematical approach is tool life equations that attempt to predict the life of the tool under certain machining conditions. With respect to TCM, these are referred to as a sensorless approach. The model is thus used without help from on-line sensors.

Three types of models are used, namely analytical (theoretical), computational (numerical) and empirical (experimental) models. Analytical models are useful to study the effects of tool geometry on the various machining parameters, but are too complex to be of any value in a real-time TCMS. The non-linear, stochastic and time invariant nature of machining processes makes theoretical modelling of machining processes very difficult [56]. Due to the complexity, modelling of the physical law representing the metal removal process cannot be performed for most cases.

Empirical models are models generated from experimental data. The output from empirical models is usually one or more empirical constants. These constants are different for every tool and workpiece combination, but are available in the literature for common combinations. Empirical modelling can be performed parametrically or non-parametrically. Parametric modelling usually represents an adaptation of the analytical model and is of limited use. This method also requires the inputs of an expert familiar with the relational mechanisms and an ability to translate these into simple rules [39]. Non-parametric modelling methods are based on a statistical description of natural phenomena. Empirical models have been used with great success for describing many manufacturing processes.

Grabec *et al.* [57] used empirical modelling for estimating turning tool sharpness, for the determination of surface roughness in a grinding process, and for classifying surface quality of paper. Ruiz *et al.* [58] used a multi-sensor empirical approach to estimate tool wear, and identified tool wear with three different empirical identification methods. The use of an analytical model for force reconstruction for wear identification was proposed by Braun *et al.* [59]. The model can also be used for the prediction of chatter onset. Ravindra *et al.* [60] proposed a mathematical model based on multiple regression analysis. The model describes the wear-time and wear-force relationships for turning operations. Good correlation was found between the cutting force and progressive tool wear. Lin *et al.* [61] describe the use of an abductive network for modelling surface roughness and cutting forces for turning. Abductive networks consist of several polynomial functions organised in layers. One advantage of this method is that the optimal network architecture is determined automatically, and requires less iterative work than Neural Networks (NNs).

Jawahir *et al.* [20] reviewed the most common tool-life relationships in 1995. A modified version of this review is presented in Table 2.1. Some problems were identified with these approaches [20]:

- The methods are not adaptable to all the different tool designs and chip-groove configurations that are used in industry, which have a large influence on the tool life.
- The influence of all the machining parameters on the tool life cannot be described in a systematic

manner.

- The methods do not incorporate the knowledge of human experts.

Table 2.1: Tool-life equations (T = tool life in minutes)

no.	*Tool life equation*	*Comments*	*Ref*
1.	Basic Taylor equation: $VT^n = C$ where V = cutting speed n = Taylor's tool life exponent C = empirical constant	Most widely used equation, however, constants C and n only apply for specific tool and workpiece combinations.	[36,49, 62]
2.	Taylor's reference-speed equation: $\left(\frac{V}{V_R}\right) = \left(\frac{T_R}{T}\right)^n$ where V_R is the reference cutting speed for reference tool-life T_R = 1 min.	n only applies to particular tool and workpiece combinations.	[62]
3.	Modified reference-speed equation with coating and groove effects taken into account: $T = T_R W_g \left(\frac{V_R}{V}\right)^{W_c \frac{1}{n}}$ where W_C = coating effect factor W_g = chip-groove effect factor	Same as reference speed equation, but more accurate and adaptable to coating and chip-groove effects.	[54,63]
4.	Extended Taylor equation: $T = \frac{C_2}{V^P f^q d^r}$ where f = feed d = cut depth C_2 = empirical constant	Better accuracy than basic Taylor equation, but requires several tool-life experiments.	[64]
5.	Temperature-based equation: $\theta T^n = C_3$ where θ = tool temperature C_3 = empirical constant	Equation is not feasible for use on the shop floor.	[20]

6.	Colding's equation based on ECT (Equivalent Chip Thickness): $y = K - \frac{(x-H)^2}{4M} - (N_0 - Lx)z$ where x = ln (ECT) y = ln (V) z = ln (T)	Many empirical constants, and machining parameters are integrated into single ECT parameter. However, tool life predictions are inconsistent.	[20]
7.	Basic Taylor equation including rake and clearance angle: $C \propto \left[(\cot\beta - \tan\alpha)^n F(\alpha,\beta)^{\frac{1}{\varepsilon}} \right]^{-1}$ where $F(\alpha,\beta)$ is a function of: α = rake angle β = clearance angle	A complicated relationship between tool-life and rake / clearance angles.	[20]
8.	Extended Taylor equation including cutting conditions and tool geometry: $T = C_4 V^n f^m d^P r^q s^t i^u j^x$	Require many tool-life tests to determine all the empirical constants.	[20,65, 66]
9.	Extended Taylor equation including cutting conditions and workpiece hardness: $V = \frac{C_5}{T^m f^y d^x (BHN/200)^n}$ BHN = Brinell Hardness Number	It is claimed to be a good approximation for tool-life ranges of 10-60 min.	[20]

Numerical approaches are methods like the Finite Element Method (FEM) and other types of computational simulations. These models are used to predict variables such as temperatures, forces, pressures and stresses in the cutting zone. Lately, many FEM approaches have been developed dedicated to certain machining operations. The FEM has many advantages, such as the fact that it can handle many different machining conditions, materials and geometries. Athavale and Strenkowski [67] recently published an overview of FEM modelling of machining. It is stated that there is a basic disagreement between researchers in the area of FEM modelling of machining. These unresolved issues are:

1. the failure mechanism at work of the workpiece material at the tool cutting edge
2. the stress distribution and frictional relationship at the tool-chip interface
3. the material flow past a rounded cutting edge, tools with wiper inserts or worn cutting tools

As a result simulation results vary close to the cutting edge and tool-chip interface. Further away, there is a good correlation among the results of the various researchers. There are three types of FEM formulations of machining operations:

1. Langranian (mesh is attached to the workpiece)
2. Eulerian (workpiece material is assumed to flow through a meshed control volume)
3. Arbitrary Langranian-Eulerian (utilise both formulations during iteration)

Thermal behaviour and material models are also required for accurate simulations of the cutting mechanics. To model tool wear with a FEM remains subject to continuous research. This is because the wear mechanics are very complex to model. Hence, FEM models are of limited usefulness for TCM. Despite this, FEM models provide the most comprehensive and accurate results compared to other existing techniques. There are many other applications for FEM modelling. Sandstrom [68] proposed a FEM for modelling the physics of high-speed machining. The model can assist in planning manufacturing processes on a sound technical foundation. Lovell *et al.* [69] proposed a FEM model for variable tool-chip interface and tool coatings. This can be used to assist to evaluate optimal tool coating parameters and wear rates. A picture of an explicit dynamic FEM model of the machining process from Lovell *et al.* [69] is pictured in Figure 2.16.

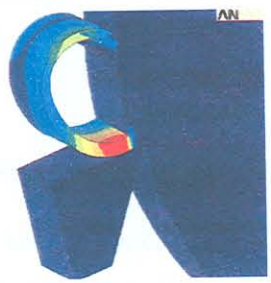


Figure 2.16: Explicit dynamic FEM model of the machining process [69]

Marty *et al.* [70] implemented a numerical simulation of machining that includes workpiece vibrations. The inclusion of vibration is very important because it has a very significant influence on the surface roughness of the machined workpiece. Lee *et al.* [71] used the FEM to determine the effect of a larger nose radius on the stress distribution in a tool insert. The areas where chipping and tool breakage will occur can be identified in this way. Marusich and Askari [72] used a numerical method to model residual stresses in machined surfaces, which is very important especially for components subject to fatigue. Very promising results were obtained.

The research works mentioned here are only a fraction of the activities in the area of FEM modelling of machining operations. In fact, Mackerle [73] presents a bibliography of FEM analysis and simulation of machining operations which covers the work from 1986-1996, and 675 research papers are included! Today, most researchers take the computational / numerical approach for research applications. Much less work is being done in the area of analytical modelling. Experimental / empirical models have the best practical application for industry. Artificial Intelligence (AI) models are also of the experimental type but are more often implemented as sensor-assisted models, and will be discussed in the next chapter. Jawahir *et al.* [74] reviewed methods for modelling turning operations at the University of Kentucky, one of most active groups in the U.S.A. The activities at the University of British Columbia are more focused on modelling of milling operations, and Altintas [75] presented an over-

view of their approaches. The reader is referred to these excellent overviews for more information on modelling of machining operations.

2.6 Optimisation of machining operations

2.6.1 Introduction

Tool wear studies are regularly included in manufacturing process optimisation studies. For this reason some of the basic concepts regarding machining process optimisation are discussed in this section. This will assist in comprehending some of the economic aspects with respect to tool wear. During the optimisation of most machining and manufacturing operations, the objective functions are related to economic criteria. Previous attempts to determine the optimal machining parameters can be divided into three main categories [65]:

- Computer Aided Design (CAD) approaches.
- Operations Research (OR) approaches.
- Artificial Intelligence (AI) approaches.

These approaches could be based on an off-line adjustment system, or an on-line Adaptive Control (AC) system.

2.6.2 Machining optimisation survey

The obvious optimisation problem for a turning operation will have feed and speed as variables, with the objective function linked to economic criteria. Ermer [76] developed a geometric programming technique to optimise the control variables for minimum cost, subject to constraints such as available horsepower, surface finish and available feeds and speeds. This very early work did not account for tool wear constraints. Da, Sadler and Jawahir [54] presented a computer aided methodology for predicting optimum cutting conditions in process planning of turning operations. This also involved the effect of the progressive tool wear on the performance of the machine.

Da *et al.* [54] also state that in an industrial machining process, machining performance varies due to tool wear. Empirical equations based on earlier research, were used to describe the behaviour of the different variables as well as their dependence on one another. Non-linear programming techniques were used to determine the constrained optimum cutting conditions for a certain tool wear state. Several papers are presented by the group of the University of Kentucky dealing with predictive modelling and optimisation of machining, mainly for turning operations [77-79].

Choudhury *et al.* [80] utilised an adapted version of the Taylor tool life equation (using force measurements as input) to predict the optimum cutting conditions in a turning process. This approach enabled them to predict the optimum conditions with the minimum number of experiments, given a database of the various material properties. A computer program reads the current machining conditions, determines the tool life from the Taylor equation, and then supplies the optimum parameters using a pre-established optimisation model. Zhou and Wysk [81] proposed a methodology for probabilistic optimisation in batch production, also using the Taylor equation.

Yen and Wright [40] proposed an optimisation procedure for adaptive control in machining. A safe working space is determined by the constraints of three different modes of failure. Control variables such as speed and feed rate are optimised for maximum metal removal rate. The gradual development of flank wear was also taken into account to update the optimisation dynamically. An important contribution was the establishment of a model that links the tool failure constraints with the control and state variables. Obikawa *et al.* [53] proposed a tool wear monitoring system integrated with an optimisation system for cutting conditions. The tool wear is estimated by monitoring the AR coefficients representing the power spectrum of the cutting force, and feeding it into two NNs. The machining parameters are optimised to ensure that a certain number of components can be manufactured reliably before the end of the tool life is reached. Jang and Seireg [35] proposed an optimisation procedure by which the machining parameters are optimised for specified surface conditions. Tool failure, tool wear, dimensional accuracy and chip formation are taken into account as constraints with a penalty function formulation. Maximum metal removal rate is achieved in conjunction with specified surface conditions.

2.6.3 Adaptive Control (AC)

Adaptive Control (AC) involves continuous changing in machining conditions by means of an on-line strategy, like the fuzzy-based AC system proposed by Tarng *et al.* [82]. Running an AC system based on one objective might cause an infraction on other constraints. This is why an AC system must be based on different control objectives, in order to optimise the process for the current machining conditions. Combining a range of sensors to interpret measured data can also extend the possibilities of an AC system. This is referred to as intelligent manufacturing [83,84]. The following monitoring and control functions are considered to be significant for such systems [50]:

- Advanced process monitoring, to protect from fatal events
- Adaptive Control Optimisation (ACO)
- Adaptive Control Constraint (ACC)

ACO attempts to adjust machining parameters in a direction that will optimise a predefined performance index. The aim of ACC systems is to adjust the machining parameters to their maximum possible values given the constraints of the process.

2.6.4 Approaches for Optimising Machining Operations

The conventional methods for selecting CNC machining parameters are based on textbooks or the experience of the operator. In most instances, the parameters are selected in a conservative manner in order to prevent failures such as tool breakage. As a result, the Metal Removal Rate (MRR) is low [65]. An optimisation strategy may consist of one or more of the following approaches:

A. Computer Aided Design (CAD) approaches

This off-line approach uses process, tool wear and cutting force models based on prior knowledge gathered from experiments. Based on these models, a computer simulation, using the Numerical Control (NC) code, can estimate the cutting force and tool wear. With these results, the MRR can be optimised without violating the machining constraints. The advantage of this approach is that it is easy to

implement and effective for most applications. A disadvantage is the fact that the approach can only be used off-line.

B. Operations Research (OR) approaches

The objective of Operations Research (OR) approaches is to minimise global machining cost by considering multiple criteria related to machining, for example the policies developed by Jeang [85] and Akturk and Avci [86]. These methods are used for off-line adjustment due to their computational difficulty. An advantage is the establishment of a reference model that can adjust to changes in the machining parameters. Gopalakrishnan and Al-Khayyal [87] demonstrated a machine parameter selection scheme based on geometric programming for turning, which is a typical OR approach.

C. Artificial Intelligence (AI) approaches

Artificial Intelligence (AI) based methods can be used to optimise a CNC machining process. AI methods can be either ACO or ACC based, or may even be an off-line system. AI based methods attempt to automatically optimise machining parameters based on sensor information. Reaction of the control system due to changes in process must be carried out within milliseconds to ensure the reliability of the process. There have been a number of studies on the application of AI techniques in on-line control [88,89]. These can be divided into three categories:

- Neural networks
- Probabilistic inference
- Knowledge-Based Expert Systems (KBES) [90]

2.6 Conclusion

In this chapter an introduction to the mechanics of metal cutting was given. Furthermore, an overview of the turning process in general together with the main variables concerned was described. An in-depth discussion on tool wear described wear mechanisms, modes, as well as methods for wear assessment and mapping. Lastly brief discussions on modelling and optimisation techniques of machining processes were given. The amount of literature in the area of metal cutting is enormous, and the brief background reviews presented here is relevant with respect to the research that follows in the further chapters.

The recent report on "Present Situation and Future Trends in Modelling of Machining Operations" [91] reviews the many different research activities in metal cutting. Many issues connected to metal cutting have not been resolved yet. In conclusion, it could be stated that numerical methods seem to be the best way to model machining operations. However, these methods are currently of limited use in on-line TCMSs.

3. Tool Condition Monitoring

3.1 Introduction

A wide variety of techniques for machining process monitoring have been developed through the years in industrial and academic projects. Due to the wide variety of manufacturing processes, it is not possible to apply a single technique to all operations. It is not uncommon for a monitoring system to be reliable for one process, but unsatisfactory for the next. However, a number of techniques exist that can be used for different processes, if the necessary adjustments are made to them.

In this chapter the various approaches to TCM are discussed. As an introduction to TCM, a brief overview of process monitoring in the area of manufacturing is given. There exist sensorless and sensor-based approaches to the problem of TCM. Sensorless approaches are not monitoring methods but are of relevance to this work. Basic sensorless approaches were discussed in Chapter 2. This chapter is concerned with sensor-based methods. It is widely accepted that intelligent, sensor based manufacturing is vital to achieve reliable operation of a manufacturing process. Sensor signals supply information about the manufacturing conditions that enables optimisation, control and decision-making. The information from sensors can be treated in numerous ways and research is aimed towards developing the best techniques to extract the relevant information from the signals. One way to utilise sensor information is through the use of Artificial Intelligence (AI) models. The use of AI in TCM will be discussed in more detail because it is the most relevant to this research. As a general case, designing a TCMS consists of the steps depicted in Figure 3.1. Various methods that could be used for each step will be discussed.

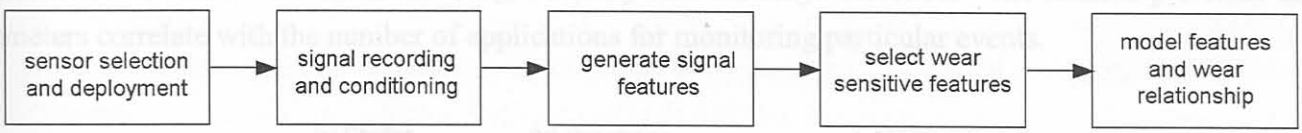


Figure 3.1: TCM steps

The reader is also referred to other overviews of sensor-assisted TCM, published by Dan and Mathew [92], Byrne *et al.* [5], Scheffer and Heyns [93] and Dimla [41]. A TCM database was also published by Teti [94]. This database includes more than 500 research papers focusing on TCM. The overview in this chapter is mainly concerned with developments in the literature. However, the commercial applicability of TCMSs is very important and as a result Appendices A and B were compiled that deals specifically with commercial systems. These two Appendices are the result of an exhaustive overview of commercial equipment and their application in industry. It is also important to compare the abnormal gap between research and industrial practice in this case, as it was an objective of this research to overcome this gap by developing a reliable TCMS for industry.

3.2 Sensors for general process monitoring

A wide variety of sensors for process monitoring are available. The most common sensors found in industry are force, power, vibration and acoustic emission sensors. Others include [5,95]:

- flame detector
- sound level sensor
- lubrication oil detector
- touch sensor
- edge position sensor
- limit sensor
- clamping force sensor
- speed sensor
- thermal deformation sensor
- coolant temperature sensor
- ph sensor
- level meter
- accelerometer (vibration)
- seismic sensor
- humidity sensor
- gas sensor
- chip monitoring sensor
- dust sensor
- temperature distribution sensor
- surface roughness sensor
- smoke sensor
- image sensor
- temperature sensor
- tool wear sensor
- tool damage sensor
- current sensor
- pressure sensor
- torque sensor
- acoustic emission (AE) sensor

These sensors and many more have found their rightful place in the manufacturing industry. Most of them are only used for a specific monitoring objective. The focus of monitoring may fall on one or more of the following areas [5,96]:

1. The machine (diagnostics and performance).
2. The tools for machining (wear, lubrication and alignment).
3. Workpiece (surface roughness, tolerance, geometry).
4. Process (chip formation, energy consumption, temperature).

Cho *et al.* [97] surveyed the different sensor approaches and their application in industry for research in Korea. A summary from [97] is shown in Figure 3.2. It is interesting to note that cutting force seems be the most popular for most applications. The second most popular method is Acoustic Emission (AE), which can also be used for different applications. Furthermore, it can be noted that the motor current is not used for wear monitoring, but only tool breakage detection. The relative pie chart diameters correlate with the number of applications for monitoring particular events.

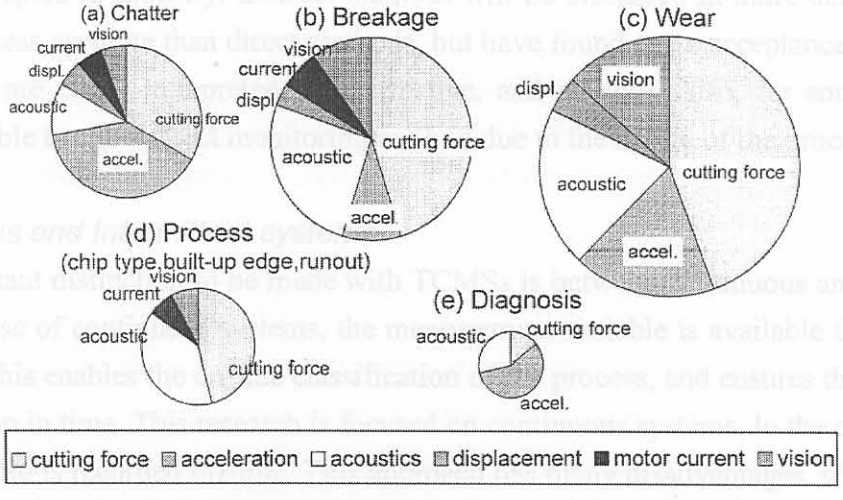


Figure 3.2: Sensor application in manufacturing process monitoring [97]

Sensor systems can communicate with the CNC control through different standards. A number of standard interfaces exist that can allow many sensor/ actuator systems to communicate with the CNC control, as was demonstrated by Pfeifer and Thrum [95]. This is very helpful to streamline the installation of sensor technology into modern machine tools. The development of smart sensor technology also presents new and exciting prospects for the manufacturing industry [5,98,99]. With smart sensors, the time needed for signal processing is reduced significantly, thus enabling faster response for on-line control. These sensors can also possess abilities such as self-calibration, self-diagnostics, signal conditioning and decision-making. In the future Analogue to Digital (A/D) converters may become obsolete for sensor systems, because this will be integrated within the sensor itself [100]. Smart sensors can also have built-in filters to filter certain vibration modes with application in intelligent structures [99]. The Transducer Electronic Data Sheet (TEDS) has also become an acceptable standard in sensor technology. This development, together with sophisticated signal processing software, makes inexpensive, fast and accurate measurements possible. The latest development in sensor technology is to develop wireless systems that can achieve high sampling rates across multiple channels.

It will be shown in this chapter that the emphasis in recent research is to integrate sensor systems. This enables more accurate and robust characterisation of a process. Integrated sensor systems can handle noisy input data, which is caused by random disturbances in the machining process. The sensor integration systems include learning schemes such as NNs, and have the ability to handle complex processes that defy analytical mathematical modelling.

3.3 Sensor-based tool wear monitoring

3.3.1 Introduction

A. Direct and Indirect systems

Approaches to monitor tool wear can be divided in two categories, namely direct and indirect. Direct methods are concerned with a measurement of volumetric loss at the tool tip, while indirect methods seek a pattern in sensor data from the process to detect a failure mode [5]. Direct methods are of less importance to this research. In general, direct methods are sensitive to dirt and chips, and are therefore not commonly accepted in industry. Indirect methods will be discussed in more detail. Indirect methods are said to be less accurate than direct methods, but have found more acceptance in industry due to the fact that they are easily interpreted, cost-effective, and reliable. Also, for some applications, it might not be possible to use a direct monitoring method due to the nature of the process.

B. Continuous and Intermittent systems

The second important distinction to be made with TCMSs is between continuous and intermittent systems [5]. In the case of continuous systems, the measurement variable is available throughout the machining process. This enables the on-line classification of the process, and ensures that sudden changes can be reacted upon in time. This research is focused on continuous systems. In the case of intermittent systems, the variable is recorded off-line. This approach has many disadvantages, which includes time losses and high costs. One practical application of an intermittent system can be a wear measurement on a magazine of tools while the machine is using a different tool.

C. Sensor requirements for tool wear monitoring

Monitoring usually takes place in very hostile environments. Subsequently, sensors used for tool wear monitoring should be robust and easy to install. Sensors used for TCM must meet certain requirements, such as [5]:

- Measurement as close to the machining point as possible.
- No reduction in the static and dynamic stiffness of the machine tool.
- No restriction of working space and cutting parameters.
- Wear and maintenance free, easy to replace and cost-effective.
- Resistant to dirt, chips and mechanical, electromagnetic and thermal influences.
- Function independent of tool and workpiece.
- Adequate metrological characteristics.
- Reliable signal transmission, *e.g.* from rotating to fixed machine components.

3.3.2 Force-based monitoring

It is well established that worn tools cause an increase in the cutting force components [5,101,102]. The dynamic and static force components generally increase with increasing tool wear (due to frictional effects). The difference between the static and dynamic components of the cutting force is shown schematically in Figure 3.3.

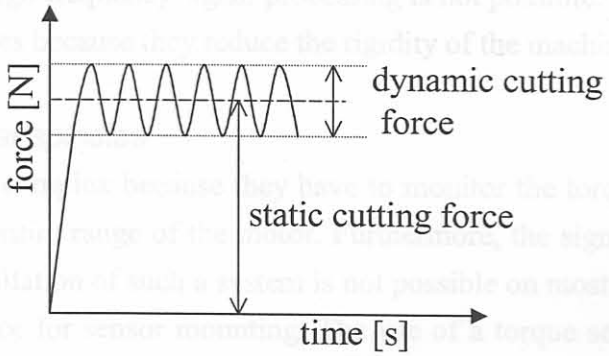


Figure 3.3: Static and dynamic forces

The different components of the cutting forces respond differently to machining parameters and tool wear modes. Depending on the type of process that is investigated and specific experimental setup, results among researchers vary. This can be contributed to dynamic effects of the machine tool and measurement equipment. Many types of sensors have been developed to measure cutting forces. These include [5] (also refer to Appendix A):

A. Direct measurement dynamometers

These sensors are based on the piezoelectric effect and can measure dynamic cutting forces very accurately. However, these sensors are expensive and in most cases not protected from overload, and therefore not used in industry. There is also some difficulty in protecting the sensors against cutting lubricant. Force-measuring tool turrets have been developed that can measure three force components, but are still very expensive.

B. Plates and rings

Force-measuring plates can be fitted with relative ease on turning machines between the turret housing and the cross slide, or between the turret disc and slide. These thin plates are fitted with piezoelectric force measuring sensors. These sensors have some advantages, but are subject to many disturbing factors, such as thermal expansion.

C. Pins, extension sensors

These sensors are suitable for tool breakage monitoring in rough machining. They are fitted on force carrying machine components to detect the cutting force indirectly. The identification of a suitable fitting position can only be determined experimentally, which is a disadvantage.

D. Measurement of displacement

Non-contact sensors to detect the displacement or bending of tools can be mounted directly on the tool [103]. However, these sensors are subjected to the high risk of damage and disturbances from chips, dirt and cooling lubricant.

E. Force-measuring bearings

Bearings and bushes can be specially fitted with strain gauges in certain positions to measure cutting forces. Force-measuring bearings require a low-pass filter due to disturbances from the ball contact frequency, and as a result high frequency signal processing is not possible. Force-measuring bushes are only accepted in special cases because they reduce the rigidity of the machine.

F. Force and torque at spindles

These systems can be very complex because they have to monitor the torque of the spindle with high resolution, and within the entire range of the motor. Furthermore, the signal must be transmitted on a non-contact basis. The installation of such a system is not possible on most machines because of a constraint on the available space for sensor mounting. The use of a torque sensor for TCM in drilling is described in [104] and [105].

3.3.3 Measurement of motor current

The measurement of motor current is an easy alternative to other sensors and can be installed without much difficulty. A wide range of sensors is available for this purpose. However, due to fluctuations in the signal due to friction, the signal is not accurate enough for wear monitoring. Also, tool breakage can only be detected after some damage has occurred. Spindle power is proportional to the cutting force in the primary motion, and is not the most sensitive direction for tool wear monitoring. The cutting process consumes only a small portion of the measured power of the spindle. However, monitoring systems based on the principle of spindle current can be successful when used with the right operation [102]. Ni *et al.* [106] used the spindle motor current to identify faults such as misalignment, oversize, undersize and wear for a tapping operation. A combination of wavelet analysis and Principal Component Analysis (PCA) of the motor current signal is used to distinguish between the faults. Tseng and Chou [107] use the reaction of the spindle motor's workload to cutting conditions to detect abnormalities. When these abnormalities are accumulated to the warning limits, the tool must be replaced. A

disadvantage is the appropriate selection of the warning level, which must be determined experimentally for different cutting conditions.

3.3.4 Acceleration

Piezoelectric accelerometers can measure the machine vibration caused by oscillations of cutting forces. Vibrations from a cutting process have components of free and forced vibration response. Furthermore, random and periodic behaviour can be observed. It has been shown by previous authors that the vibration levels change with tool wear (see references below). Industrial accelerometers fulfil the environmental requirements for tool wear monitoring because they are resistant to the aggressive media present in machining operations. Accelerometers are less expensive than most force sensors, and can measure vibration levels within a very wide frequency range. For these reasons, accelerometers are often used for TCM [38,102, 108-112]. Kim and Klamecki [113] also reported the use of torsional vibration (using a Laser-Doppler Vibrometer) for monitoring the wear of milling cutters.

One of the main difficulties of monitoring the tool life with acceleration is to identify the frequency range that is influenced by tool wear, since machining processes comprise of many factors that produce vibrations that are not related to tool wear. Bonifacio and Diniz [38] suggest that the useful frequency range falls between 0 – 8 kHz. It would seem that the frequency range sensitive to tool wear depends on the specific machining operation, and must be determined experimentally. A 'global' range that would satisfy all machining operations does not exist.

3.3.5 Acoustic emission

Cutting processes produce elastic stress waves that propagate through the machine structure. Different sources in the cutting process generate these stress waves known as Acoustic Emission (AE). Sources of AE in metal cutting are:

- Friction on the tool face and flank.
- Plastic deformation in the shear zone.
- Crack formation and propagation.
- Impact of the chip at the workpiece.
- Chip breakage.

The fact that crack formation generates AE, makes AE very useful for tool breakage detection. Generally, collection of the AE requires special hardware that can bandpass filter the signals to the AE range (between approx. 50kHz – 250kHz). Amplification is also required and an analogue root mean square (rms) circuit with a short time constant is often included to collect the rms AE level. The different steps required to collect AE for are depicted in Figure 3.4 (adapted from [114]).

Araujo *et al.* [115] investigated the sliding friction as a possible source of AE during metal cutting. The AErms in different frequency ranges was collected for different widths of cut and also with the tool rubbing against the workpiece (without cutting). It was found that the level of AE remains almost constant for all width of cut conditions, and hence it can be concluded that the main mechanism for AE

during metal cutting is the sliding friction between the tool and workpiece. Consequently, an increase or decrease of AE can be expected with tool wear depending on the effect on the sliding friction due to tool wear. It is also believed that the cutting temperatures will affect the level of AE due to thermal expansion effects. The effect of plastic deformation with other materials is currently under investigation. Chio and Liang [116] investigated AE with tool wear and chatter effects in turning. A model is presented that can predict the chatter AErms amplitude with certain severities of flank wear. Good correlation was found between the model and experimental results. Kim *et al.* [117] reports on the use of AE to monitor the tool life during a gear shaping process. The AErms is collected and used in a software program to predict the remaining tool life.

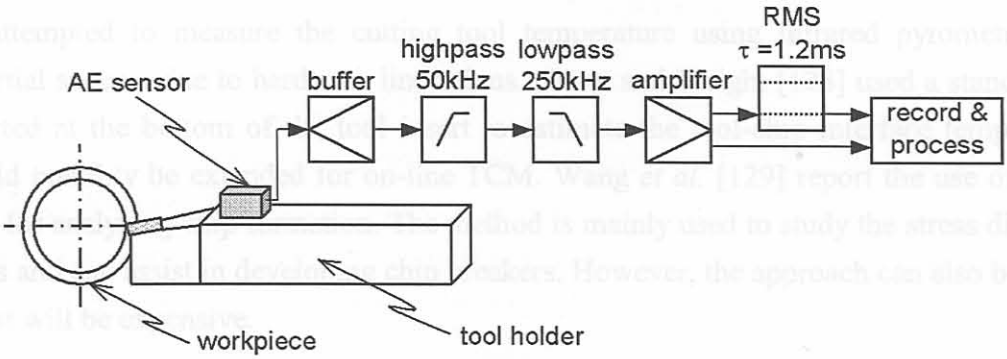


Figure 3.4: Steps for collecting AE during turning

Although a wide range of AE sensors exist, only a few can withstand the hostile environments of machining processes. AE sensors specially designed for use on machine tools are available, and these can be attached anywhere on the machine tool. A new concept is to use a coolant stream to transmit the AE waves from the tool to the sensor, for example the system presented by Dollinsek and Kopac [118]. The advantage is that the distance between the cutting area and the sensor is reduced, and thereby damping effects are minimised. Some problems with this approach are that bubble free coolant is required, and monitoring may be disturbed when chips pass through the coolant stream. Dollinsek and Kopac compared different tool insert types and found that the AE is most sensitive to tool wear but is also affected by the insert type. Another approach is to use non-contact transmission of the signal, allowing measurement near the process.

One problem still lies with an appropriate interpretation of the AE frequency spectrum. In most studies, an explanation for the choice of certain frequencies and their advantages are not given or not investigated. In fact, Jemielnaik [114] found that using the average value of AE (or AErms) is most suitable for TCM. A similar result was found during the course of this research (refer Chapter 4). Li [119] presents an overview of using AE for TCM in turning operations. It is stated the AE is heavily dependant on cutting conditions, and as a result methods should be employed to handle this problem effectively. Some methods are proposed that include advanced signal processing, sensor fusion and modelling techniques for tool wear and breakage monitoring [102,114,120-126]. There are also industrial implementations, and these are described in Appendix A.

3.3.6 Temperature monitoring

The high temperatures around the cutting edge during machining have a direct influence on the tool wear. The cutting temperature also affects chip formation and surface quality. The high frictional forces when cutting with worn tools cause higher temperatures. The heat is removed from the process by the chip (approx. 90%) and the workpiece and tool itself (approx. 10%). If the temperature in the cutting zone can modelled or measured, it will provide a complete solution to many problems encountered with machining. However, measuring the temperature directly is virtually impossible. Accurate temperature modelling for some machining operations is now possible by numerical techniques such as the FEM.

Lin [127] attempted to measure the cutting tool temperature using infrared pyrometry, but only achieved partial success due to hardware limitations. Chow and Wright [128] used a standard thermocouple inserted at the bottom of the tool insert to estimate the tool-chip interface temperature. The method could possibly be extended for on-line TCM. Wang *et al.* [129] report the use of an Infrared (IR) camera for analysing chip formation. The method is mainly used to study the stress distribution in cutting chips and can assist in developing chip breakers. However, the approach can also be considered for TCM, but will be expensive.

Using a remote thermocouple technique seems to be the only practical method for temperature monitoring for machining. This renders the temperature approach very ineffective. If an accurate and cost effective method can be established to estimate the temperature in the cutting zone, the technique will be very useful for TCM. Klocke and Hoppe [130] used a special fibre-optic pyrometer embedded into the tool insert to measure the temperature directly in the secondary shear zone for high-speed machining. The result was correlated with a FEM model and a good agreement was found. It is unclear if this approach could be used for TCM, but seems to be the best attempt up to date.

3.3.7 Ultrasonic methods

Abu-Zahra *et al.* [131,132] describe the use of an ultrasonic system for indirect tool wear measurement. With this approach, an ultrasonic signal is transmitted through oil to the tool insert. The reflection / echo of the ultrasonic waves is then collected with the transceiver. When the tool wears, more ultrasonic energy is reflected. The use of a calibration mark on the tool insert assists to quantify the severity of flank wear and eliminates temperature effects on the ultrasonic signals. The ultrasonic measurements are made when the tool is not engaged to the workpiece. The approach is very refreshing but somewhat limited in application and not yet cost-effective enough for industrial implementation.

Cho *et al.* [97] also report on the use of an ultrasonic sensor in a very interesting overview paper dealing with the research and developments in Korea. The methodology is similar to that of [131], but a thermocouple for temperature measurements is also included. A diagrammatical layout of the ultrasonic approach is shown in Figure 3.5 [97].

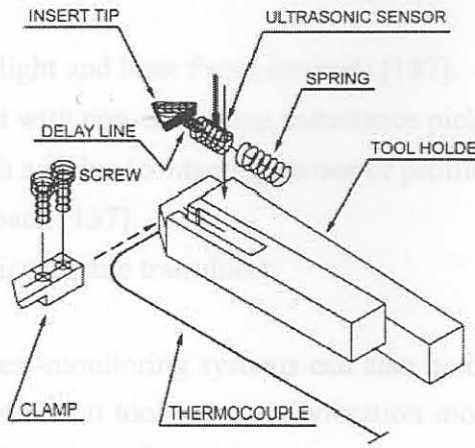


Figure 3.5: Ultrasonic TCM approach [97]

3.3.8 Vision systems

The use of vision systems for tool wear measurement is described by Kurada and Bradley [133] as well as Novak *et al.* [134]. In these approaches, a special camera is installed on the machine tool to assess the tool wear when the tool is not engaged in cutting. The digital picture taken by the camera is processed with special techniques and can yield the sizes of the flank and nose wear. The vision systems are accurate but have some disadvantages. One difficulty is to determine the flank and crater wear simultaneously with one camera. This problem was overcome by Karthik *et al.* [135]. In this case, a 3-D vision system was developed using only one camera that takes pictures from different angles. It was shown that the system could determine the average crater depth for different wear geometries. Another problem is the costs involved in installing and calibrating such a system on a CNC machine tool. Furthermore, chips, cutting fluid or components of the machine tool can restrict the line of sight of the camera.

3.3.9 Surface roughness monitoring

A. Introduction

Surface roughness is one of the most important factors in evaluating the quality of the machining operation. Because it is sometimes easier to measure the surface roughness of the machined component than to measure the tool wear, surface roughness estimation can be utilised to monitor the tool wear [136]. Cutting conditions, such as cutting speed, feed rate, depth of cut, tool geometry and material properties of the tool and workpiece, significantly influence the surface finish of the workpiece material. If these factors are known and set correctly, an in-process surface roughness measurement system can also indicate a worn tool [137,138].

Surface inspections in industry are typically done as a post-process operation, which is time consuming and uneconomical since a number of non-conforming parts can be produced prior to inspection. This underlines the importance of devices to monitor surface finish continuously without interrupting the machining process. Several methods have been proposed to estimate surface roughness on-line in flexible manufacturing systems. Some of these methods are [139]:

- Correlation between surface roughness and cutting vibration to develop an on-line roughness

measuring technique.

- Image processing, stray light and laser focus methods [137].
- Roughness measurement with non-contacting inductance pick-up.
- Direct measurement with a stylus (contacting sensor or profilometer).
- Ultrasonic sensing approach [137].
- Sensing with a special air pressure transducer.

As with TCM systems, roughness-monitoring systems can also be divided into direct and indirect approaches. This section concentrates on tool wear and vibration monitoring with relevance to surface roughness monitoring. The calculation of surface roughness parameters for machined parts is discussed in Appendix I.

B. Surface roughness analysis and tool wear

The surface roughness of machined components holds direct correlation with tool wear [138,140]. A logical consequence is to use the roughness information to control the machining operation as the tool wears. To maintain a certain roughness, the feed and cutting depth must either be increased or decreased to maintain the workpiece quality. For this, relatively simple geometric control systems can be developed that measures the roughness, calculates an error value, and then changes certain machining parameters accordingly. The ultimate goal is to develop an automated in-process monitoring system that would counteract any troublesome external factors. Process parameters could be varied in process with an adaptive or geometric control scheme, which would ensure consistent part quality [137].

Bonifacio and Diniz [38] found that vibration of the tool is a reliable way to monitor the growth of surface roughness in finish turning, and can be used to establish the end of tool life for these operations. Flank and groove wear mostly influence surface roughness. Some researchers found that there is increased amplitude of roughness at the beginning stages of cut, a lesser tendency in the middle and again an increasing tendency at the end of tool life.

C. Vibration monitoring and surface roughness analysis

The average surface roughness of a machined part can be assumed to be the result of the superpositioning of a theoretical profile computed from cutting kinematics, and of the oscillatory profile determined by the relative vibration between the cutting edge and the workpiece [141]. The random resistance against cutting (stick-slip process between the chip and the tool) causes the relative vibration between the tool and workpiece. The ideal or theoretical surface profile can be easily calculated from the cutting kinematics. The actual surface profile can be measured, or it can be estimated by measuring the relative vibration between the tool and the workpiece. This makes it possible to determine the surface roughness on-line without interrupting the machining process. However, there are a lot of practical problems involved when working in a real manufacturing environment.

One problem is that chatter between the tool and workpiece causes large vibrations that cannot be superimposed on the surface roughness. Another problem is that loose metal parts and other external factors easily distort signals from the sensors. However, the method has been successfully implemented in

dry turning with ferrous metals by Jang *et al.* [139]. In this case the kinematics of the machine tool are taken into account to estimate the roughness from vibration signals collected during cutting. They suggested that further research be done in this field.

Bonifacio and Diniz [38] did experiments with coated carbide tools in finish turning, measuring in the 0 – 8 kHz range. The vibration was measured on two channels, one in the cutting direction and one in the feed direction. The rms value was used to compare sets of measurements. They also varied the feed and cutting speeds during different experiments. It was found that cutting speed had a much larger influence on the tool life than the feed, and that vibration and roughness measurements correspond to a certain severity of tool wear at a given time.

3.3.10 Other methods

Some of the other methods for indirect / direct tool wear monitoring (excluding surface roughness approaches) are:

- Use of a non-contact capacitive sensor [142]
- Laser scatter methods [143]
- Fibre-optic sensor [144]
- Audible Emission [9]

Future research should be directed towards directly comparing different sensor methods for certain machining processes. Choi *et al.* [145] developed a single sensor for parallel measurement of force and AE. A FEM analysis was carried out to determine the optimal position for the sensor away from the tool holder. The reason for a more indirect measurement is because dynamometers sometimes restrict the working space of the machine tool. The approach was successful for breakage detection but no wear estimations are reported. Barrios *et al.* [102] compared AE, vibration and spindle current for TCM during milling. It was found that the spindle current is the most sensitive sensor for detecting tool wear, and found that AE is the least sensitive. However, contradictory results are reported in other publications, and hence more research would be required to ultimately determine which method will yield the best results for continuous estimation of tool wear. Govekar *et al.* [146] compared force and AE methods for TCM, and concluded that the best result is achieved when the sensor information is combined. Dimla and Lister [147] compared the use of force and vibration signals for TCM and also combined the features obtained in a single decision making technique [148]. Similar comparative studies were conducted during this research and are reported in Chapter 4.

3.4 Decision making in sensor-assisted TCM

3.4.1 Introduction

With the sensor information from the different sensor systems described in the previous section, a decision must be made regarding the tool condition. In complex problems it is advantageous to combine knowledge from sensor data to achieve the best results. Sick [6] recently proposed a generic sensor fusion architecture for TCM, shown in Figure 3.6.

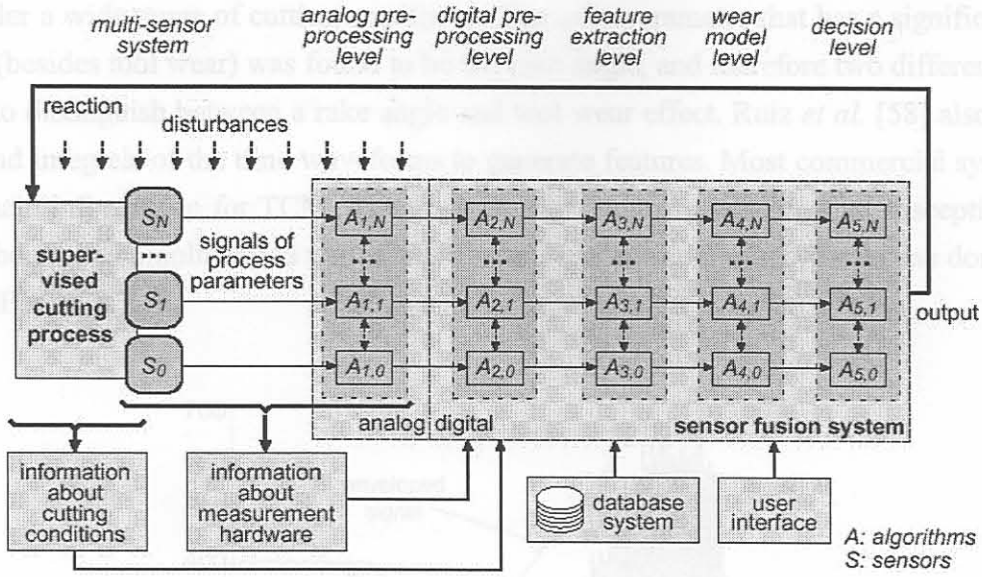


Figure 3.6: Sensor fusion architecture for TCM by Sick [6]

Fusion of sensor information can occur at any of the levels. Analogue and digital pre-processing consist of signal amplification, conditioning, filtering, calibration, temperature compensation *etc.* The feature extraction step is one of the most important steps, because here the sensor signals must be reduced to only a few appropriate wear-sensitive values. Many different methods are available to achieve this and will be discussed in further detail. The wear model level establishes a relationship between the chosen features and the tool condition. In many cases NNs are used in this step and sensor fusion also occurs. A decision level can also be included where a final decision can be made with respect to the tool condition, *e.g.* a "competing experts" approach if a TCMS is used in conjunction with a tool-life equation. In many cases the decision is made directly from the NN output. A discussion on the techniques for feature extraction, wear model and decision-making for TCM follows.

3.4.2 Feature extraction

Most decision-making techniques for process monitoring are based on signal features. Through appropriate signal processing, features can be extracted from these signals that show effective and consistent trends with respect to tool wear. Once these features are extracted through preliminary processing of the signal, the tool condition can be predicted with pattern recognition or other classification techniques [149]. Features are mainly derived through time, frequency, joint time-frequency domain signal processing or statistical analysis.

A. Time domain

Features extracted from the time domain are mostly basic values such as the signal average, mean, or a root mean square (rms) value. Other techniques include the shape of enveloping signals, threshold crossings, ratios between time domain signals, peak values and polynomial approximations of time domain signals. It has been found that some of the time domain features are very useful and they are easy to implement. Bayramoglu and Dungel [150] investigated the use of several different force ratios (calculated from the static cutting forces). It was found that certain force ratios can be used to monitor

tool wear under a wide range of cutting conditions. The only parameter that has a significant influence on the ratios (besides tool wear) was found to be the rake angle, and therefore two different force ratios are required to distinguish between a rake angle and tool wear effect. Ruiz *et al.* [58] also report using derivatives and integrals of the time waveforms to generate features. Most commercial systems rely on the time domain information for TCM. The time domain features are somewhat susceptible to disturbances and should be complimented with features from another domain. Typical time domain features are shown in Figure 3.7.

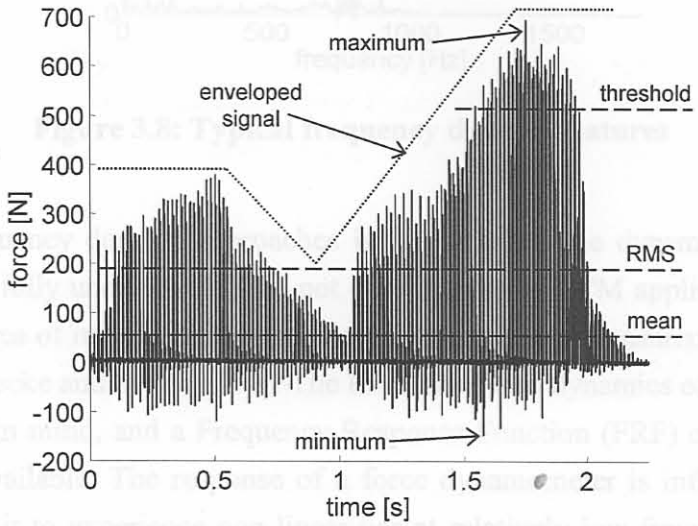


Figure 3.7: Typical time domain features

B. Frequency domain

The most common frequency domain features are the power in certain frequency bands. It is often difficult to identify spectral bands that are sensitive to tool wear. It is even more difficult to determine exactly why these frequencies are influenced by tool wear. Power in certain bands will generally increase due to higher excitation forces because of the increase in friction when the tool wears. Sometimes a peak in the Fast Fourier Transform (FFT) will also shift due to changing process dynamics when the tool wears. An early frequency domain approach is reported by Jiang *et al.* [110], in which a frequency band energy is determined from the Power Spectral Density (PSD) function as a feature for tool wear monitoring.

Some authors suggest that two frequency ranges must be identified from the original signal [38]. The one range must be sensitive to tool wear, the other must be insensitive. For instance, if the measurement was made from 0 – 8000 Hz, it must be split (using appropriate filters) into a 0 – 4000 Hz signal, and a 4000 – 8000 Hz signal. If the lower range is more sensitive to tool wear, a ratio between the two ranges can be calculated. If this ratio exceeds a certain pre-established value, it can be deducted that the end of the tool life has been reached. This can also apply for a ratio between the signals recorded from a fresh tool to that compared with a worn tool. Typical frequency domain features are shown in Figure 3.8.

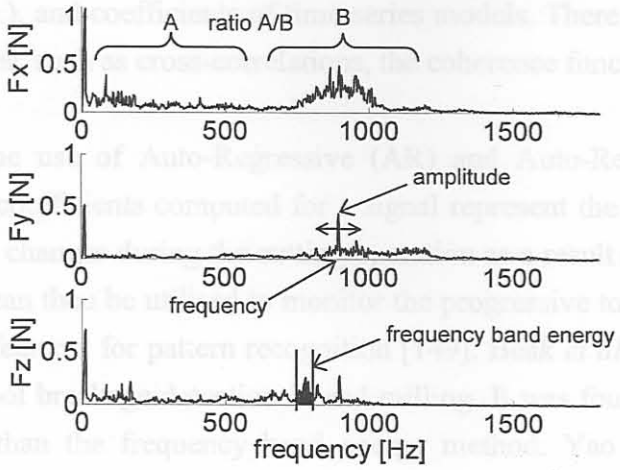


Figure 3.8: Typical frequency domain features

One problem with frequency domain approaches is the fact that the dynamics of the measurement hardware is not always fully understood. This not only applies to TCM applications but also to other research topics in the area of machining that requires dynamic force measurements. This problem was also identified by Warnecke and Siems [151]. The limitations and dynamics of measurement hardware should always be kept in mind, and a Frequency Response Function (FRF) of the installed hardware should preferably be available. The response of a force dynamometer is influenced by its clamping condition, which cause it to experience non-linearities at relatively low frequencies. There also exist some uncertainties when using these instruments, relating to their calibration and varying parameters. A model for expressing the uncertainties when collecting cutting forces with a dynamometer was proposed by Axinte *et al.* [152]. The identified properties might be responsible for the scatter of force components often reported in the literature. An interesting study is also reported by Bahre *et al.* [112] to determine the natural frequencies of the machine tool components using the FEM. These are taken into account for interpretation of the vibration / AE signal.

Choi and Kim [153] describe the use of the spectral energy from the PSD for both vibration and force to identify different stages of wear in diamond tools. Lee *et al.* [154] investigated the correlation between the dynamic cutting force and tool wear. Two important frequencies are identified: The 1st natural frequency of the tool holder and the frequency of chip formation. Normally, the tool holder natural frequency will dominate the results of a dynamic analysis of the force signals. A rough estimation of the frequency can be obtained by modelling it as a cantilever beam (refer to Chapter 5). This frequency can be used as a feature for TCM and is unrelated to the chip formation frequency. The chip formation frequency can be monitored if process stability problems are encountered.

C. Statistical processing

In the case of statistical features, signals are assumed to have a probabilistic distribution. Hence, the signal is regarded as a random process. Generally, machining processes are non-stationary but are assumed stationary for the short periods during which these features are calculated. Several statistical features have been investigated for TCM and can be applied to machining operations. The main features are those that describe the probability distribution of a random process (variance, standard devia-

tion, skewness, kurtosis *etc.*), and coefficients of time series models. There are also various other miscellaneous statistical features, such as cross-correlations, the coherence function and harmonic mean.

One useful approach is the use of Auto-Regressive (AR) and Auto-Regressive Moving Average (ARMA) coefficients. AR coefficients computed for a signal represent the characteristic behaviour of the signal. When the signal changes during the cutting operation as a result of tool wear, the model coefficients also change and can then be utilised to monitor the progressive tool wear. Hence, AR coefficients can also be used as features for pattern recognition [149]. Beak *et al.* [155] report the use of an 8-th order AR model for tool breakage detection in end milling. It was found that the AR approach is somewhat more accurate than the frequency band energy method. Yao *et al.* [156,157] used the ARMA method to decompose the dynamic cutting force signals and wear sensitive frequencies were identified. This assisted to identify the importance of certain vibration modes with respect to tool wear monitoring.

El-Wardany *et al.* [108] found that the instantaneous Ratio of Absolute Mean Value (RAMV) was useful in eliminating false alarms that occur when monitoring drill wear and breakage in conjunction with kurtosis and cepstrum analysis. They state that the kurtosis value is useful in identifying transients and spontaneous events within vibration signals. Cepstrum analysis is used to identify a series of harmonics or side bands in the power spectrum and to estimate their relative strength. Drill breakage consistently caused a peak at the quefrency corresponding to one spindle revolution. The RAMV was used to trigger the onset of kurtosis and cepstrum analysis.

Li *et al.* [109] found that the coherence function of two crossed accelerations can be used as an easy and effective way to identify tool wear and chatter. They found that with progressive tool wear, the autospectra of the two accelerations and their coherence function increase gradually in magnitude around the first natural frequencies of the cross-bending vibration of the tool shank. As the tool approaches a severe wear stage, the peaks of the coherence function increase to values close to unity. This was also proved in theory by the authors. However, there are two conditions to be fulfilled when using this approach: The first is the careful selection of sensor locations on the tool shank. The second is the high-speed computation required for real-time monitoring on the tool performance, as well as the need for a fast FFT co-processor.

The use of Statistical Process Control (SPC) methods were also reported by some authors. Jun and Suh [158] consider the X-bar and Exponentially Weighted Moving Average (EWMA) for tool breakage detection in milling. Jennings and Drake [159] use statistical quality control charts for TCM. Different statistical parameters are calculated and examples of one-, two- and three-variable control charts are given.

D. Time-frequency domain

Several types of time-frequency domain analyses will be encountered in the literature. The most common time-frequency domain processing method in TCM applications is wavelet analysis. A comprehensive description of the advantages and disadvantages of wavelet analysis for TCM can be found in

[6]. It is often stated that wavelets are used because they provide information about the localisation of an event in the time as well as in the frequency domain. However, the time domain information is either not used or is not important. Furthermore, a localisation of events in the time domain is rarely of importance if the aim of the model is wear estimation. A breakage event, which will have a large local effect in the time domain, can be better detected and reacted upon with pure time domain techniques. Furthermore, wavelets are time variant and the exact contribution of a particular frequency at any given time can never be determined accurately due to Heisenberg's uncertainty principle.

The use of wavelet analysis is reported in several publications, such as [37,160-165]. A combination of wavelet decomposition and NNs is described by Hong *et al.* [161] as well as by Xiaoli *et al.* [162]. Lee and Tarng [160] use the discrete wavelet transform for cutter breakage detection in milling and found that the technique is reliable even under changing machining conditions. Scheffer [164] implemented the approach suggested by Wu and Du [163] and showed how wavelets can be used as a digital filter to enhance the reliability of features obtained from statistical analysis of the time waveform. It was found that statistical processes of certain wavelet packets can yield features that correlate well with tool wear.

An advantage is that feature selection from wavelet packet analysis can be done automatically and does not require a large amount of processing time. Luo [166] recently published results of a TCMS using wavelet analysis of vibration signals. In this case the wavelet is used as a filter to enhance wear sensitive features in the signals. However, the results are not compared with conventional digital filtering. A comparative study was carried out during this research and is discussed in Chapter 6.

Another method of time-frequency analysis rarely found in the literature in the area of TCM is spectrograms. Spectrograms are more conventional time-frequency analysis methods and are very useful to identify stationarity in the dynamic signal. They are also useful for detection of disturbances that may be time-localised in signals. The use of the Choi-Williams time-frequency distribution for TCM during multi-milling is described by Li and Tzeng [167]. Wear sensitive regions on the time-frequency distribution are calculated and used as inputs to a NN for wear classification. Although not applied to TCM, the use of the Choi-Williams time-frequency distribution for machining process monitoring is also described by Gu *et al.* [168]. It was shown that the method could be applied on-line for transient monitoring and diagnosis, for example chatter detection. Several examples of spectrogram analysis will also be found in Chapter 4. Although no features were derived from spectrograms, they should always be included as an exploratory step before further processing can commence.

D. Other

There are also a few other techniques that cannot be categorised as either time or frequency domain. One interesting technique is the use of entropy functions. An example is described by Fu *et al.* [169], where the entropy of the frequency spectrum is calculated. The result is one value that is used for pattern classification of different faults that may occur during machining. It is stated that the advantage is the entropy function's insensitivity towards new geometries of cutting. This might however be more related to the characteristic of the FFT than the entropy function! Chungchoo and Saini [170] use the total energy and total entropy of force signals in the frequency domain for TCM. It is stated that the

entropy is more related to the distribution of energy in the spectrum, and is relatively insensitive to changing machining parameters. However, the total energy was found to be the only parameter sensitive to progressive tool wear. The entropy is often also used as a time domain feature, and also represents the energy contained within the wave, thus more or less the same as the signal rms.

3.4.3 Feature selection

It is often found in the literature that authors attempt to generate features that are sensitive to tool wear but insensitive to changing machining parameters. The choice might depend on the particular application, but the sensitivity of a feature towards machining conditions is not of utmost importance because machining conditions can be included in a wear model. There are also other techniques for normalising sensor data with respect to machining conditions, for instance the use of a theoretical model [171-174]. This is very useful if the machining conditions change so often that not enough data can be collected for training or calibrating a model. Numerous techniques exist for selecting the most wear sensitive features, or reducing the input feature matrix to a lower dimension. The main techniques (not necessarily often applied in the area of TCM) are:

- Principal Component Analysis (PCA)
- Statistical Overlap Factor (SOF)
- Genetic Algorithm (GA)
- Partial Least Squares (PLS)
- Automatic Relevance Determination (ARD)
- Analysis of Variance (ANOVA)
- Correlation Coefficient
- Simulation error calculations

Al-Habaibeh *et al.* [175] presented a TCMS for a parallel kinematics machine tool for high speed milling of titanium. An interesting approach to feature selection is employed, called Self-Learning Automated Sensors and Signal Processing Selection (ASPS). This approach is based on an on-line self-learning methodology, whereby a certain feature will be selected automatically based on a correlation with tool wear. A linear regression is performed on each feature in the sensory feature matrix to detect the sensitivity of each feature with respect to tool wear. A very interesting cost analysis is then performed to determine if the installation of a sensor justifies the costs involved.

Ruiz *et al.* [58] proposed the use of a discrimination power for feature selection in a TCM application. The method is similar to that of the SOF. An automated version is proposed that also checks for linear correlation between features. It is difficult to assess the success rate of the automated procedure because the experiments / simulations are not described in enough detail. Quan *et al.* [176] reported the use of the correlation coefficient to assist in feature selection. Lee *et al.* [177] describe the use of the ANOVA to determine the best force ratio for TCM statistically. Several ratios between the three main cutting forces are computed and the influence of controllable parameters (*e.g.* machining conditions) on these ratios are investigated by means of the ANOVA technique. The use of ANOVA as well as the correlation coefficient was also reported by Scheffer [164]. Du [33] describes the use of a blackboard

system, which is a knowledge-based approach for feature selection and decision-making. An advantage is the fact that a physical interpretation of feature can be linked to phenomena in the machining operation. The method is also flexible, but suffers from the disadvantage of requiring a large amount of data and expertise to establish the knowledge-based rules. Some of these techniques can also be automated for a faster implementation. In the opinion of the author, engineering judgement plays a vital role in feature selection for TCM. Some of the techniques and their role in feature selection will become apparent in the chapters that follow.

3.4.4 Wear model / Decision making

A. Time domain signature

The techniques used in commercial systems are often based on the time domain history (also called the 'part signature' in industry). If the time domain history of a vibration sensor yields values outside the limits from a reference cut, a decision is made with respect to the condition of the tool. Two methods are used, namely static and dynamic limits. Examples of these methods can found in Appendix B.

B. Trending, threshold

Instead of investigating the complete time domain signal, a very simple decision making technique can be based on trending features derived from the signals. When a certain feature, or a set of features, reach certain pre-established set limits, an estimation of the tool condition is made. Threshold values for the features can be established that can be related to a certain tool condition. Unfortunately, these thresholds can only be determined through experiments, and problems are encountered under diverse cutting conditions. Furthermore, the features typically exhibit high variance due to disturbance and consequently cause false alarms. An example of trending and thresholds in shown in Figure 3.9.

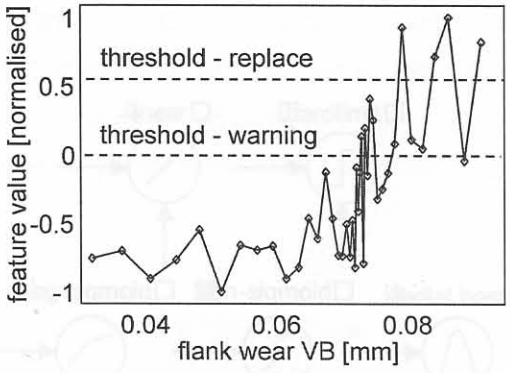


Figure 3.9: Thresholds

C. Neural networks

The use of NNs as a secondary, more sophisticated signal processing and decision making technique have been investigated by many authors in various areas of manufacturing. This is also very true for TCM [171-195]. A NN is usually used to model the input-output relationship between signal features and tool wear. Due to the many complexities involved, NN modelling is ideal for TCM problems because it utilises a matrix of independent data simultaneously to make a classification. The extraction of underlying information and the robustness towards distorted sensor signals are two of the most attrac-

tive characteristics of NNs.

This also applies to sensor fusion schemes for TCM. Combining features from the vibration, AE, force and current signals results in a model that can predict the tool condition with improved accuracy [56]. The successful implementation of NNs is dependent on the proper selection of the network structure, as well as the availability of reliable training data. It is also important to make a distinction between supervised and unsupervised network paradigms. Unsupervised NNs are trained with input data only and are usually used for discrete classification of different stages of tool wear. Supervised NNs are trained with input and output data and these are used for a continuous estimation of tool wear.

Because NNs form an integral part of this study, some comments on their formulation are necessary here. A simple single neuron is shown in Figure 3.10 [196]. In this case, p is a scalar.

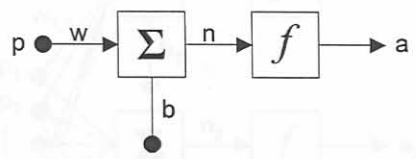


Figure 3.10: Single neuron with bias

The value for a is determined by:

$$a = f(wp + b) \tag{3.1}$$

where w is referred to as the weight value and b the bias value of the neuron. The function f is called the activation function and many different activation functions for NNs exist. The most popular are the linear, hardlimit, log-sigmoid, tan-sigmoid and radial basis function. Examples of activation functions are shown in Figure 3.11.

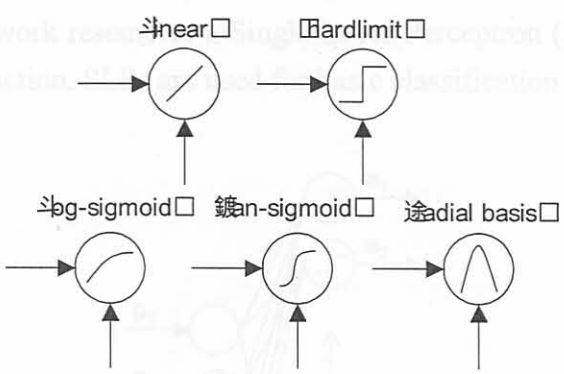


Figure 3.11: Common activation functions

The input to the neuron is normally a vector, and would then resemble the layout in Figure 3.12. The weight would then also be a vector, and will be multiplied with the vector input. The neuron could now be trained to reach a required value for a resulting from the input vector $\mathbf{p}$. Adjusting the weight and bias values with an unconstrained optimisation algorithm until the target is reached, will achieve neuron training. Depending on the type of activation function, different optimisation algorithms are used.

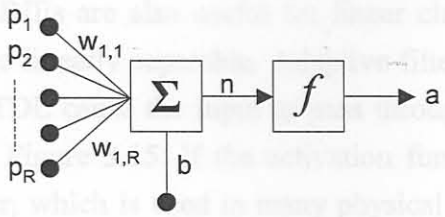


Figure 3.12: Vector input to a neuron [196]

For most NN modelling applications, more than one neuron are required to achieve proper training. An example of a layer of neurons is shown Figure 3.13.

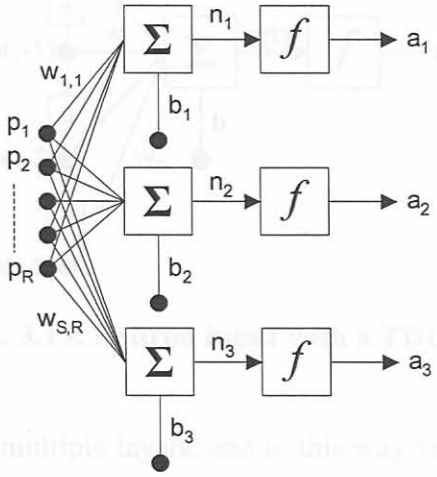


Figure 3.13: A layer of neurons [196]

For simplicity, it is easier to refer to schematic figures representing a layer of neurons such as the illustration in Figure 3.14. The dotted arrow depicts the layer of neurons that may consist of a number of neurons. In this case, the network resembles a Single Layer Perceptron (SLP) network, due to the use of the hardlimit activation function. SLPs are used for basic classification problems.

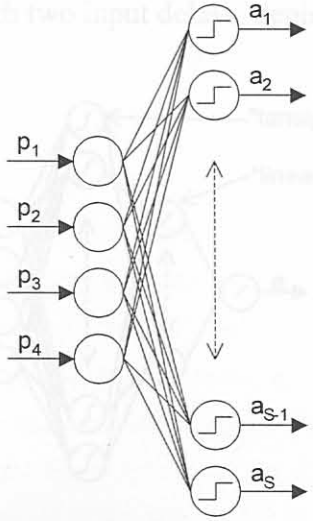


Figure 3.14: A layer of neurons with "hardlimit" activation functions

Adaptive Linear Neuron Networks (ADELINEs) are similar to perceptron networks, but these use the linear activation function. ADELINEs are also useful for linear classification, which means that the classification information must be linearly separable. Adaptive filtering can be achieved by adding a Tapped Delay Line (TDL). The TDL cause the input to pass through a number of delays before it is entered to the neuron, shown in Figure 3.15. If the activation function in this example is linear, it represents an adaptive linear filter, which is used in many physical applications. The TDL can be included in any network type, and such a network is called a Time Delay Neural Network (TDNN). The function of a TDNN is to model a time series resulting from the inclusion of temporal (time) information.

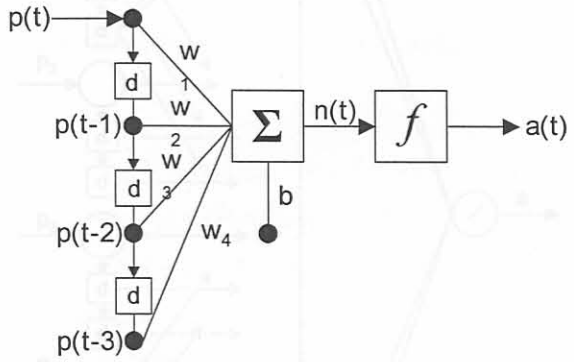


Figure 3.15: Neuron input with a TDL [196]

Neurons can also be combined in multiple layers, and in this way very complex non-linear models can be created. These can be either Multilayer Perceptron (MLP) networks or Multilayer Feedforward (FF) networks. An example of a network with multiple layers is shown in Figure 3.16. Normally, a non-linear activation function should be used in the first layer and a linear neuron in the subsequent layers. In the case of the FF networks, the backpropagation algorithm is used to train the networks. Backpropagation can generally be described as an optimisation algorithm based on steepest gradient descent. The algorithm is quick and efficient, but it is obvious that it can only be used if the gradient of all the activation functions can be determined analytically. If not, for example when using perceptron neurons, other training methods must be used. Delay elements can also be included in multilayer networks, for example the FF network with two input delays, depicted in Figure 3.17.

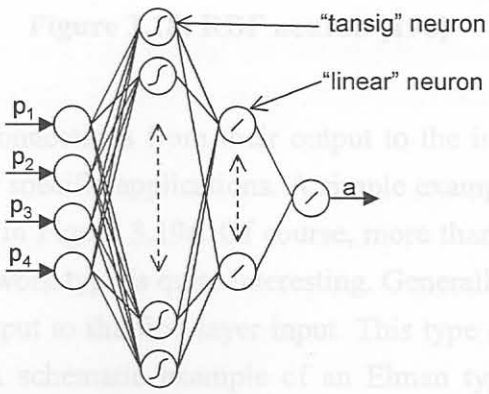


Figure 3.16: Multilayer FF network

There are many variants on the basic NN formulation. One is the Radial Basis Function (RBF) network. These networks may require more neurons but training is much faster. An RBF neuron is depicted in Figure 3.18. Note that the input of the neuron differs from the FF type. In this case, the input is the vector distance between the input vector **p** and the weight vector **w**. The activation function is called a RBF and resembles a normal distribution. The user normally selects the spread of the distribution.

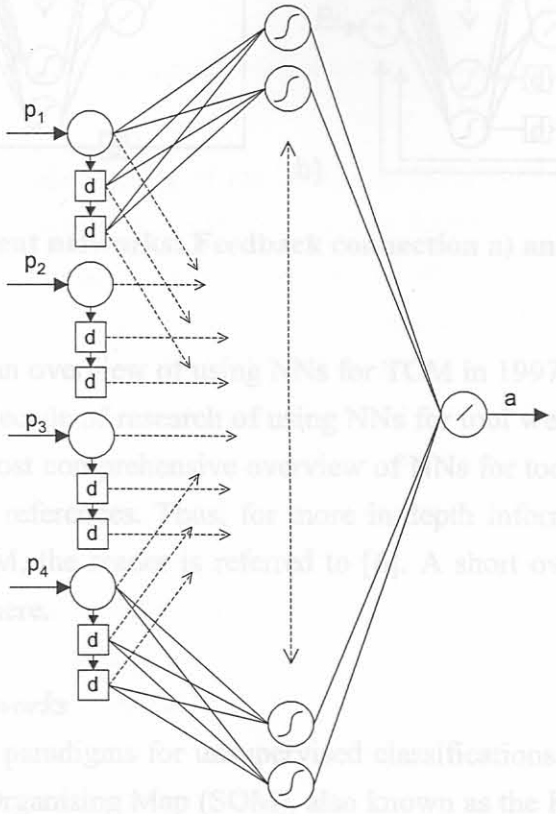


Figure 3.17: Multilayer FF network with TDLs

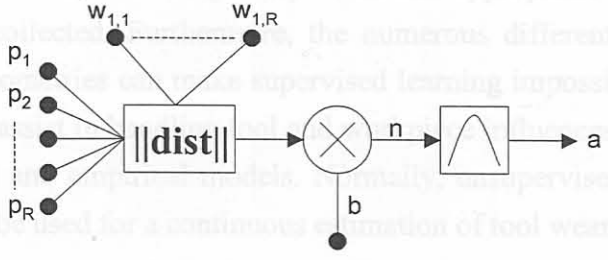


Figure 3.18: RBF neuron [196]

Recurrent NNs have feedback connections from their output to the input. There are various types of recurrent NNs that are useful for specific applications. A simple example of a multilayer network with a feedback connection is shown in Figure 3.19a. Of course, more than one delay or feedback connection can be used. The Elman network type is quite interesting. Generally, it is a two-layer network with feedback from the first layer output to the first layer input. This type of network can be used to learn and model temporal patterns. A schematic example of an Elman type network is shown in Figure 3.19b. For further reading on the theory of NNs, the reader is referred to [196], which also lists many other useful references.

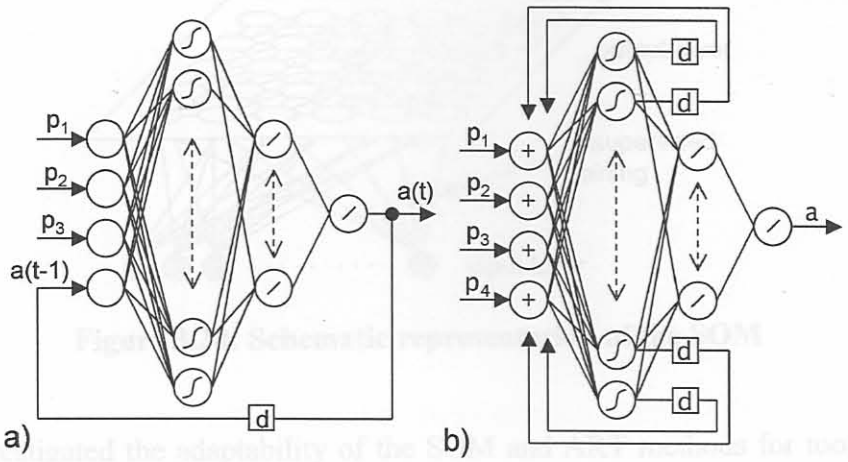


Figure 3.19: Recurrent networks: Feedback connection a) and Elman network b)

Dimla *et al.* [197] presented an overview of using NNs for TCM in 1997. Recently, Sick [6] presented an overview of more than a decade of research of using NNs for tool wear monitoring in turning. This exhaustive overview is the most comprehensive overview of NNs for tool wear monitoring up to date, and includes more than 200 references. Thus, for more in-depth information regarding previous research in using NNs for TCM, the reader is referred to [6]. A short overview of different NN paradigms for TCM is presented here.

C.1 Unsupervised networks

There are two basic network paradigms for unsupervised classifications, namely Adaptive Resonance Theory (ART) and the Self-Organising Map (SOM), also known as the Kohonen Feature Map (KFM). There are many practical advantages for using unsupervised networks. One is the fact that the machining operation is not interrupted for wear measurements. There is also the advantage of practical implementation if machining conditions change very often and appropriate training samples for supervised learning cannot be collected. Furthermore, the numerous different combinations of tool and workpiece materials and geometries can make supervised learning impossible. It should be mentioned that other methods exist to assist in handling tool and workpiece influences but these are subject to the disadvantages of analytical and empirical models. Normally, unsupervised NNs are used to identify discrete classes and cannot be used for a continuous estimation of tool wear.

ART is based on competitive learning, addressing the stability-plasticity dilemma (i.e. overfitting versus generalisation) of NNs. The main advantage is its ability to adapt to changing conditions. ART networks also have self-stability and self-organisation capabilities. The SOM is actually a data-mining method used to cluster multi-dimensional data automatically. A high dimensional feature matrix can be displayed on a two-dimensional grid of neurons that are arranged in similar clusters. Clusters for new and worn tools can be formed and these are used for automatic classification of the tool condition. A SOM is schematically depicted in Figure 3.20 (also refer to Appendix H).

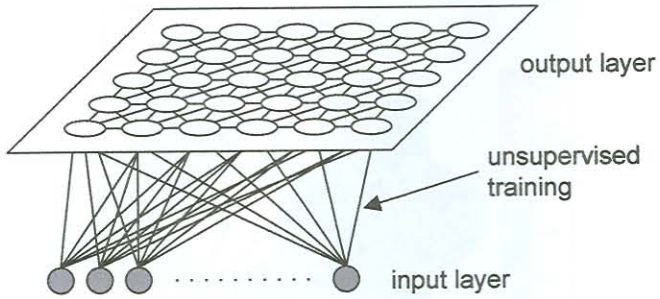


Figure 3.20: Schematic representation of the SOM

Silva *et al.* [9] investigated the adaptability of the SOM and ART methods for tool wear monitoring during turning with changing machining conditions. It was found that with appropriate training the methods have enough adaptive capabilities to be employed in industrial applications. Govekar and Grabec [198] used the SOM for drill wear classification, where the SOM is used as an empirical modeller. It was found that the adaptability of the SOM and its ability to handle noisy data makes the technique feasible for on-line TCM. Jiaa and Dornfeld [199] used the SOM for prediction and detection of tool wear during turning. Scheffer and Heyns [165,200] showed how a TCMS can be adaptable using SOMs. Different network sizes were compared to define discrete classes of new and worn tools. Larger networks yielded more continuous results. The TCMS using SOMs was applied to monitoring synthetic diamond tools for a turning operation in industry, and data mining by using the SOM was also carried out to assist in feature selection. It was found that the SOM can be used for industrial applications, especially if tool wear measurements are not available. Examples from [165] are shown in Figure 3.21 and Figure 3.22. If an accurate value of the tool wear is required, supervised networks can be used, but these will require suitable training data.

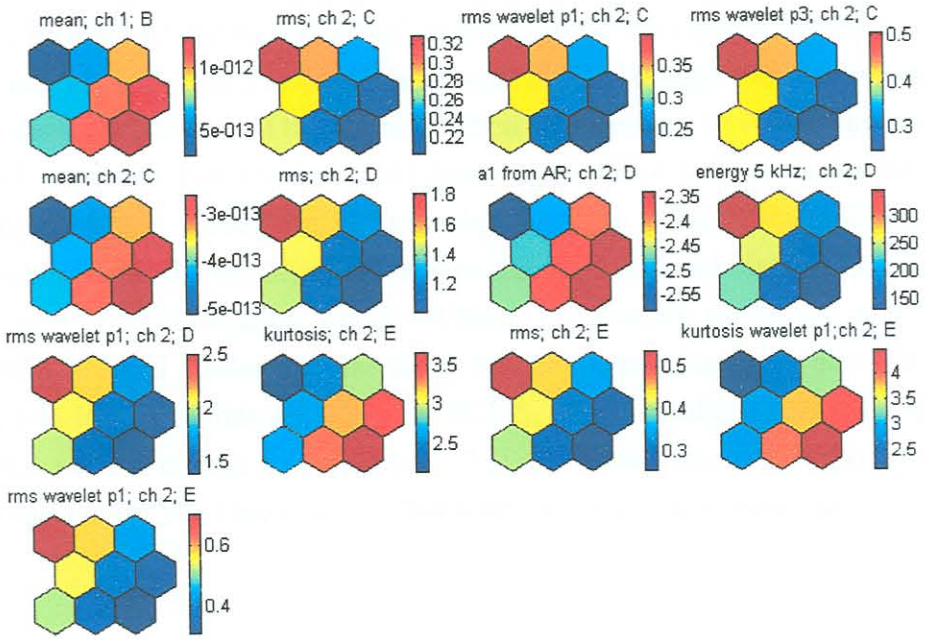


Figure 3.21: Data mining with SOM [165]

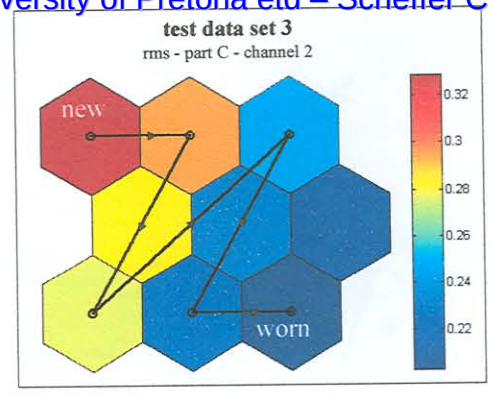


Figure 3.22: Wear classification with SOM [165]

C.2 Supervised networks

The most common supervised NNs for TCM is the Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), Supervised Neuro-Fuzzy System (NFS-S), Time Delay Neural Network (TDNN), Single Layer Perceptron (SLP) and the Radial Basis Function (RBF) network. The use of an SLP for TCM is described by Dimla [178], using the perceptron learning rule. The SLP can only be used to identify discrete classes of the tool condition. MLPs are usually trained with the backpropagation algorithm, for example [101]. However, backpropagation should not always be the preferred choice because other methods are known that outperform this technique in terms of training time and generalisation capabilities. The size of the hidden layer of the MLP network should be optimised or at least investigated for performance [180,187]. Many contradictory statements about the use of MLP networks can be found in the literature. One of the main problems is the choice of the number of input features, size of the network and the number of training samples. In fact, the structure of an MLP network should always be optimised for performance [6].

The use of FF networks with the backpropagation training rule are reported by authors such as Zhou *et al.* [47], Das *et al.* [181,183] and Zawada-Tomkiewicz [193]. The sigmoid is often used as the activation function in the hidden layer and the linear function in the output layer. Cutting conditions can also be included in FF networks. Lou and Lin [201] describe the use of a FF network using a Kalman filter to avoid training problems encountered with backpropagation training for a TCM application. The method is less sensitive to the initialisation values of the weights and biases that often cause convergence problems with backpropagation. Lui and Altintas [189] report on the use of a FF network using a combination of TDLs and feedback connections. Machining conditions are also included. It is stated that the system was integrated into an industrial TCMS, but no results are reported, due to "the availability of robust, practical cutting force sensors..." [189]. It can thus be concluded that the system is not operational in industry yet. However, the NN formulation is quite unique and is depicted in Figure 3.23.

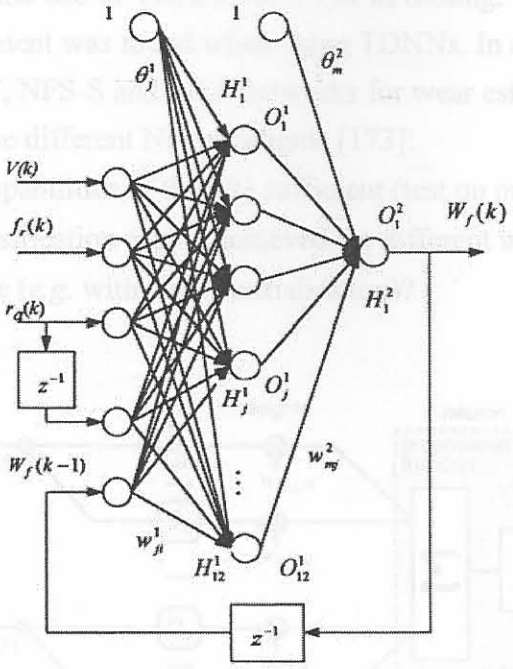


Figure 3.23: NN formulation with TDLs and feedback [189]

Recurrent NNs can be classified as dynamic systems due to the use of feedback connections. Generally, monitoring a dynamic system such as cutting processes should be done through a dynamic modelling technique such as dynamic NN paradigms. Using recurrent networks, or even combining recurrent networks with other NN types can achieve this. Luetzig *et al.* [182] reported the use of recurrent networks for TCM, using a two-layer perceptron in combination with a SOM and RBFs. Ghasempoor *et al.* [8,34,192] reported the use of a non-linear observer technique based on NNs for TCM under variable cutting conditions for estimating two wear modes. It was shown that the technique works quite well for the range of cutting that was considered. One drawback was that the technique was only applied to a laboratory setup and no significant feature generation and selection is employed.

Neuro-Fuzzy Systems (NFS-S) attempts to combine the learning ability of NNs with the interpretation ability of fuzzy logic. A TCMS using NFS-S can be generated almost automatically because the generated fuzzy rules can be learned by the NN. A combination of supervised and unsupervised training is used for NFS-S. An in-process NFS-S system to monitor tool breakage were designed and implemented successfully by Chen and Black [202], concentrating on end milling operations. Xiaoli *et al.* [97] as well as Chunchoo and Saini [195] also propose some of the advantages of using a NFS-S for TCM.

TDNNs are also dynamic systems, for example the formulation shown in Figure 3.24 (also refer to Figure 3.15). One advantage of TDNNs above RNNs is that stability problems are avoided. An investigation towards the inclusion of one and two phase delays for a TCM application was reported by Venkatesh [179]. Different network sizes were also investigated, and it was found that the NNs with temporal memory (time delays) generally perform better than those without memory. It is also stated that new algorithms should be investigated for training (refer to Chapter 4 and Appendix D). Sick and

Sicheneder [174] also describe the use of TDNNs for TCM in turning. The TDNN is compared to the MLP and a significant improvement was found when using TDNNs. In another paper, Sick *et al.* [173] compares the SOM, Fuzzy ART, NFS-S and MLP networks for wear estimation. The following critical questions are used to evaluate the different NN paradigms [173]:

- Are the generalisation capabilities of the NN sufficient (test on previously unseen data)?
- What rate of correct classification can be achieved for different wear stages?
- Are the results repeatable (*e.g.* with a new initialisation)?

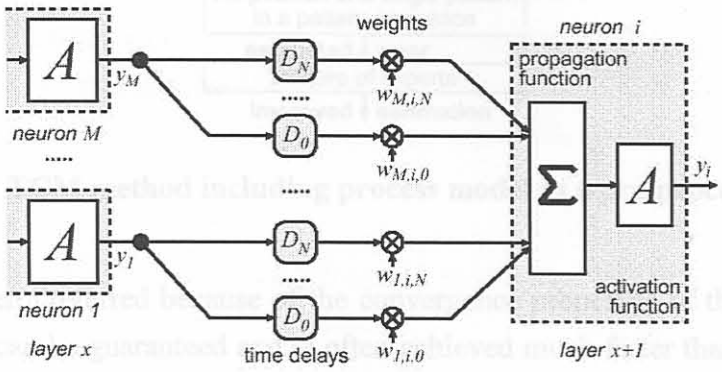


Figure 3.24: TDLs for a TCM application [171]

In the case of [173], MLPs were found to yield the best results. It is stated however that the results can be improved when using TDNNs, which is just a different formulation of MLPs. Such results are reported in [172]. Furthermore, a very novel combined approach is suggested by Sick [171,172] to handle the effect of machining parameters on TCM data. An empirical model is used to normalise the data with respect to machining conditions before the data is entered to the NN. Thus, machining conditions are not included within the NNs. This approach solves the problem of extrapolation of NNs to classify with previously unseen machining parameters. Although many authors test their NNs paradigms in such a way, NNs cannot be expected to extrapolate in this way. The NNs should be tested with previously unseen data under the same conditions (hence interpolate instead of extrapolate – refer [6]). This is a huge problem because training and testing patterns for each condition must be supplied. However, if the data can be normalised with respect to machining conditions, the NN only requires training for the normalised condition. This was in effect achieved by Sick [6], and the method is shown diagrammatically in Figure 3.25. A difficulty still lies with establishing an appropriate model to achieve this, and in many cases it will still require many experiments to develop such a model. However, if an accurate and reliable model is available, the combined approach presents the best solution. The model should preferably be completely analytical to avoid excessive experimentation. An overview of combined techniques for wear monitoring can be found in [203].

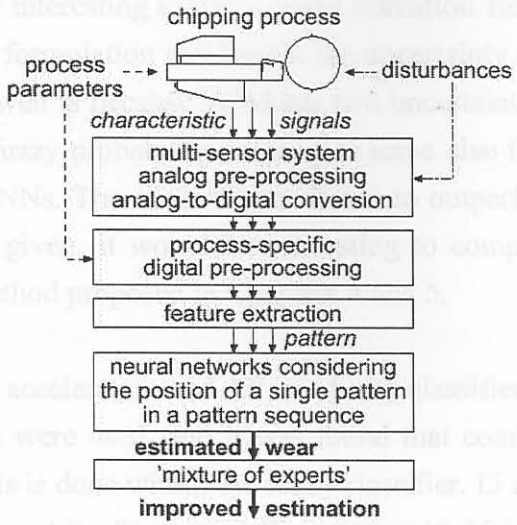


Figure 3.25: TCM method including process model as a pre-processing step [6]

RBF networks are often preferred because of the convergence properties of the training algorithm. In essence, convergence can be guaranteed and is often achieved much faster than with MLPs. However, the accuracy of RBFs depends on the choice of the centres for the basis functions, and should be treated with care. Pai *et al.* [191,194] reported the use of a Resource Allocation Network (RAN) for TCM. The RAN is a RBF network with sequential learning. The RAN is compared to the MLP for wear estimation during face milling. It was found that the RAN has faster learning ability but the MLP is more robust.

In summary, it could be stated that many supervised NN paradigms yield good results for TCM applications, but dynamic paradigms are preferred. Despite a decade of research, an industrial TCMS using the advantages of NNs does not exist. There are a number of possible reasons for this, one being the fact that a laboratory setup differs significantly to an industrial situation. Furthermore, some of the methods and results presented in the literature are not very realistic – for instance the training, validation and testing data sets are not treated properly. The reason for this can probably be contributed to the expense of conducting tool life tests. A cutting test should be repeated at least three times under the same conditions for adequate training, validation and testing. Unfortunately, this is rarely possible. Also, in many cases, the NNs are not subjected to repeatability tests and methods of testing are questionable in some cases.

D. Fuzzy logic

Many authors [204-208] have investigated the use of fuzzy logic to classify tool wear. It has been shown that fuzzy logic systems demonstrate great potential for use in intelligent manufacturing applications. While NN models cannot directly encode structured knowledge, fuzzy systems can directly encode structured knowledge in a numerical framework. Additionally, fuzzy control systems are capable of estimating functions of systems with only a partial description of the systems' behaviour. This is very difficult to construct by simply using NNs.

Du *et al.* [209] propose a very interesting method called transition fuzzy probability, which was applied to a boring process. This formulation can handle the uncertainty of process conditions. The reason why the method performs well is because TCM has two uncertainties: that of occurrence and that of appearance. The transition fuzzy probability solves this issue also through the use of temporal information, similar to dynamic NNs. The method was shown to outperform a backpropagation NN although only minor details are given. It would be interesting to compare this method with dynamic NNs, such as TDNNs or the method proposed in Chapters 4 and 5.

Fu *et al.* [205] combined force, acceleration and AE in a fuzzy classifier for TCM during milling. Time and frequency domain features were used, and it was found that combining the sensory information achieved the best result, and this is done within the fuzzy classifier. Li and Elbestawi [206], Kuo [190] and Kuo and Cohen [207,208] combine fuzzy modelling steps with NNs at different levels for TCM.

E. Other methods

There are also a number of other decision-making and modelling methods that have been applied to TCM, and these include:

- Knowledge Based Expert Systems (KBES) [33,39].
- Pattern recognition algorithms [149].
- Dempster-Shafer theory of evidence [210].
- Hidden Markov Models [211,212].

Of these four approaches, only Hidden Markov Models have the potential to possibly outperform NNs and fuzzy systems. However, not enough comparable research has been conducted in this area and is certainly a worthwhile topic for future research.

3.5 Conclusion

In this chapter, the most important issues regarding sensor-based tool wear monitoring were discussed. The advantages and disadvantages of various sensor systems were discussed with relevance to TCM. Furthermore, an in-depth investigation of different signal processing methods for TCM were given and these will be encountered in further chapters. The formulation of NNs and different NN paradigms were discussed in detail and are especially relevant with respect to this research. From the overviews in Chapters 2 and 3 it can be concluded that sensor-based monitoring using an AI modelling scheme such as NNs is the only way to achieve reliable and accurate TCM. Other approaches cannot achieve the objectives of effective TCM stated in Chapter 1. In the remaining chapters, the focus is on the development of a industrial TCMS using the best techniques available in a unique way.

4. Using AI for tool wear monitoring during hard turning

4.1 Preface

This chapter describes the development of a tool wear monitoring method for hard turning. Hard turning is used in the manufacturing industry as an economic alternative to grinding. Unfortunately, the reliability of hard turning processes are often unpredictable. One of the main factors affecting the reliability of hard turning is tool wear. Conventional wear monitoring systems for turning operations cannot be used for monitoring hard turning tools because a conglomeration of phenomena like chip formation, tool wear and surface finish during hard turning exhibit unique behaviour compared to regular turning operations. During this research, various aspects connected to hard turning were studied with the aim to design an accurate and reliable tool wear monitoring system for hard turning. From this experience it was clear that the best method to monitor tool wear during hard turning is by means of an Artificial Intelligence (AI) approach.

4.1.1 Introduction to hard turning

Accurate monitoring of the hard turning process is essential to ensure part quality. Experience from industry and various research papers have shown that reliable tool wear monitoring during hard turning processes is even more difficult to achieve than wear monitoring during regular turning operations. Influences of changing machining parameters and disturbances on the process are unpredictable. It is the aim of this chapter to present a new methodology for monitoring tool wear during hard turning. To attain a deeper understanding of the hard turning process, some general observations are necessary: It is known that the tool flank wear in hard turning has a major influence on the surface- and subsurface quality of the workpiece, but crater wear governs the reliability of the operation. Crater wear eventually leads to fracture of the tool. Diffusion and chemical reactions are responsible for the crater wear in hard turning. Major wear on the tool flank can also cause formation of the so-called white layer. White layers are the result of microstructural alteration in the workpiece subsurface. It is called the 'white' layer because it appears white under an optical microscope, and has higher hardness than the bulk material [140,213]. Another phenomenon common to hard turning is material side flow, and is caused by tool wear. Material side flow is the flow of workpiece material opposite to the feed direction due to the high temperatures caused by the process. This affects the quality of the finished workpiece. Material side flow is also influenced by cutting speed, and will increase with increasing tool wear [214].

4.1.2 Tool wear with relevance to hard turning

Dawson and Kurfess [215] investigated the wear trends of CBN cutting tools in hard turning. Experiments with different grades of CBN tools were conducted under different cutting conditions. It was found that the specific geometry of the wear on the tool is very important, especially to improve current FEM models of hard turning. It was found that crater wear changes the nominal cutting geometry. It was also found that the flank wear on the same grade of CBN tools under the same conditions differ.

It is unfortunate that experiments were not repeated with the same tools. It was concluded that cutting speed had the most dramatic effect on the tool life, which improved with slower cutting speeds. Significant changes in the cutting geometry were recorded due to flank and crater wear. Zimmermann *et al.* [216] investigated the wear of CBN tools during hard turning using Spectroscopy and X-Ray techniques. In this case, it was concluded that the wear mechanism of the initial crater wear is due to a tribochemical mechanism. The remaining region is then abraded. The wear mechanism for flank wear could not be positively identified.

Chou and Evans [140] also made investigations towards the tool wear mechanisms during hard turning, and found that the carbide sizes of workpieces have a significant effect on fine scale attrition, which appears to have a dominant effect on low CBN content tools. The work was extended to interrupted hard turning and it was shown that tool life is sensitive to cutting speed and rate of interruption [217]. Experiments also revealed that depth of cut has the least influence on surface roughness in hard turning. Generally, it has been found that the thrust force is a good indicator of tool wear in hard turning.

Ozel and Nadgir [218] propose the use of backpropagation NNs for prediction of the flank wear during hard turning. Orthogonal cutting tests were performed using a dynamometer that can collect the torque and the three component cutting force. A force ratio and cutting conditions are included in the input layer; the hidden layer has 30 neurons and the output layer consist of 8 neurons, which is a binary representation of the experimentally measured flank wear. Interestingly, the use of a predictive model is also described to calculate the force ratios for cutting conditions that were not included in the experiment. Good results for flank wear predictions were obtained using the NN approach. The authors want to extend the work to a sensorless approach for on-line TCM, which in the opinion of this author will yield only a tool life equation based on NNs, and will be of limited use.

The scope of this chapter is to develop a monitoring system for hard turning by using a dynamic AI technique, such as dynamic NNs. Passive modelling of manufacturing operations does not meet the requirements for an industrial system. Kothamasu and Huang [219] identified the same problem and developed a dynamic method for hard turning using Neuro-Fuzzy modelling. In this chapter, another method is proposed based on a combination of static and dynamic NNs.

4.2 Experimental setups

4.2.1 Physical configurations

A total of eight experiments were conducted, all with the same basic configuration. However, in order to isolate disturbances and to identify appropriate features for monitoring, some of the experimental conditions were varied. Thus, care was taken to ensure that all experimental conditions remain the same, except parameters that were controlled changes. The basic configuration is shown diagrammatically in Figure 4.1. It consisted of a CNC turning machine equipped with a tool holder dynamometer for 3-component cutting force measurements. In some experiments, Acoustic Emission (AE), 3-component acceleration and temperature measurements were also made.

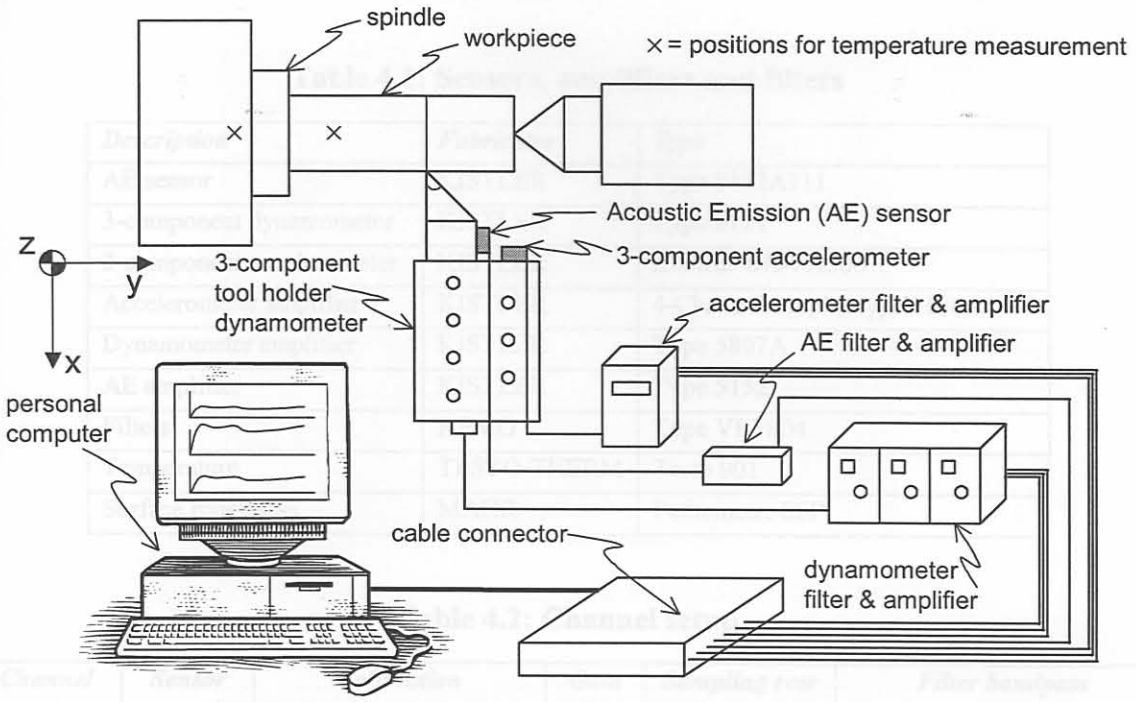


Figure 4.1: Schematic representation of experimental setup

A detailed description of the hardware is given in Table 4.1, with the conditions for each channel in Table 4.2. The variables that were measured during each experiment are summarised in Table 4.3. It is known that the static cutting forces are good indicators of tool wear, but a proper dynamic analysis of the cutting forces can also generate reliable features for wear monitoring. To achieve this, a high sampling rate and analogue low-pass filtering must be used. AE and acceleration also possess properties for wear monitoring during hard turning and these were included during some experiments.

Temperature measurements were made to investigate the effect of fluctuating temperatures on the experimental results. From an industrial point of view, the surface finish of the turned components is the most important feature of the workpiece, and consequently the surface finish of all the workpieces were also measured with a stylus. This enabled an investigation into the effect of tool wear on the surface finish.

In all experiments an analogue trigger was used to start the recordings. In the event where two PCs were used simultaneously the same trigger channel and threshold level were used to ensure alignment of the data. The accelerometer and AE sensor were attached close to the tool tip on the tool holder. The dynamometer was attached on the index turret, and the tool holder containing the CBN insert was clamped in the dynamometer. All sensors are piezo-based, with the dynamometer a special arrangement that can measure both DC and AC signals reliably up to about 1.2 kHz. Many previous studies have found that it is important to mount the AE and vibration sensors as close as possible to the tool tip. In fact, this is more important than the actual mounting orientation, especially in the case of AE.

Table 4.1: Sensors, amplifiers and filters

Description	*Fabricator*	*Type*
AE sensor	KISTLER	Type 8152A111
3-component dynamometer	KISTLER	Type 9121
3-component accelerometer	KISTLER	*K-shear* 8794 A500
Accelerometer amplifier	KISTLER	4-Channel coupler type 5134A1
Dynamometer amplifier	KISTLER	Type 5807A
AE amplifier	KISTLER	Type 5152
Filters	KEMO	Type VBF804
Temperature	TESTO-THERM	Testo 901
Surface roughness	MAHR	Pethometer S8P

Table 4.2: Channel setup

Channel	*Sensor*	*Calibration*	*Gain*	*Sampling rate*	*Filter bandpass*
1	AE_{real}	$57dB_{ref\ 1V/(m/s)}$	10	400 kHz	50 kHz -100 kHz
2	AE_{rms}	$57dB_{ref\ 1V/(m/s)}$	1	20 kHz	50 kHz -100 kHz
3	Fx	50 N/V	1	20 kHz	DC - 3 kHz
4	Fy	50 N/V	1	20 kHz	DC - 3 kHz
5	Fz	50 N/V	1	20 kHz	DC - 3 kHz
6	Ax	10 mV/g	100	20 kHz	10 Hz - 6 kHz
7	Ay	10 mV/g	100	20 kHz	10 Hz - 6 kHz
8	Az	10 mV/g	100	20 kHz	10 Hz - 6 kHz

Table 4.3: Measurements per experiment

	Exp. 1	*Exp. 2*	*Exp. 3*	*Exp. 4*	*Exp. 5*	*Exp. 6*	*Exp. 7*	*Exp. 8*
AE_{real}	✓							
AE_{rms}	✓							
Three cutting forces	✓	✓	✓	✓	✓	✓	✓	✓
Three cutting vibrations		✓	✓	✓				
Surface roughness	✓	✓	✓	✓	✓	✓	✓	✓
Workpiece diameter	✓	✓	✓	✓				
Machine temperature	✓							
Room temperature	✓							
Workpiece temperature	✓							

In the following pages, pictures of the experimental setups are shown. Figure 4.2 is a picture of the machine and the electronic equipment for experiments 1-4. In Figure 4.3 and Figure 4.4 the sensors and workpiece for experiments 1-4 can be seen. The machine used for experiments 4-8 is shown in Figure 4.5. The layout of the workpiece and the dynamometer for experiments 4-8 is depicted in Figure 4.6.

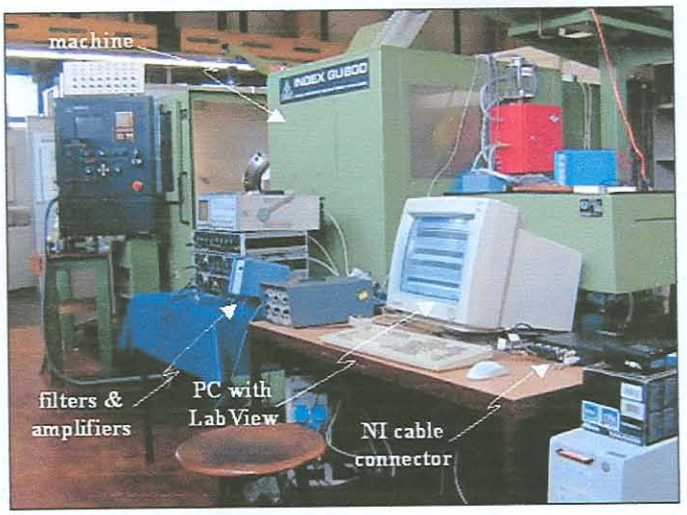


Figure 4.2: Machine and electronic equipment experiment 1-4

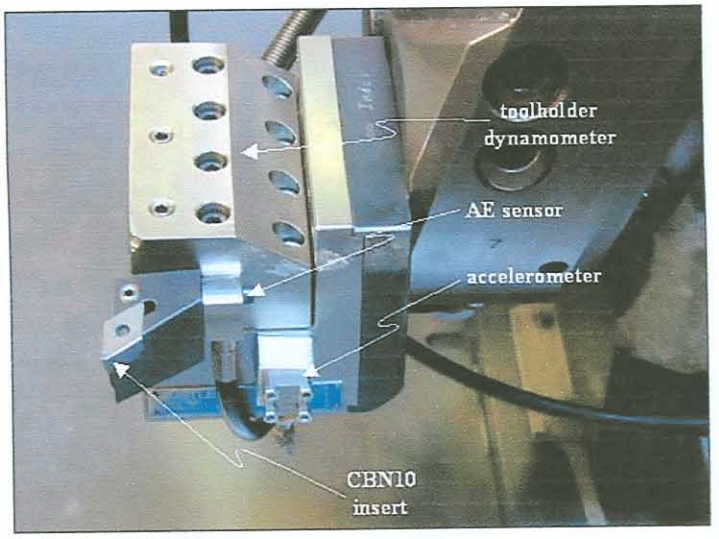


Figure 4.3: Sensors experiment 1-4

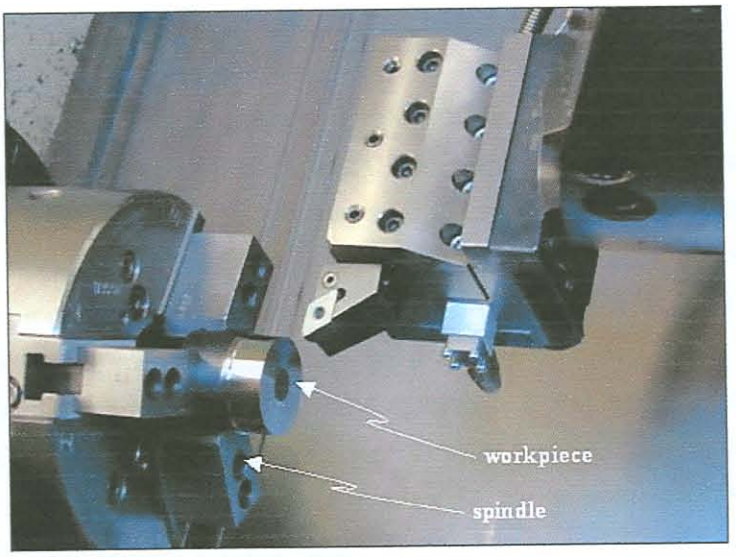


Figure 4.4: Workpiece setup experiment 1-4

Figure 4.5: Machine experiments 4-8

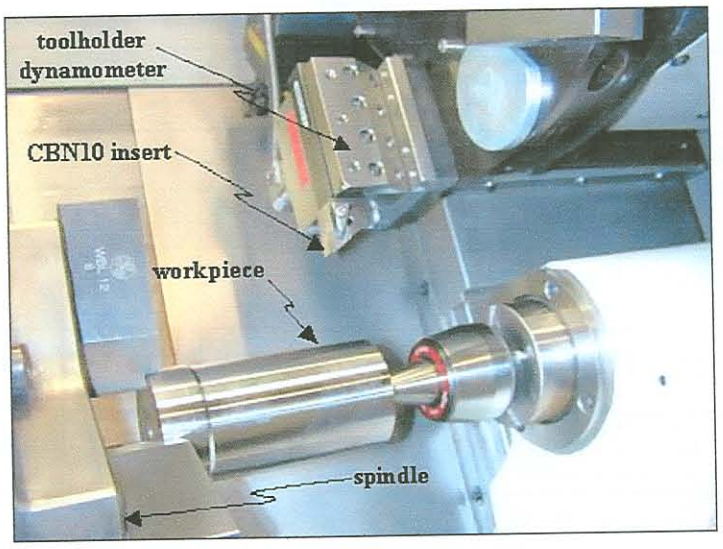


Figure 4.6: Layout of workpiece and dynamometer experiments 4-8

4.2.2 Procedure and process parameters

The experimental procedure consisted of several steps. With each new workpiece, a preparation cut with depth 0.1 mm was made with a preparation tool. This was to ensure that during subsequent runs, a constant depth of cut is maintained. After the preparation cut the workpiece remains in the chuck and the preparation tool is replaced with the tool under test. Also, before each machining test, the temperature of machine, workpiece and room were measured with a temperature probe. Care was taken to measure in the same positions every time. After the runs were completed, the temperature of the machine and workpiece (still clamped) were measured again. Then, the tool was removed for wear measurement under a tool maker's microscope. As a last step the surface roughness of all the workpieces were measured with a stylus. The tempo of the experiments varied somewhat, with some running over several days and others running over a single day.

The machining parameters were selected to resemble industrial practice. The machining parameters for

strategy, experiments were repeated with the same parameters. In essence, the cutting speed and feed rate was varied, and experiments were conducted on two different lathes and workpiece types. Hence, the effect of the machine type, workpiece hardness and geometry can be investigated. The type of tool insert (CBN 10) was not varied, and the depth of cut was kept constant and continuous at 0.1mm for all experiments.

Table 4.4: Experimental parameters

Description	*Exp. 1*	*Exp. 2*	*Exp. 3*	*Exp. 4*	*Exp. 5*	*Exp. 6*	*Exp. 7*	*Exp. 8*
Machine	INDEX GU800				Montforts RNC 400 plus			
Holder	SECO PDJNR/L 2525M11				SECO PDJNER/L 25 25M11			
Insert	SECO DNGA 150612SL1 (CBN 10)							
Depth of cut [mm]	0.1mm							
Speed [m/min]	140						160	
Feed rate [mm/rev]	0.04	0.08				0.16	0.08	0.16
Workpiece	100Cr6				51CrV4			
Passes	210	256	210	216	76			
Total time [min]	74.4	43.0	50.2	50.3	187.0	91.7	157.4	77.1
Total length [km]	10.42	6.01	7.03	7.04	26.18	12.84	25.18	12.34
Diameter [mm]	41.7	40.7	41.8	40.6	72.8	72	71.2	70.4
Pass length [mm]	15		20		120			

4.2.3 Method for tool wear measurement

A combination of flank and crater wear governs the hard turning process. In order to correlate the sensor information with the formation of tool wear, tool wear measurements were made between each new workpiece. The shape of the wear was described with the parameters depicted in Figure 4.7. These are:

VB_{avg} - Average flank wear
VB_{max} - Maximum flank wear
VB_L - Approximate length of the flank wear
K_W - Width of the crater wear
K_D - Depth of the crater wear
K_L - Approximate length of the crater wear

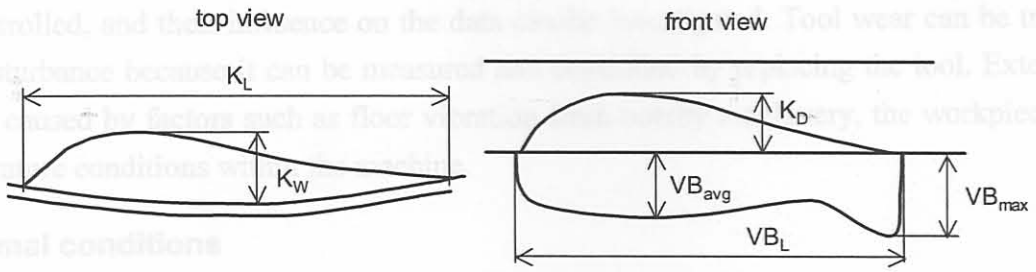


Figure 4.7: Tool wear parameters

For the purpose of the NN implementation, it was decided to attempt to describe to flank wear in terms of an approximate area of flank wear, and the crater wear in terms of an approximate volume of crater wear. Though these are very crude approximations, it is a step in the right direction for more accurate monitoring of tool wear during hard turning. The parameters were formulated as follows:

Area of flank wear:

$$VB_A = VB_{avg} \times VB_L \tag{4.1}$$

Volume of crater wear:

$$K_V = K_W \times K_D \times K_L \tag{4.2}$$

4.2.4 Discussion

The experimental procedures were selected to enable the design of a proper methodology for monitoring tool wear during hard turning that can eventually be treated as a general case. A further reason for the selected experiments was to properly identify and isolate disturbances on the process and experimental setup. The identification of disturbances is a very important issue when monitoring systems for tool wear are concerned. It is known that tool wear will cause a change in sensor signals, but it is extremely difficult to establish whether a change is caused by the tool wear or some other influence on the process (or even something inherent to the experimental setup). Disturbances often have a large influence on the measured signals, whereas the effect of tool wear is relatively small. Several processing methods were used, and are discussed in the following sections.

4.3 Experimental results - disturbances

4.3.1 Introduction

Due to the very small depth of cut, the cutting forces during hard turning are low compared to regular turning. Many studies only consider optimising the hard turning process, such as investigating the lifetime of different tools, analysing the surface and subsurface of hard turned pieces, investigating the influence of interrupted cutting on tool life, *etc.* The purpose of these experiments was to determine which parameters could be used to monitor the tool wear, and identify disturbances on the tool wear data. Methods to treat the effects of disturbances should then be employed.

Disturbances are not only phenomena common to the shop floor but can also be caused by the way in which experiments are conducted in the laboratory. Internal disturbances are caused by known settings inherent to the process, *e.g.* machining parameters. Internal disturbances can be modelled because they

can be controlled, and their influence on the data can be investigated. Tool wear can be treated as an internal disturbance because it can be measured and controlled by replacing the tool. External disturbances are caused by factors such as floor vibration from nearby machinery, the workpiece clamping and temperature conditions within the machine.

4.3.2 Normal conditions

Figure 4.8 shows a typical measurement of cutting forces, with the corresponding Gabor spectrogram (up to 40 Hz) presented in Figure 4.9. It is clear that the forces consist of a static part (DC value) and a dynamic part (AC values). Both the static and dynamic parts of cutting forces are important for TCM [41].

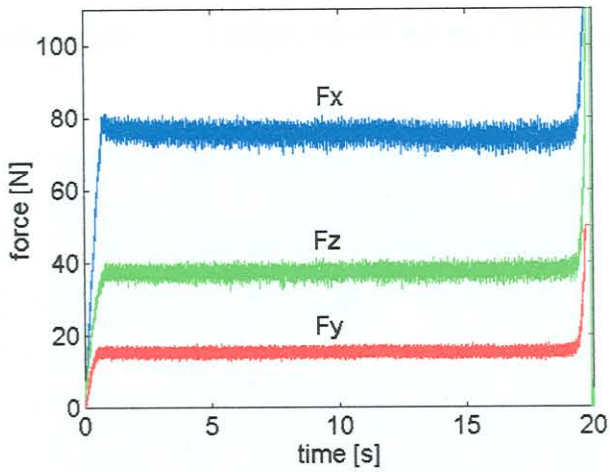


Figure 4.8: Response history of forces under normal conditions

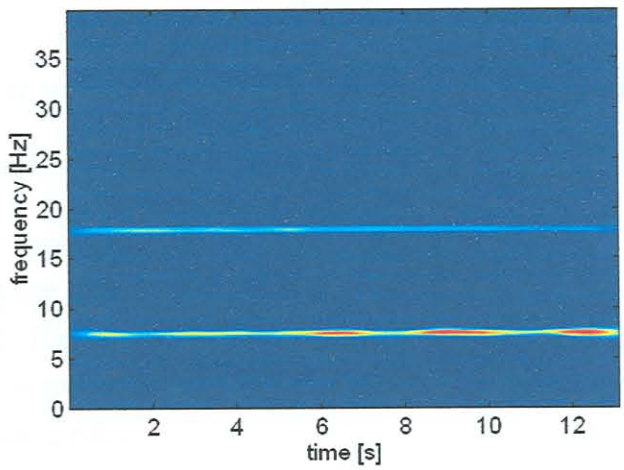


Figure 4.9: Gabor spectrogram of thrust force under normal conditions

4.3.3 External disturbance - clamping condition

In some cases, the workpieces were clamped marginally skew, and this has a big influence on the cutting forces. This influence was taken into account during analysis of the signals. The effect can be monitored by a frequency domain analysis of the cutting force, and was removed from the signal with digital filters. An example of such a measurement is presented in Figure 4.10, with its Gabor spectrogram shown in Figure 4.11.

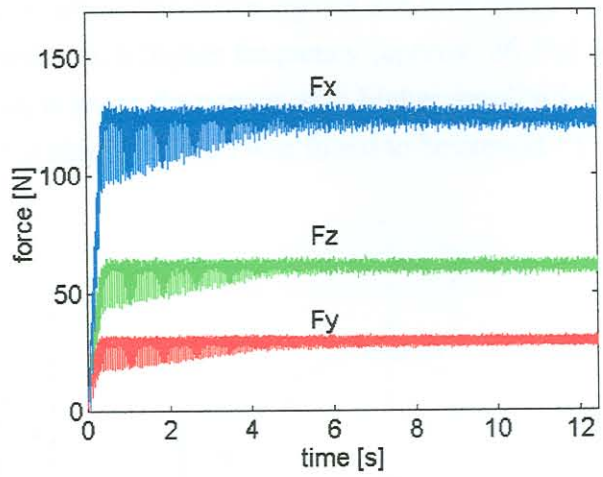


Figure 4.10: Response history of cutting forces when workpiece is clamped skew

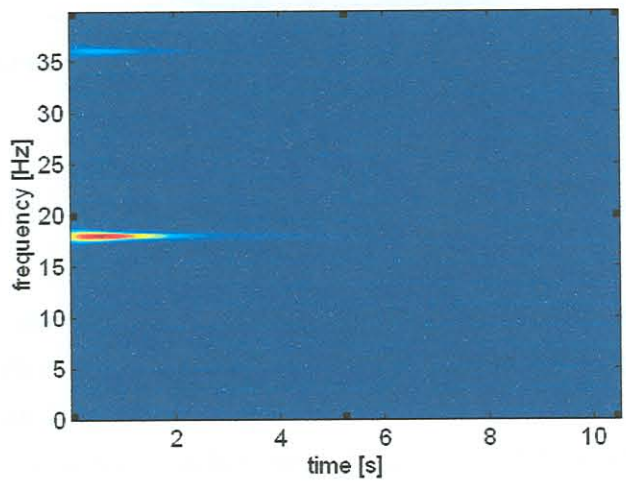


Figure 4.11: Gabor spectrogram of thrust force when workpiece is clamped skew

4.3.4 Internal disturbance - chatter

In some cases, tool chatter appeared due to tool wear effects. Chatter also caused a very irregular chip formation. Upon analysing the data, it was found that chatter cause a high vibrations at about 800 Hz, as depicted in Figure 4.12. The signal number / index in the figure refers to each new measurement. The 800 Hz peak occurred in two experiments on different machines. The effect was removed from the signals by using a digital notch filter.

4.3.5 Internal disturbance – chip formation

During all experiments, the cutting process was observed closely and notes were kept on the chip behaviour. In some cases, a very irregular formation of the chip was observed. Sometimes a very big and irregular curled chip was observed, and in other cases a very smooth behaviour was observed. Sometimes the chips seemed to 'stick' to the tool. Figure 4.13 is a smoothed plot (logarithmic scale) of the thrust force spectrum in the region influenced by chip formation. The behaviour of the chips seems to be totally random, with no connection to tool wear or any other measurable process parameter. However, the instability of chip formation increases with increasing tool wear. This behaviour can also be

seen in the recorded signals, where the force signals are dominated by the frequency related to chip formation. For a very smooth cut, a higher frequency (approx. 26 Hz) with lower amplitudes was observed. For an irregular cut, a lower frequency with higher amplitudes was observed (approx. 8 Hz). Another important region for chip formation was found to be around 15 - 20 Hz.

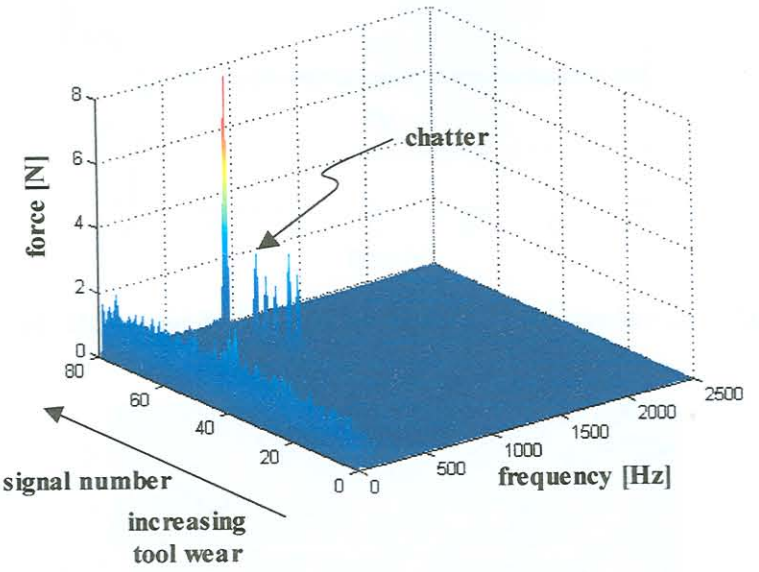


Figure 4.12: Spectra of thrust forces experiment 1

An example of a measurement where the chip suddenly changed from a smooth formation to an irregular formation with big curls is presented in Figure 4.14. A Gabor spectrogram of the thrust force reveals the frequency domain behaviour in Figure 4.15. It is clear that the dominant frequency jumps quite suddenly from 26 Hz to 8 Hz. Another example of cutting forces when the chip curl was visibly big and unstable is plotted in Figure 4.16 with its Gabor spectrogram in Figure 4.17.

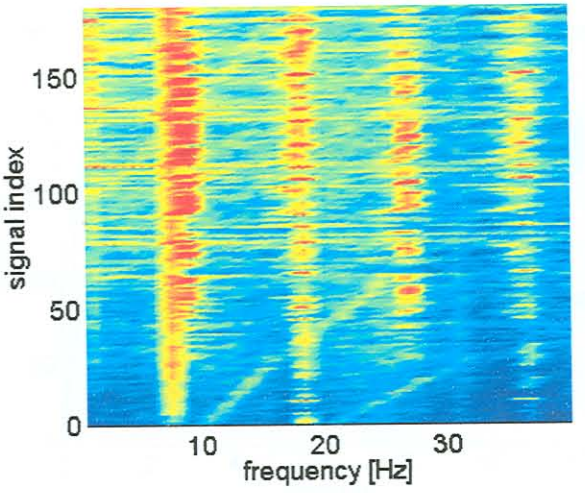


Figure 4.13: Chip formation from forces - thrust force experiment 4

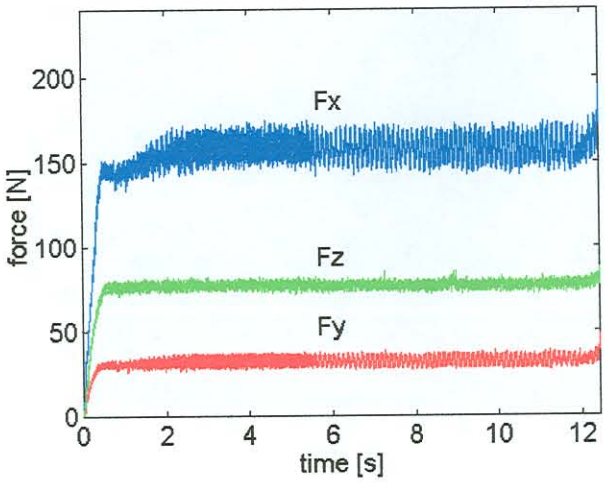


Figure 4.14: Response history of forces during irregular chip formation

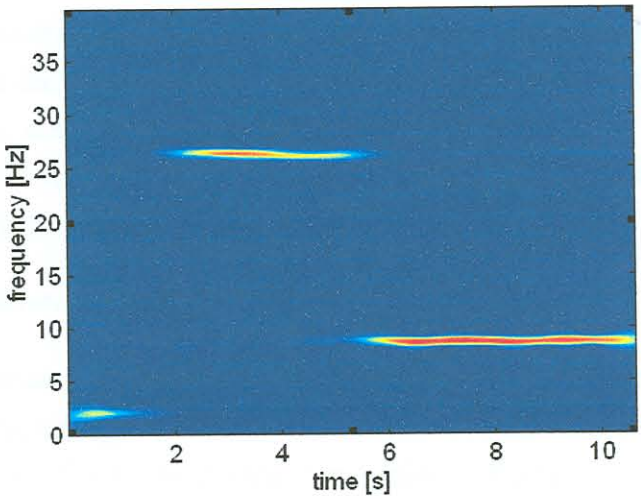


Figure 4.15: Gabor spectrogram of thrust force during irregular chip formation

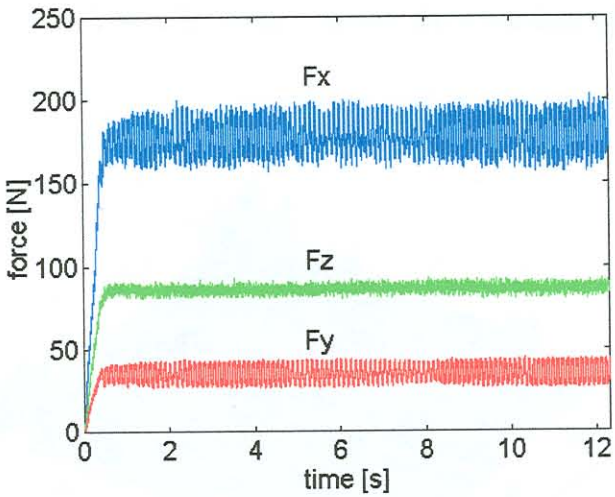


Figure 4.16: Cutting forces with big and unstable chip curl

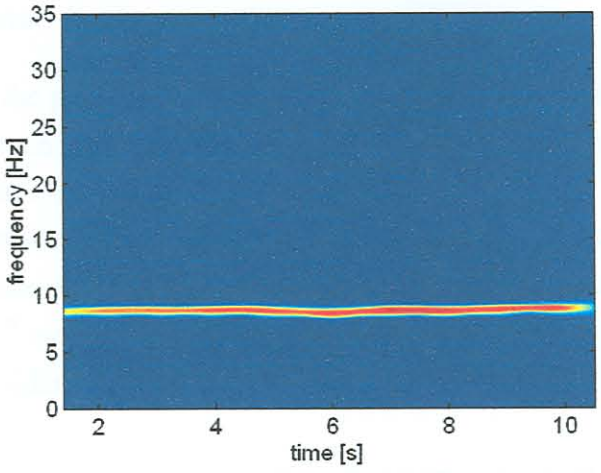


Figure 4.17: Gabor spectrogram of thrust force with big and unstable chip curl

It was concluded that time-frequency analysis can be used to monitor the behaviour of the chip, but cannot be used to enhance the performance of the tool wear monitoring system. No reliable features for wear monitoring could be derived with this technique, and the behaviour of the chip is chaotic with respect to tool wear. It is however important to identify stationary sections in the signals with the aim towards further analysis in the frequency domain (to follow).

4.3.6 External disturbance - electrical noise

Another external disturbance is that caused by the environment. Environmental disturbances are usually due to electromagnetic energy emitted from nearby machinery or even the machine under test. In this case, a 50 Hz (corresponding to the electrical transmission frequency in Germany) peak was found in the force data. The disturbance amplitude is low in the morning and grows as the day progresses. This could be the influence of nearby machinery as more machines are switched on during the day. An example from experiment 2 is shown in Figure 4.18, where data over two days is shown. The disturbance is shown in terms of volts. This effect was removed from the data with digital filtering.

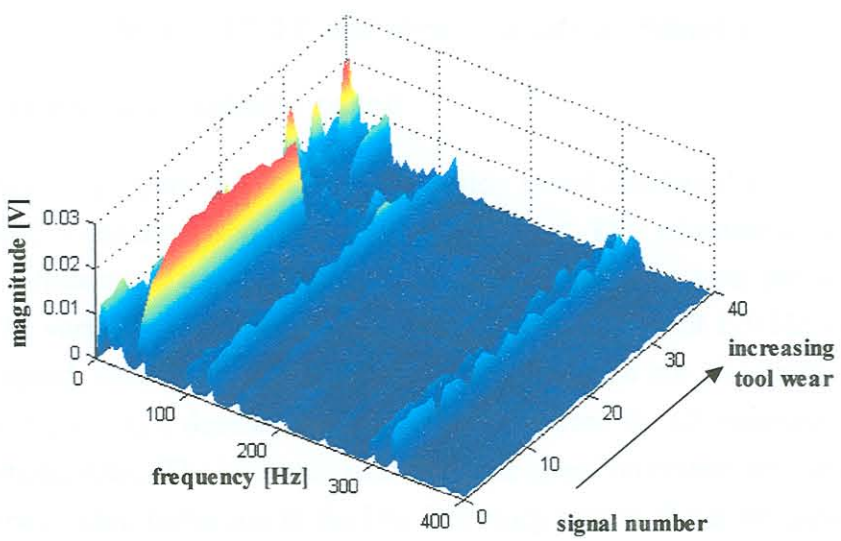


Figure 4.18: Spectrum of feed force data experiment 2

4.3.7 Internal disturbance - tool wear

An inherent disturbance to the process is of course the tool wear. With all other disturbances identified and isolated, it would be much easier to identify parameters influenced by the inherent disturbance of tool wear. As can be seen in Figure 4.19, some regions on the frequency spectrum show a definite increase with progressive tool wear. It was found by this (and previous) investigation that the frequency regions most likely to increase due to tool wear are those corresponding to the tool holder natural frequencies. In this case the 1st mode is near 1 kHz, which is the resonance frequency of the fixed dynamometer clamping a tool holder.

The overhang length of the tool holder is another important consideration when searching for wear sensitive frequencies. In the case of a dynamometer setup, there is a prescribed length (35 mm for the Kistler 9121) prescribed by the sensor calibration specifications. The stiffness, geometry and overhang length of the tool holder determines the values of the natural frequencies. Generally, it can be determined experimentally by an impact (with modal hammer) test, and verified analytically. Some authors also suggest the use of finite element modelling [97].

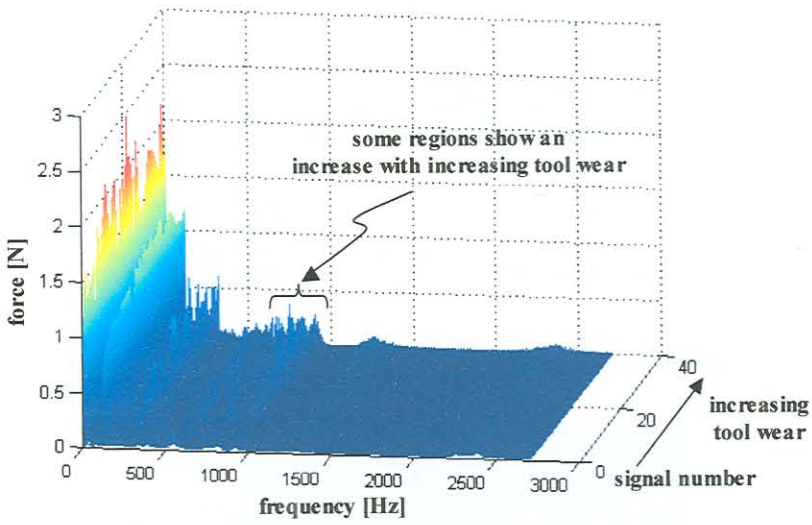


Figure 4.19: Filtered thrust forces experiment 1

4.3.8 External disturbance - spindle speed

Most CNC machines are programmed to keep the cutting speed constant during overruns. This means that the spindle speed is adjusted according to the depth of cut. This adjustment of spindle speed has a minor but notable influence on cutting test data. The results from two tests are shown in Figure 4.20 and Figure 4.21. The variation in cutting speed is clearly visible at about 120 Hz, corresponding to the spindle rotational speed. Other related frequencies in the spectrum are also influenced, but the influence is very small. Figure 4.22 depicts the thrust force spectrum for all measurements during experiment 4 on a logarithmic scale. The influence of the spindle speed on certain frequencies can be seen, as well as the influence of chip formation in the low frequency region. It can be concluded that the spindle speed does not have an influence on chip formation, because the cutting speed remains constant.

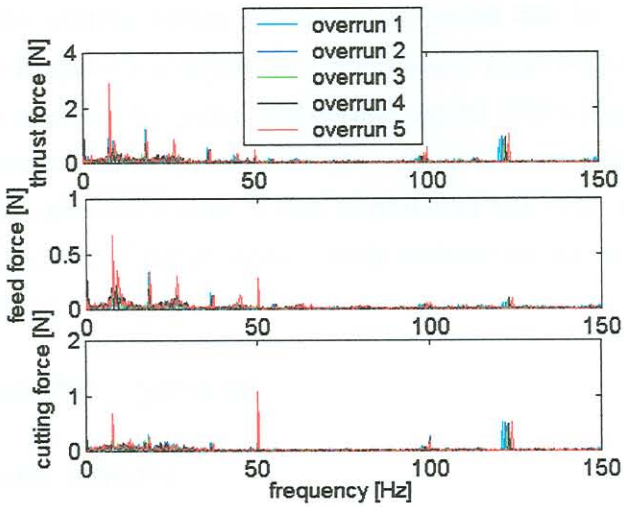


Figure 4.20: Experiment 4 piece 11 forces

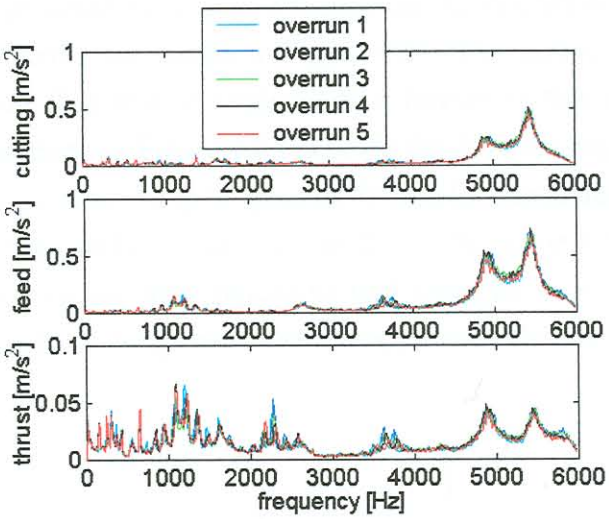


Figure 4.21: Experiment 4 piece 11 acceleration

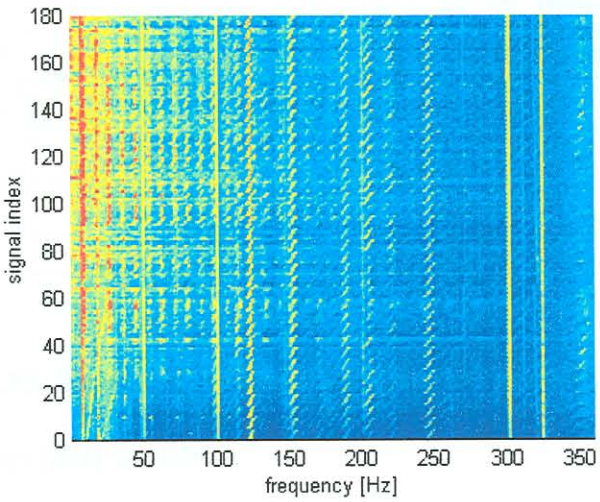


Figure 4.22: Thrust force spectrum experiment 4 (logarithmic scale)

4.3.9 Discussion

An interesting external disturbance proved to be removal of the tool insert for wear measurements. The slightly different clamping situation of the tool insert has an influence on vibration and the static and

dynamic components of the cutting forces. These disturbances can be identified with ease by frequency and / or joint time-frequency analysis of vibration and force signals. Furthermore, their influence on the signals can be removed by using appropriate digital filters. Further investigations revealed that the identified disturbances had no apparent effect on the tool wear growth rate. It should be kept in mind that the analysis results presented here is vital to practical tool wear monitoring. A generic monitoring approach can only succeed if the process is fully understood and the effects of the various disturbances are compensated for.

4.4 Experimental results - general

4.4.1 Temperature measurements

During experiment 1, temperature measurements were taken to determine if external temperatures (those not inherent to the process) have any influence on the operation. The results of temperature measurements from experiment 1 are shown in Figure 4.23. The temperature measurements revealed some interesting correlations that will be explored in a further section using Self Organising Maps (SOMs). The vertical dotted lines on Figure 4.23 indicate the different days over which the experiment was conducted. The start of each new day is apparent from a low temperature that increases as the day progresses. The room and the machine temperatures follow the same trends. The temperature of the workpiece after machining increase with increasing tool wear, as expected. However, investigations showed that the room and machine temperature had no apparent influence on the rate of tool wear.

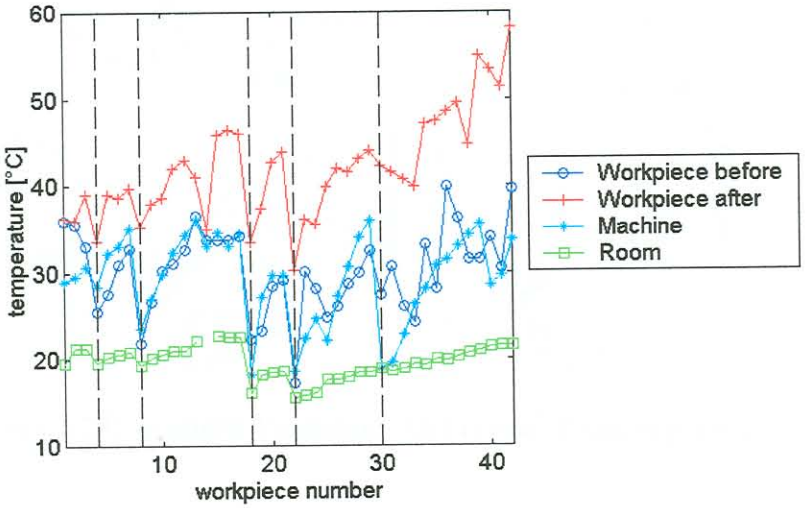


Figure 4.23: Results of temperature measurements experiment 1

4.4.2 Acoustic Emission (AE) results

The results of the AE recordings during experiment 1 are shown in Figure 4.24 (the solid lines are 3rd order polynomial fits of the scattered data). The AE signals were treated with four different analysis methods, all of which eventually displayed the same tendencies. It can be seen that the AE energy starts at a high level, and then decreases to a minimum at about 0.04 mm flank wear. Then it starts to increase again and reaches the highest levels towards the end of the experiment. This result corresponds to that of a very recent study by Tönshoff *et al.* [220]. They provide the following explanation:

"A tool without flank wear generates high AE amplitudes from the sharp-edged tool and the lower damping rate of the tool-workpiece system. With increasing flank wear, the damping rate grows and the generated AE amplitude of the process decreases to a lower bound. If the mechanical and thermal load during machining increases in combination with the appearance of white layers and tempered zones, a continuous increase of the amplitude can be recorded." The same result was also found with different machining parameters.

It can thus be concluded that AE can be used to identify the moment during hard turning where mechanical and thermal loads start to increase. An added result of this case study to that of Tönshoff *et al.* is the fact that the AErms, directly recorded from the rms sensor output without filtering, can also be used to trend the AE energy. The AErms is much easier to record and process compared to the real-time AE signal that requires a sampling rate above 400kHz. Parameters calculated from frequency and time domain analysis did not show significant improvement of the results compared to the AErms. A difficulty with the AE and AErms signal is the fact that it exhibits a lot of variance, and is therefore prone to create false alarms in a production environment. It was decided to omit monitoring of AE during further experiments.

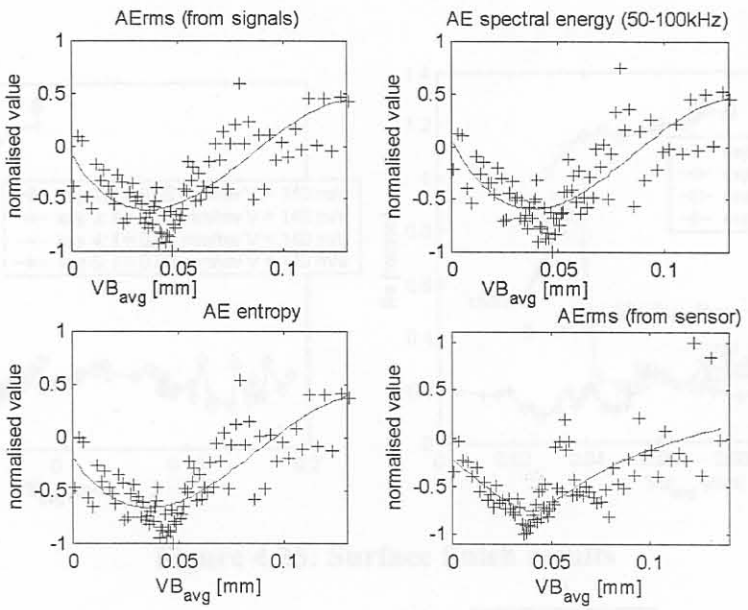


Figure 4.24: Acoustic Emission (AE) results from experiment 1

4.4.3 Surface finish and diameter

Measurements of the surface roughness and diameter of the machined workpieces revealed some interesting facts. It was found that there is not a very clear relationship between the tool wear and the workpiece geometry. The reader is referred to Appendix I for notes on surface roughness measurement and analysis. As general rule the following behaviour is expected:

- Surface roughness will increase initially, then reach a maximum, then decrease suddenly and then start to increase again towards the end of the tool life. This is caused by micro-breakage of the cut-

ting edge after a considerable length of cut. The breakage caused the creation of a new, sharper cutting edge. This process can repeat a few times, after which the edge will deteriorate until complete failure. Thus, more unstable behaviour can be expected towards the end of the tool life.

- The diameter of the workpieces after machining will increase with increasing tool wear, due to the fact that tool material is removed from the cutting edge.

The behaviour of the surface roughness, displayed as the measured Ra values with respect to tool wear in Figure 4.25 (grouped into experiments with the same and different machining parameters), seemed to follow the above-mentioned assumptions. However, the workpiece diameter measurements, shown in Figure 4.26, did not clearly exhibit the expected behaviour. This was due to the fact that the tool was regularly removed from the holder for wear measurements, and it was concluded that this is also responsible for the large variation in the surface roughness values. Although the surface roughness generally increases with increasing tool wear, the relationship is very complex and difficult to generalise. The surface finish of hard turned components can be improved by a secondary process called roller burnishing. However, worn tools will cause weaknesses in the workpiece subsurface and therefore tool wear monitoring is still very important. From the figures it is obvious that it will be extremely difficult to estimate the degree of wear using surface finish information only (and vice versa).

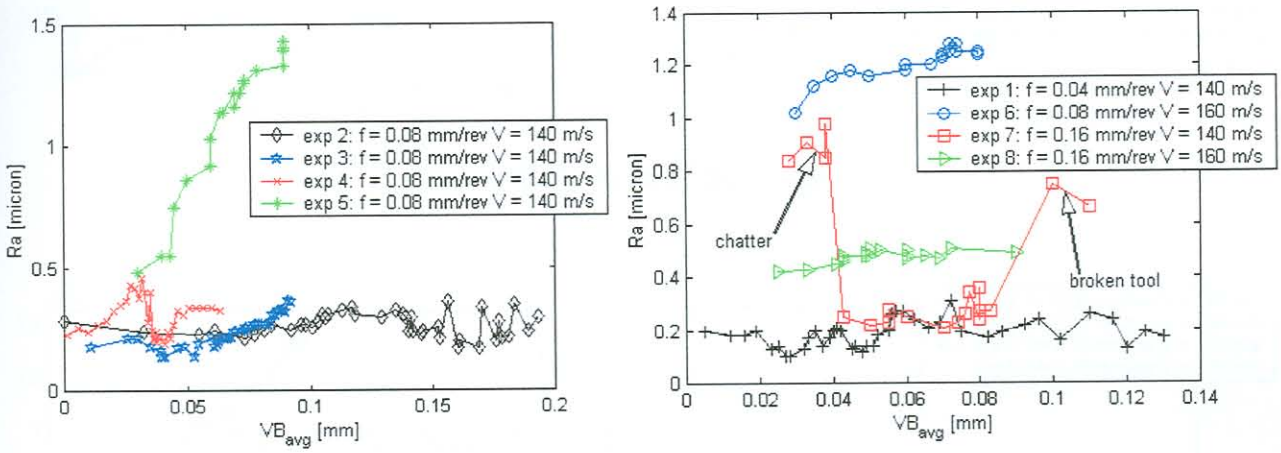


Figure 4.25: Surface finish results

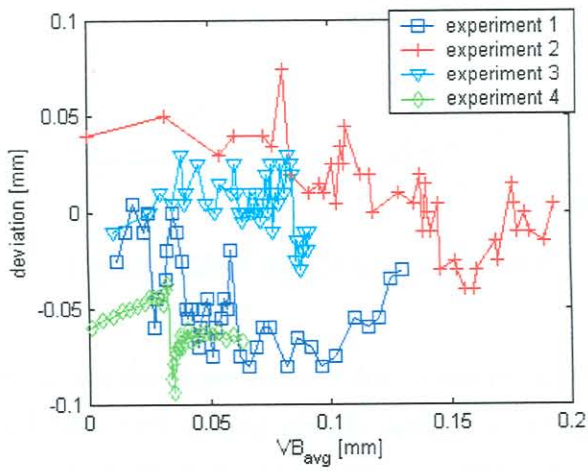


Figure 4.26: Workpiece diameter deviation

4.4.4 Tool wear

The area of the flank wear for the different experiments is plotted in Figure 4.27, with the volume of the crater wear in Figure 4.28. A very interesting fact is revealed in these figures: Experiments 2 and 3 were conducted with the exact same machining parameters, tool type and workpieces, but the wear behaviour is completely different. The only difference between the two experiments is the number of engagement impacts on the tool, which were more during experiment 2. Also, the time delays between overruns were a little longer during experiment 3. Thus, the number of engagements and cool-down time allowed between overruns is an important consideration when considering the life of hard turning tools. A slight difference in quality of the tool coating and workpiece material could also have an influence on the tool life, but the margin of difference is very small.

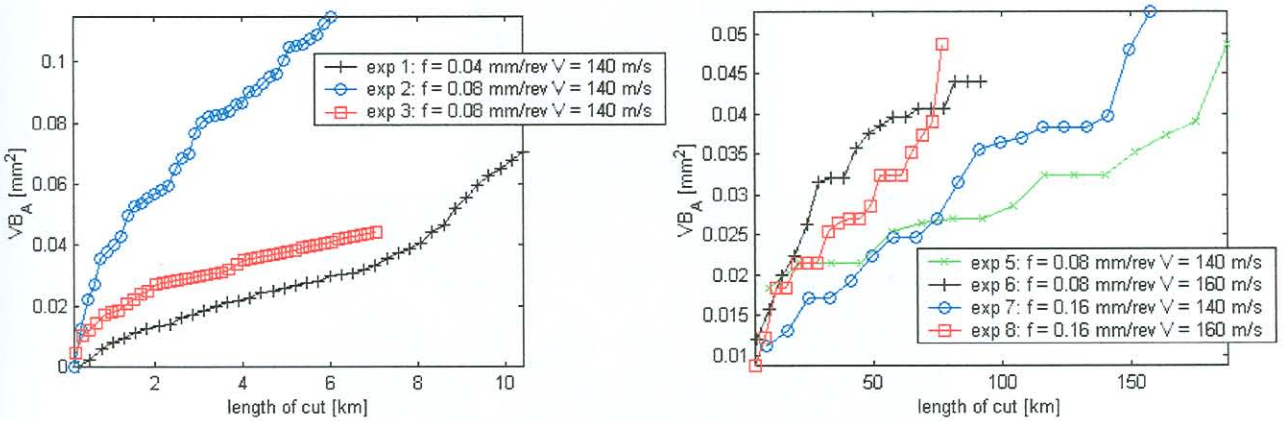


Figure 4.27: Area of flank wear

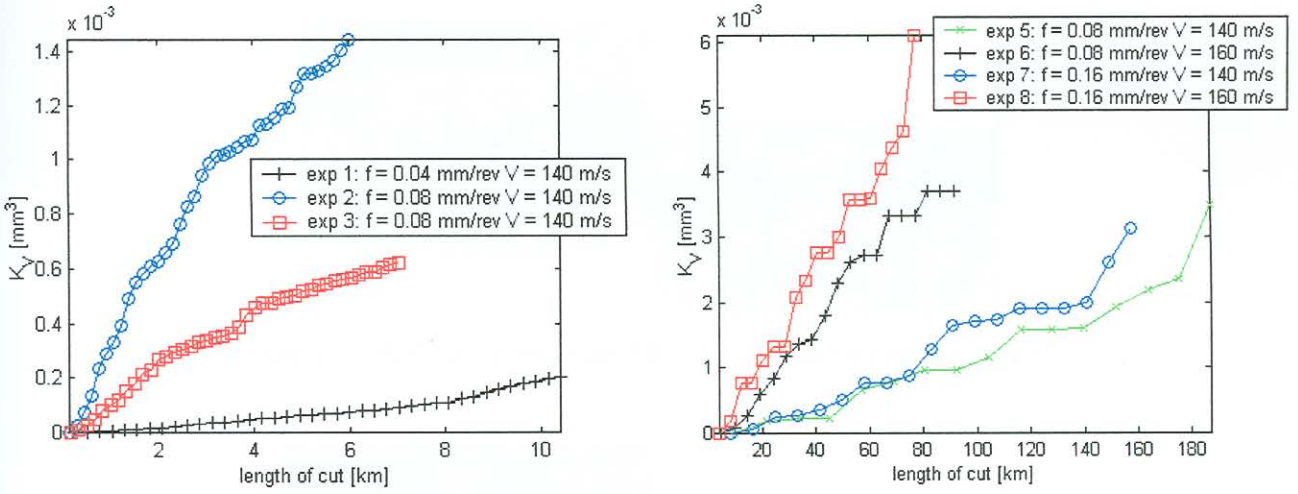


Figure 4.28: Volume of crater wear

The growth of tool wear was slower for experiment 1 due to the lower feed rate. The tool wear is also influenced by the higher feed rates and cutting speeds used in experiments 7-8. The machining parameters for experiment 5 are exactly the same as during experiments 1-4 (besides the workpiece hardness which is slightly different) but the rate of tool wear is much slower.

Scanning Electron Microscope (SEM) pictures of the cutting edges from experiments 1-3 are shown in Figure 4.29. The big differences in the wear geometries of flank and crater wear from the different experiments are clear from these images, especially when comparing the crater wear which was formed under the exact same machining parameters (experiments 2 and 3). Further pictures of the wear on the tools used in experiments 7 and 8 are shown in Figure 4.30 and Figure 4.31, respectively. These pictures were taken with an optical microscope with a digital camera attachment. The geometries of the flank and crater wear in each case can be compared, and it is clear that the particular geometry is unique in each case. Both the flank and crater wear have an influence on the stability of the process. The influence of the wear geometry on the parameters for wear monitoring will be discussed in a following section.

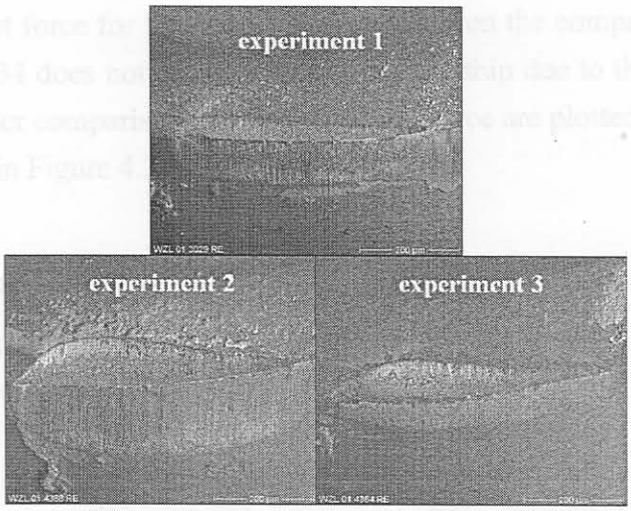


Figure 4.29: SEM pictures of cutting edges

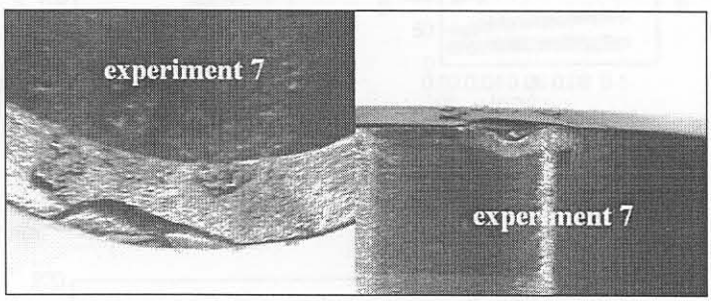


Figure 4.30: Crater wear and flank wear insert experiment 7

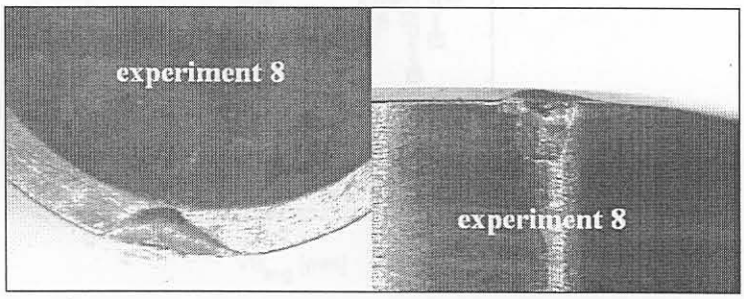


Figure 4.31: Crater wear and flank wear insert experiment 8

4.4.5 Static forces

The static cutting forces are commonly used as tool wear monitoring features. The results of this and previous studies show that the thrust force is the most reliable static force direction to monitor tool wear during hard turning. However, the static parameters alone are not sufficient for a reliable monitoring system. This is due to the fact that they display much variation and are influenced by the disturbances described earlier. The behaviour of the static forces is depicted in Figure 4.32. Apart from changes in machining parameters, there were two major influences on the static forces. The first influence is the removal of the tool for wear measurement. The second major influence seems to be the 3-D geometry of the tool wear itself. When referring to a direct comparison of the static cutting forces, such as the thrust forces in Figure 4.33, it can be seen that there is not a consistent relationship between VBavg and the thrust force for the four experiments. Even the comparison of different machining conditions in Figure 4.34 does not reveal a simple relationship due to the large degree of variance in the measurements. Further comparisons of the static feed force are plotted in Figure 4.35 and Figure 4.36, and the cutting force in Figure 4.37 and Figure 4.38.

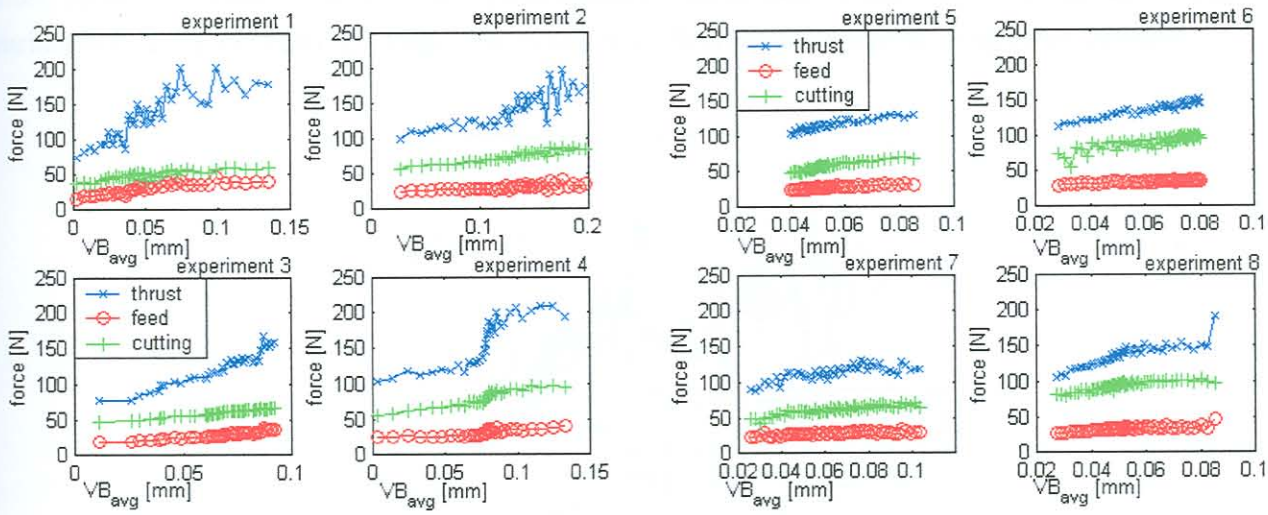


Figure 4.32: Static forces

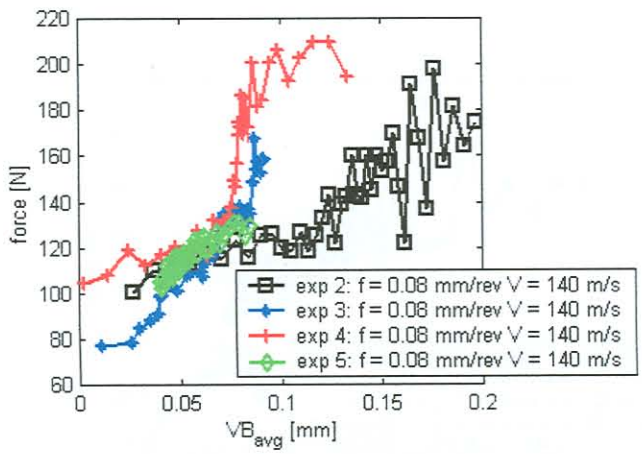


Figure 4.33: Static thrust forces (1)

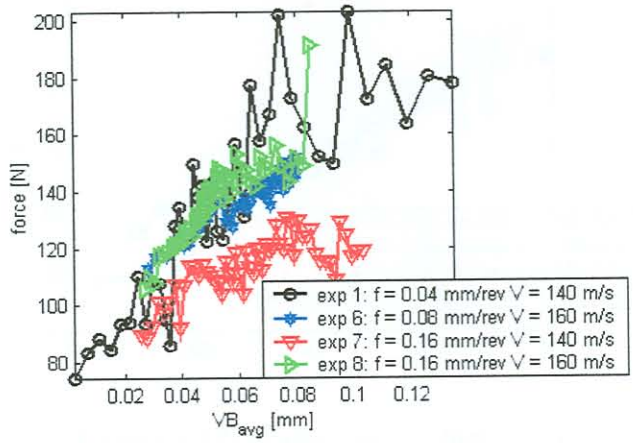


Figure 4.34: Static thrust forces (2)

Due to the fact that all controllable parameters during experiments 2, 3 and 4 remained the same, it can be postulated that the shape of the crater can cause the process to be more unstable and thus cause more random variations in the static (and dynamic, discussed in the following section) forces. The variation in the static cutting forces creates major difficulties from a modelling point of view.

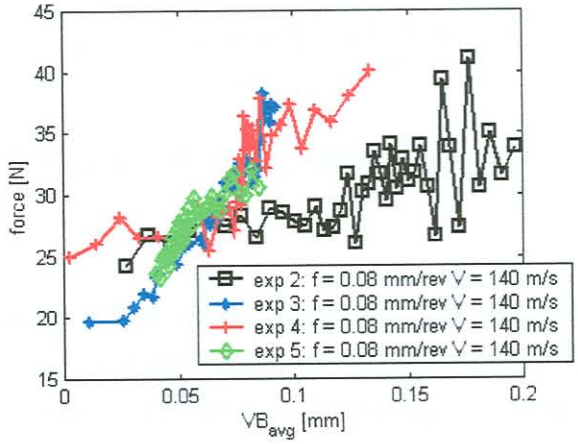


Figure 4.35: Static feed forces (1)

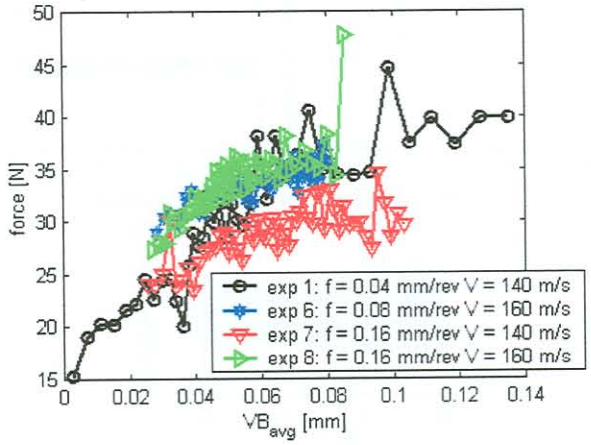


Figure 4.36: Static feed forces (2)

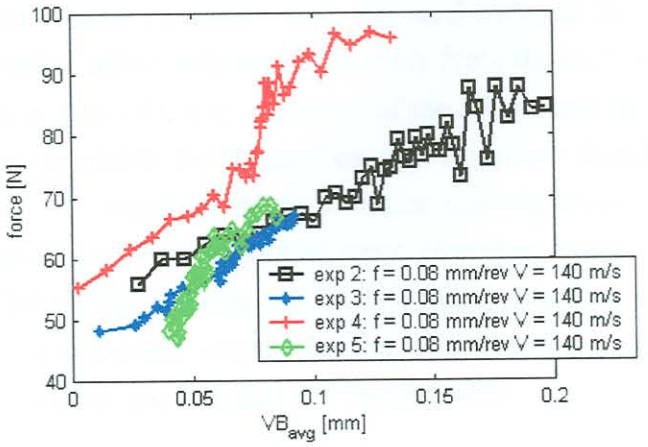


Figure 4.37: Static cutting forces (1)

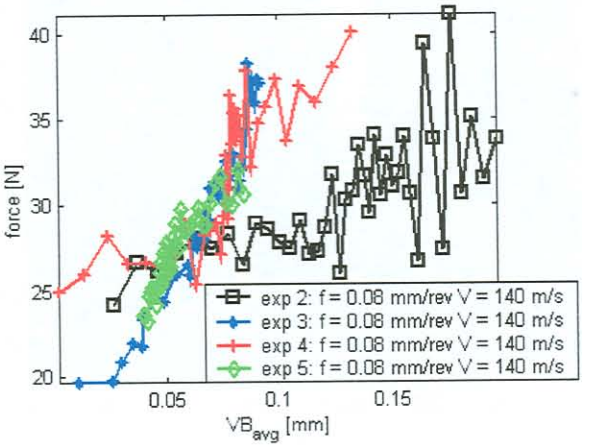


Figure 4.38: Static cutting forces (2)

4.4.6 Vibration analysis

A typical response history of acceleration during experiment 2 is shown in Figure 4.39. It can be seen that the vibration levels are fairly low, especially in the thrust direction. The basic method of analysing vibration signals in TCM is to check the vibration frequency spectrum for regions sensitive to tool wear.

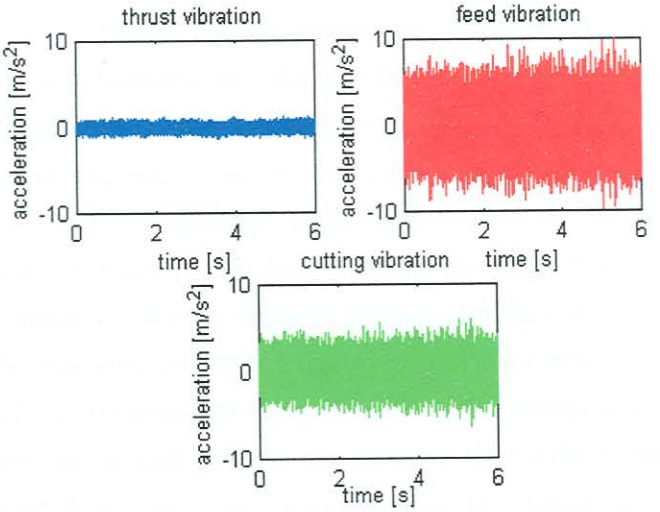


Figure 4.39: Typical acceleration signals experiment 2

Figure 4.40 depicts the thrust vibration measurements for the 5 overruns per workpiece during experiment 4 in a wear sensitive region of the frequency spectrum. Note the increase in vibration amplitudes at 1.2 kHz from measurement 1 to 180. The influence of the adjustment in cutting speed (due to decreasing workpiece diameter) is visible, but is small enough to ignore when looking at spectral energy values in a certain bandwidth. The results from the vibration analysis revealed that certain features derived from the vibration signals are sensitive to tool wear. However, it was difficult to determine features that were consistent for all the experiments. Furthermore, vibration is also sensitive to disturbances, and it is often difficult to explain why a particular vibration feature is influenced by tool wear. More results of the vibration signals are included in a further section.

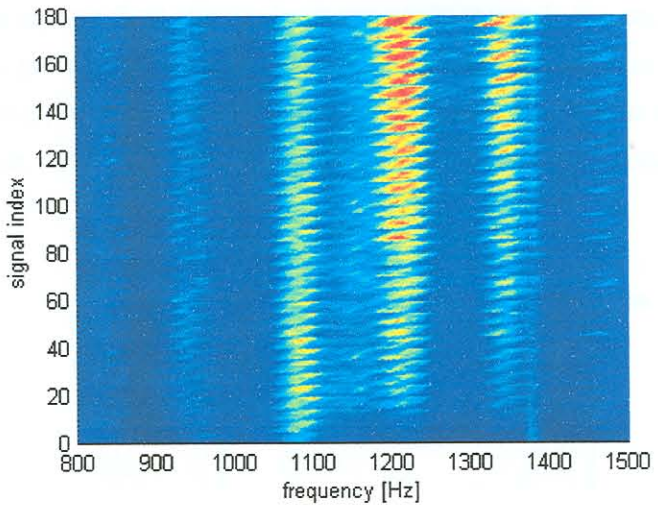


Figure 4.40: Thrust vibration spectra experiment 4 (linear scale)

4.5 Experimental results - correlation analysis with SOMs

4.5.1 Purpose of SOM analyses

To investigate the effect of all the different variables described in the previous section, several Self-Organising Map (SOM) analyses were conducted. The SOM is an unsupervised NN, also known as a Kohonen map, after their inventor [221], and is rapidly becoming a well-known tool for data exploration. The SOM automatically arranges multi-dimensional observation data on a two dimensional grid of neurons where similar observations are placed close to one another and dissimilar ones further away. In this way, hidden relationships in multi-dimensional data can be identified with a SOM analysis. In the case of the hard turning experiments, it was necessary to identify possible relationships in all the measured parameters. These relationships or correlations in the sensor information can also assist in the identification of disturbances in the data. Due to the large amount of sensory information, a SOM analysis is an appropriate method to explore the data. For instance, it was necessary to determine if the temperature of the machine or its surroundings have an effect on the process performance. Furthermore, the effect of tool wear on the surface finish of the components can also be investigated in this way. A detailed formulation and help on interpreting the SOM is included in Appendix H. For help on the interpretation of SOMs, refer to the recent work by Vesanto [222].

4.5.2 Results

A SOM analysis with data from experiment 1 is shown in Figure 4.41. The SOM reveals the relationship between surface finish (Ra), tool flank wear (VB), entropy (refer to Appendix F) of the Acoustic Emission (AE) signal, the static thrust force (Fz) and the temperature of the machine. From the SOM analyses it can be concluded that Ra does not degrade linearly with tool wear. It rather seems to reach a maximum value during medium wear conditions, and then decrease again. However, VB > 0.1 mm generally induces the formation of white layers, and should be avoided. Other observations are that the machine temperature has no influence on the workpiece surface finish, and that the static thrust force correlates well with tool wear. The results from the AE signal shows that AE reaches a minimum value at about 0.05 mm flank wear, and then start to increase again towards the end of the tool life, as shown in Figure 4.24. A sharp edged tool will cause high values of AE. During the normal worn phase of the tool, there is more damping present in the process, which cause the AE to decrease. During the final stage of the tool's lifetime, very high temperatures in the deformation zones and the onset of white layers cause an increase of AE.

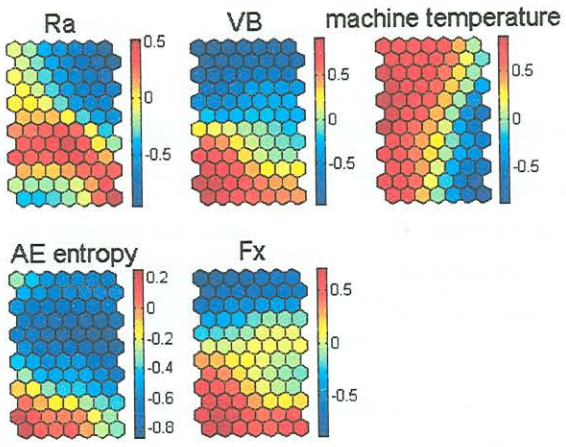


Figure 4.41: SOM analysis experiment 1 (1)

The result of a SOM analysis involving all the significant parameters from experiment 1 is shown in Figure 4.42. The SOM analysis basically confirms what was mentioned in the previous paragraphs dealing with experiment 1. From this analysis, the influence of (for example) tool wear on the other measured parameters can be derived. These influences are summarised in Table 4.5.

Table 4.5: Influence of VB on other variables:

	VB low	**VB high**
Surface Rz	Lower	Higher
Surface Ra	Lower	Higher
Workpiece diameter	Low and high values	More high values
Temp WP after	Towards lower values	Towards higher values

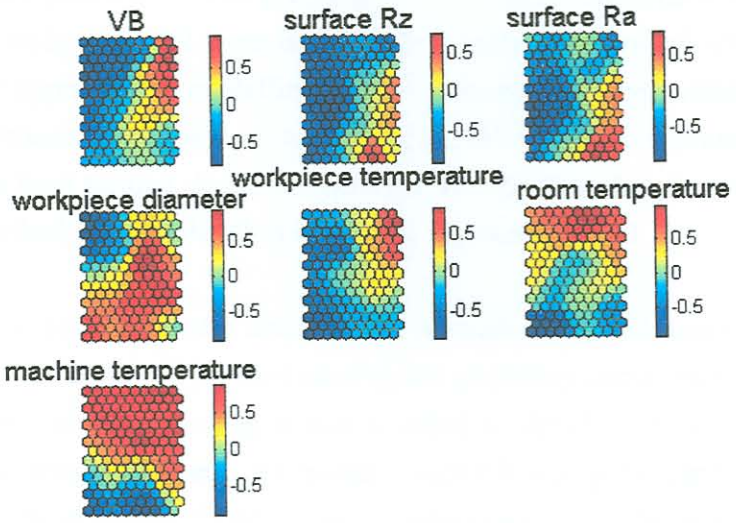


Figure 4.42: SOM analysis experiment 1 (2)

Furthermore, the influence of the machine temperature on the other parameters can also be derived. These results are summarised in Table 4.6.

Table 4.6: Influence of machine temperature on other variables:

	Temperature Machine low	Temperature Machine higher
Surface Rz	No apparent influence	No apparent influence
Surface Ra	No apparent influence	No apparent influence
Workpiece diameter	Show generally high values	High and low values

Another important conclusion from the SOM analysis is that the workpiece diameter is not influenced much by tool wear or the temperatures. Some recommendations resulting from this analysis were that the measurements of the workpiece and room temperatures be omitted for further experiments. The room temperature does not play a significant role, and the workpiece temperatures are difficult to measure accurately because the workpieces cool down very quickly after machining. The workpiece temperatures are also more governed by the time delay between two overruns than tool wear. Signal features were also investigated for sensitivity towards tool wear with SOM analyses.

4.6 AI approach for wear monitoring

4.6.1 Introduction

In order to establish a reliable method for monitoring the tool wear during hard turning, an appropriate wear model is required. During the course of this research many different types of modelling methods were investigated for their feasibility towards hard turning. Analytical and empirical methods have been developed through the years that can model many different machining operations. Analytical models are often deemed as too complex to be practical for a production environment. On the other hand, recently developed analytical and combined models can accurately model the behaviour of quite

complex tool / workpiece geometries. The problem with analytical methods is that they cannot model the behaviour of worn tools, and tool wear is a fact that cannot be avoided with present technology. Recently, a mechanistic approach for modelling worn tool forces appeared, which is probably the most novel combination approach developed up to date [223]. However, implementation of the method proved unsuccessful for hard turning due to some of the assumptions that do not apply to hard turning operations. This is described in more detail in Chapter 6 and Appendix G.

Hard turning experiments revealed that a combination of crater and flank wear govern the stability of the process. A worn tool force model for hard turning not including crater wear will not succeed. Considering this experience with hard turning it was decided to develop an Artificial Intelligence (AI) method with NNs to monitor tool wear. AI models resemble non-parametric empirical models, but have some advantages above conventional empirical equations. With the aim to ultimately estimate tool wear on-line with force measurements, it was decided to develop a NN monitoring strategy that utilises sets of static and dynamic NNs. The advantages of various NNs paradigms above other methods were described in Chapter 3. However, before the AI monitoring method can be described, certain signal features should be identified by means of signal processing and feature selection.

4.6.2 Signal processing

The usual signal processing procedure for TCM is to generate many signal features and then apply some kind of feature selection strategy to identify features that are most sensitive to tool wear. There are four domains from which features from sensor signals can be generated:

1. Statistical analysis
2. Time domain analysis
3. Frequency domain analysis
4. Joint time-frequency analysis (*e.g.* spectrograms and wavelet analysis)

Before choosing any of the domains for feature extraction, the following critical questions should be considered:

- Can the signal be approximated as stationary?
- Is a localisation of events in the time series important?

Signals taken from a cutting process will rarely be stationary, and hence simple time-frequency analyses with spectrograms are always necessary to identify the most stationary parts of the signal for frequency domain analysis (refer to Section 4.3). Investigations revealed the force signals are fairly stationary above about 50 Hz. In the lower frequency range, unstable chip formation causes chaotic frequency shifts. The type of chip that forms is governed by the severity of wear on the cutting edge, and there is no apparent 'reverse' effect. For the purpose of tool wear monitoring, a localisation of an event in the time series is of no importance. For monitoring tool breakage or other events, an accurate localisation could be of more importance. Therefore, for the purpose of a wear monitoring strategy for hard turning, time-frequency analysis was only conducted to investigate the general dynamic behaviour of the signals. No signal features were extracted from the time-frequency analysis.

The static components of the force signals are often used for TCM purposes, because static forces increase due to increasing friction between the tool and workpiece when wear is present. The static component is the mean of the force signal. The mean value of a signal *s(t)* over an interval T is:

$$\bar{s} = \frac{\int_0^T s(t)\,dt}{T} \tag{4.3}$$

The dynamic component of the force signal contains very useful information about the tool wear as well. The reason is because tool wear causes increasing vibration amplitudes in certain frequency ranges due to the larger frictional forces when the tool is worn. Previous research has suggested four possibilities from dynamic analyses that could be used to monitor tool wear:

1. Response at the natural frequencies of the tool holder.
2. Frequencies related to chip formation (*e.g.* chip serration and chip curl frequencies).
3. Natural frequencies of the dynamometer.
4. Vibration frequency of the tool insert when fixed in the tool holder.

Only option 1 is a practical possibility, especially when considering an industrial implementation of the TCMS. Investigations of option 2 did not reveal useful information, and option 3 is dependant on the presence of a dynamometer. Option 4 is subject to disturbance due to removal of the insert. Examples of FFTs of the force signals are shown in Figure 4.43. It is possible to identify certain frequency ranges in the FFT that are sensitive to tool wear, and the energy in such a range can be used as a signal feature with:

$$\psi_s^2 = \int_{fl}^{fh} S_s(f)\,df \tag{4.4}$$

where $S_s(f)$ is the one-sided Power Spectral Density (PSD) function of the signal and *fl* and *fh* chosen to reflect the energy in the regions of interest. The values of *fl* and *fh* should be chosen to correspond with a frequency range that will be sensitive to tool wear, and must be independent of the type of the workpiece and machine type. Investigating the tool holder response with the different experiments identified such a range between 1.20 – 1.25 kHz. This range is a natural frequency of the clamped tool holder. Figure 4.44 shows the result of the vibration (from accelerometer) energy in this range with respect to tool wear for one of the experiments. Though this feature is better measured with the accelerometer (high frequencies generate high acceleration levels), it can be extracted from the force signals as well (which is more related to displacement and hence lower amplitudes). However, to simplify an industrial implementation, it was decided to extract all the features for the method from the force signals only, and hence only one sensor would be required.

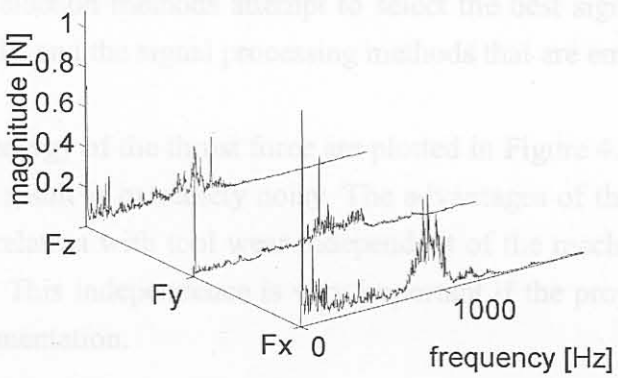


Figure 4.43: FFTs of force signals

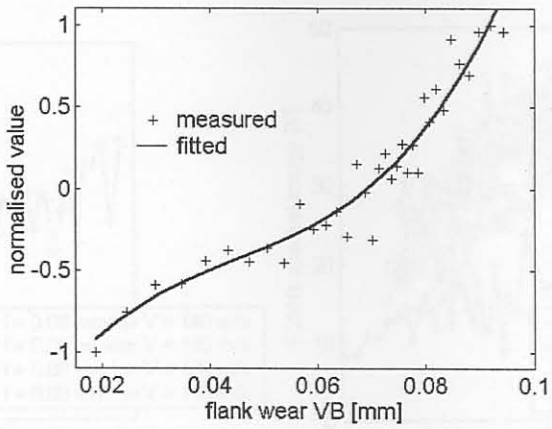


Figure 4.44: Vibration energy in the 1.20 – 1.25 kHz range (normalised)

Further signal features were generated by means of time domain and statistical analysis, but none revealed a consistent correlation with tool wear for all the experiments (typical features that were tested for feasibility are described in Section 5.6.3). Hence, based on linear correlation analyses between tool wear (crater and flank) and the various signal features, the following features were used as inputs to the NN that showed the best correlation with wear:

- static thrust, feed and cutting force (*Fx*, *Fy* and *Fz*)
- dynamic energy around the natural frequency (1.20 – 1.25 kHz) of the tool holder in the thrust force direction (*Fxd*)

These four features showed consistently increasing trends with tool wear for all the experiments (as shown previously). Linear or exponential models relating the static cutting forces with feed, speed, depth of cut and tool wear could be constructed. Such models could assist in selecting one of the static cutting forces as a feature for wear monitoring through a sensitivity analysis. Consequently a single static force component could be chosen for wear monitoring. However, for the purpose of this research, the route of automated feature selection is more appropriate. The reason for this choice is the fact that not enough data can always be gathered to construct and verify such models (also refer Chapter 2), especially if typical shop floor conditions are to be considered (refer for example to Chapter 4). Generally, features for NN models are selected through automated procedures such as correlation analyses and not through empirical verification models (there is also no need for NNs if such models are available or easily constructed). For this case study, the four features listed above were chosen because they displayed the best correlation with tool wear throughout the range of conditions investigated. Automated feature selection methods attempt to select the best signal features for NN models using the experimental results and the signal processing methods that are employed.

The results of the spectral energy of the thrust force are plotted in Figure 4.45. Once again, an increasing trend is visible but the result is extremely noisy. The advantages of the selected four features are that they showed good correlation with tool wear independent of the machine, workpiece, tool holder and machining parameters. This independence is very important if the proposed method is to be considered for industrial implementation.

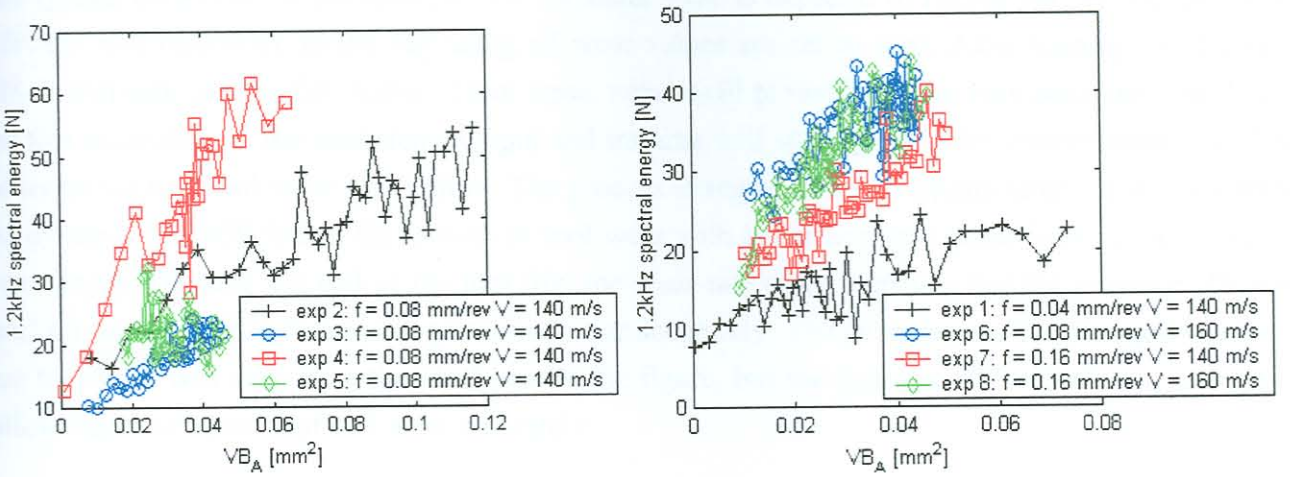


Figure 4.45: Spectral energy for thrust force

4.6.3 Formulation of method

The monitoring strategy was formulated in such a way that it utilises a combination of static and dynamic NNs. The advantages of using dynamic NNs for TCM were discussed in Chapter 3. In this research, a unique combination of static and dynamic NNs is proposed and will be shown to outperform other NN paradigms. The static NNs are trained to each model a particular feature with the flank wear, crater wear, cutting speed and feed rate as inputs. Early stopping is used to ensure generalised training, and separate training, validation and testing data sets were used (by means of cross-training and validation). The static networks each consist of a three-layer FF network, with ten sigmoid neurons in the middle layer and a tansig neuron in the output layer.

The dynamic NNs are trained on-line to estimate the wear values. The training target of the dynamic network is to minimise the difference between on-line force measurements and the output of the static networks. The advantage of this method is the fact that any development of wear can be followed on-line. The dynamic NNs use the previous three estimated wear values to estimate the next value of tool wear. This creates the ability to estimate an accurate value for the tool wear even if the training does not converge within the specified tolerance. This may happen if there was a problem with the measurement or disturbances in the signal. The behaviour of the dynamic NNs is thus to follow the curve or line of the present wear estimations, with the help of the on-line measurements and static networks. However, the method is able to generalise if errors or random disturbances occur. Another advantage is the fact that the dynamic NNs are able to follow the unique development of wear on the tool insert through on-line training. Thus, the dynamic NNs have delay and recurrent characteristics.

The method utilises sets of inner and outer steps or time-increments. The inner steps are training steps of the dynamic NNs to achieve a specified convergence. During the inner steps, the tool wear is assumed constant and the combined NNs attempt to estimate this value. When this is achieved, an outer step is taken, and in this case it is an incremental step in the tool wear. A schematic representation of the method is shown in Figure 4.46. There are two dynamic NNs: One to follow flank wear (DN_1) and the other to follow the growth of crater wear (DN_2).

The typical behaviour of the dynamic NN for flank wear is depicted in Figure 4.47 (to be read from left/right and up/down). In the beginning all wear values are set to zero. After training the dynamic NN, it will estimate the first value of tool wear, which will probably not be very accurate. It will then use this estimation in the next step as input and training will start again. After convergence, it will be closer to the true tool wear than before. The process is repeated and normally after a few outer steps the dynamic NN will follow the growth of tool wear with better accuracy (especially in the "regular" wear stage). Towards the end of the tool life, the wear rate often increases but the dynamic NN will react on this and again follow the growth of wear accurately. The advantage is that the growth of the true tool wear will vary from the ideal case in the figure, but the dynamic NN will always attempt to follow any development of tool wear accurately.

It is obvious that the dynamic NN will become more accurate if the outer steps are taken as close as possible to one another. An outer step is taken each time a new force measurement becomes available. This is advantageous for on-line applications, because a measurement can be taken during each over-run, and hence there is enough data to keep the outer steps very close to one another.

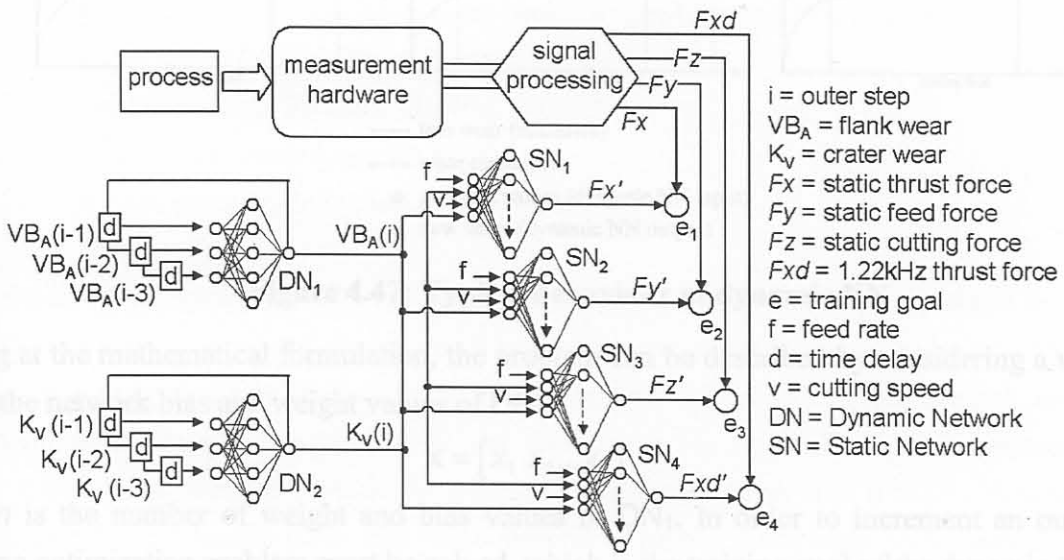


Figure 4.46: Schematic representation of monitoring method

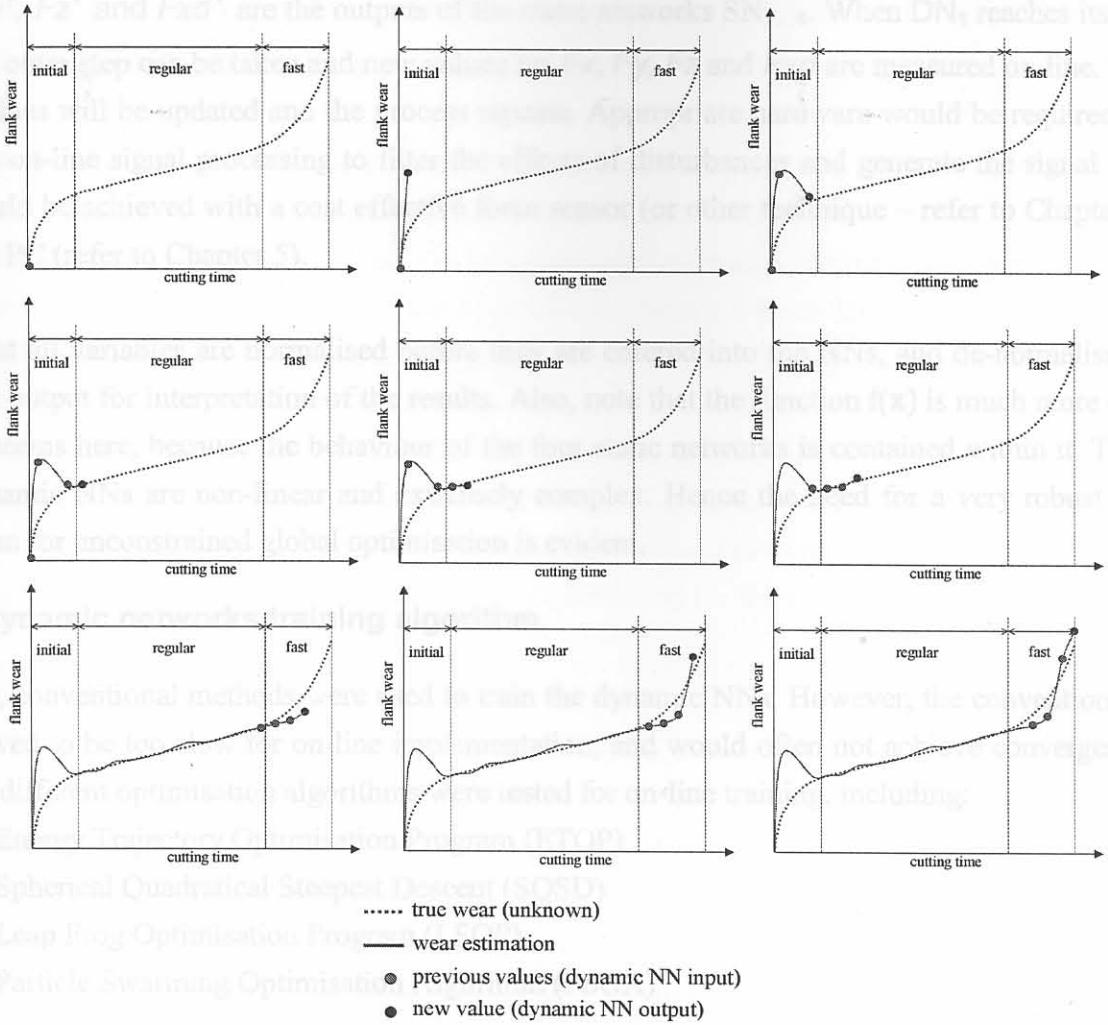


Figure 4.47: Typical behaviour of dynamic NN

Looking at the mathematical formulation, the problem can be described by considering a vector **x** containing the network bias and weight values of DN_1:

$$\mathbf{x} = [x_1\ x_2 \ldots x_n] \tag{4.5}$$

where n is the number of weight and bias values in DN_1. In order to increment an outer step, the following optimisation problem must be solved, which is the training goal of the dynamic NNs:

$$\text{minimise } f(\mathbf{x}) = \sum_{j=1}^{4} e_j \text{ such that } f(\mathbf{x}) \le tol \tag{4.6}$$

with the initialisation space for a new tool starting at:

$$D = \left\{ (x_1 \ldots x_n) \in \Re^n : -1 \le x_i \le 1,\ i = 1..n \right\} \tag{4.7}$$

and *tol* a suitable convergence tolerance on the function value. The design space for a worn tool is taken from the solution of the previous outer step. Initially, all wear values are set to zero. The error functions in Equation 4.6 are defined as:

$$\begin{aligned} e_1 &= \sqrt{(Fx' - Fx)^2} \\ e_2 &= \sqrt{(Fy' - Fy)^2} \\ e_3 &= \sqrt{(Fz' - Fz)^2} \\ e_4 &= \sqrt{(Fxd' - Fxd)^2} \end{aligned} \tag{4.8}$$

Fx', *Fy'*, *Fz'* and *Fxd'* are the outputs of the static networks $SN_{1...4}$. When DN_1 reaches its training goal, an outer step can be taken and new values for *Fx*, *Fy*, *Fz* and *Fxd* are measured on-line. The tool wear values will be updated and the process repeats. Appropriate hardware would be required as well as some on-line signal processing to filter the effects of disturbances and generate the signal features. This could be achieved with a cost effective force sensor (or other technique – refer to Chapters 2 and 5) and a PC (refer to Chapter 5).

Note that all variables are normalised before they are entered into the NNs, and de-normalised at the network output for interpretation of the results. Also, note that the function **f(x)** is much more complex than it seems here, because the behaviour of the four static networks is contained within it. The static and dynamic NNs are non-linear and extremely complex. Hence the need for a very robust and fast algorithm for unconstrained global optimisation is evident.

4.6.4 Dynamic networks training algorithm

Initially, conventional methods were used to train the dynamic NNs. However, the conventional methods proved to be too slow for on-line implementation, and would often not achieve convergence. Numerous different optimisation algorithms were tested for on-line training, including:

- Energy Trajectory Optimisation Program (ETOP)
- Spherical Quadratical Steepest Descent (SQSD)
- Leap Frog Optimisation Program (LFOP)
- Particle Swarming Optimisation Algorithm (PSOA)

A description and a more detailed formulation of the above-mentioned algorithms can be found in Appendix D. The algorithms listed above are all relatively new and unique methods for mathematical optimisation and are especially applicable to engineering-related problems. The formulation of the general mathematical optimisation problem is also documented in Appendix D. After exhaustive investigations, it was found that the Particle Swarming Optimisation Algorithm (PSOA) yielded the best result for training the dynamic NNs. The basic method of particle swarming optimisation was first suggested by Kennedy and Eberhart [224]. Training NNs is in essence an unconstrained global optimisation problem, and recent literature also states that the PSOA outperforms methods like Genetic Algorithms (GAs) in unconstrained global optimisation. The algorithm does not utilise gradient evaluations and its efficient use of random information is advantageous for training NNs. Furthermore the method is easy to implement and computationally efficient. Consider a flock of p particles (or birds). For each particle *i*, Kennedy and Eberhart [224] proposed that the position $\mathbf{x}^i$ be updated as:

$$\mathbf{x}^i_{k+1} = \mathbf{x}^i_k + \mathbf{v}^i_{k+1} \tag{4.9}$$

where **x** represents the position of each bird, **v** the velocity, and *k* a **unit** pseudo-time increment. The velocity $\mathbf{v}^i$ is updated with:

$$\mathbf{v}^i_{k+1} = w\mathbf{v}^i_v + c_1 r_1 \left(\mathbf{p}^i_k - \mathbf{x}^i_k\right) + c_2 r_2 \left(\mathbf{p}^g_k - \mathbf{x}^i_k\right) \tag{4.10}$$

where $\mathbf{p}^i_k$ represents the best ever position of particle *i* at time *k*, and $\mathbf{p}^g_k$ represents the global best position in the swarm at time *k*. The numbers r_1 and r_2 are random uniform numbers between 0 and 1. The

inertia term w is chosen as 1 and then decreased linearly during optimisation. The cognitive and social scaling factors c_1 and c_2 are chosen as $c_1 = c_2 = 2$ in order to allow a mean of unity.

With reference to equation 4.5, the dynamic NN biases and weights were taken as the position vector **x**, and the PSOA solves the optimisation problem described with equation 4.6. In this way, on-line training of the dynamic NN is achieved. On-line training is stopped when f(**x**) reaches the convergence value *tol*. If the convergence criterion is not reached, the vector **x** is randomly re-initialised for a new tool (refer equation 4.7). This is necessary for the case when an old tool is replaced with a new tool, because the swarm moved with the wear of the older tool, and will be too far from the solution to converge within limited steps for a new tool. The algorithm usually achieved convergence within approximately ten inner steps. This is possible in a matter of seconds on a Pentium III processor. If convergence is not achieved within ten steps, the training also terminates and the vector **x** is randomly re-initialised for a new tool. The same procedure is implemented for flank and crater wear.

4.6.5 Results

Results of wear estimations using this technique are shown in Figure 4.48 and Figure 4.49 for experiments 3 and 7, respectively. In each case, the method was tested on previously unseen data. The cutting speeds and feed rates were also included in the static NNs, but were not included in the dynamic NNs. From the results it can be seen that the method estimates the wear values quite accurately, despite the noisy nature of the training and testing data. Hence, the approximate shape and growth of the flank and crater wear can be determined through this method. The effect of a certain quantity of tool wear on the surface and subsurface quality of parts can be obtained from the literature, for example [225]. Thus, a decision can be made whether it would it safe to continue machining or not. In this case, only four estimation results are shown, but the method was successful in estimating the wear of the other experiments (by means of cross-training and cross-validation).

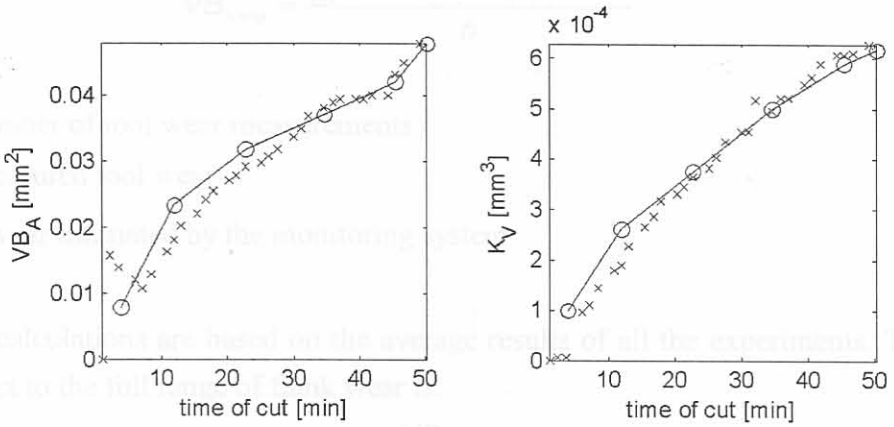


Figure 4.48: Estimation result experiment 3

o = measured values **x** = estimated values

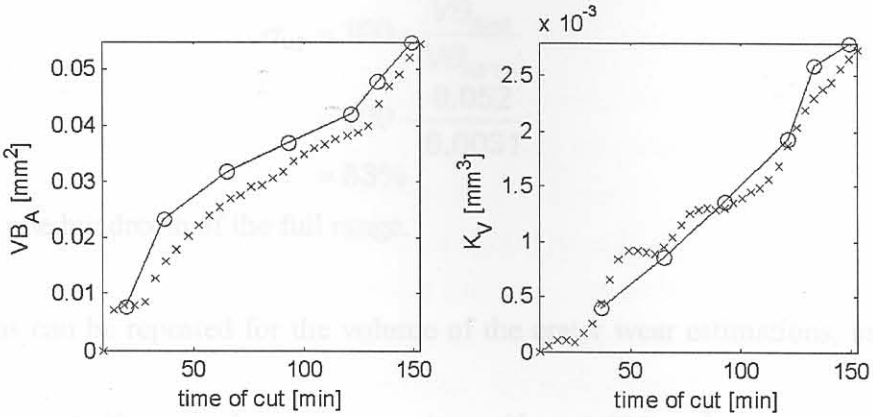


Figure 4.49: Estimation result experiment 7

o = measured values **x** = estimated values

4.6.6 Accuracy

It is essential to evaluate the accuracy of the proposed method if it will be considered for industrial implementation. The following terms and definitions are necessary to evaluate the overall accuracy of the TCMS:

$$\text{Full range of flank wear area } VB_{Afr} = 0.12\text{mm}^2 \tag{4.11}$$

$$\text{Maximum deviation from true wear } VB_{Amd} = 0.01\text{mm}^2 \tag{4.12}$$

$$\text{rms deviation from true wear } VB_{Arms} = 0.0025\text{mm}^2 \tag{4.13}$$

These values were calculated from the results of experiments 1-8.

The rms deviation can be described as:

$$VB_{Arms} = \frac{\sum_{i=1}^{n}\sqrt{(VB_{ATCMS} - VB_{Atrue})_i^2}}{n} \tag{4.14}$$

where:

n = the total number of tool wear measurements

VB_{Atrue} = the measured tool wear

VB_{ATCMS} = the wear estimated by the monitoring system

The following calculations are based on the average results of all the experiments. The average accuracy with respect to the full range of flank wear is:

$$\begin{aligned}\eta_{VB} &= 100 - \frac{VB_{Arms}}{VB_{Afr}} \times 100 \\ &= 100 - \frac{0.0025}{0.12} \times 100 \\ &= 97.9\%\end{aligned} \tag{4.15}$$

The minimum accuracy with respect to the full range of flank wear can be calculated as:

$$\begin{aligned}\sigma_{VB} &= 100 - \frac{VB_{Amd}}{VB_{Afr1\%}} \\ &= 100 - \frac{0.052}{0.0031} \\ &= 83\%\end{aligned} \tag{4.16}$$

where $VB_{Afr1\%}$ is one hundredth of the full range.

These calculations can be repeated for the volume of the crater wear estimations, using the following definitions:

$$\text{Full range of crater wear volume } K_{Vfr} = 0.006\text{mm}^3 \tag{4.17}$$

$$\text{Maximum deviation from true wear } K_{Vmd} = 0.0004\text{mm}^3 \tag{4.18}$$

$$\text{rms deviation from true wear } K_{Vrms} = 0.0001\text{mm}^3 \tag{4.19}$$

The rms deviation can be described as:

$$K_{Vrms} = \frac{\sum_{i=1}^{n} \sqrt{\left(K_{VTCMS} - K_{Vtrue}\right)_i^2}}{n} \tag{4.20}$$

where

n = the total number of tool wear measurements

K_{Vtrue} = the measured tool wear

K_{VTCMS} = the wear estimated by the monitoring system

The following calculations are based on the average results of all the experiments. The average accuracy with respect to the full range of crater wear is:

$$\begin{aligned}\eta_K &= 100 - \frac{K_{Vrms}}{K_{Vfr}} \\ &= 100 - \frac{0.0001}{0.006} \times 100 \\ &= 98.4\%\end{aligned} \tag{4.21}$$

The minimum accuracy with respect to the full range of crater wear can be calculated as:

$$\begin{aligned}\sigma_K &= 100 - \frac{K_{Vmd}}{K_{Vfr}} \times 100 \\ &= 100 - \frac{0.0004}{0.006} \times 100 \\ &= 93.3\%\end{aligned} \tag{4.22}$$

It can thus be stated that the method will estimate the area of flank wear correctly with an average accuracy of 98%, and in extreme cases not less than 83%. For crater wear, the average accuracy is also 98%, and the minimum accuracy is 93%. These calculations are based on all the experiments described in this chapter. It is possible that the accuracy might be lower during industrial implementations of the method, which is described in the next chapter.

4.7 Conclusion

A first important conclusion is that identification and isolation of disturbances on experimental data is essential during hard turning. Furthermore, some parameters influence one another in an unknown and complex manner, and these correlations were identified with a SOM analysis. This assisted in removing the effect of disturbances from the experimental data. The effect of disturbances could be removed from the data with appropriate signal processing methods. Possible features for a wear monitoring method were generated from the force signals after investigating numerous features from force, vibration and AE signals. The best features were selected after examining the effect of machining parameters, machine type and workpiece type on the data.

This chapter also demonstrated how AI could be utilised to effectively monitor crater and flank wear during hard turning. The method is formulated with a combination of static and dynamic NNs, and the training of the dynamic NNs is done with the PSOA. The advantage of this formulation is the fact that the dynamic NNs can follow any progression of wear, and protects the monitoring system against possible disturbances that may cause the system to estimate erroneous values of wear. The behaviour of the dynamic NNs is to follow on its present path, but use information from the force measurements to update its path. Thus, the low signal to noise ratio common to cutting force data has a very small impact on the accuracy of the method.

It was shown that the method is reliable and can be expected to operate with more than 90% accuracy under laboratory conditions. For a shop floor implementation, data from an actual shop floor would be required to verify the accuracy. The method is however applicable to the shop floor due to the low sensitivity towards disturbances. The method was trained for different machining conditions on different machines and it was shown that the method applies to the whole range of experimental conditions. Due to the unique formulation of the proposed AI method, it provides an accurate solution for monitoring crater and flank wear during hard turning. The suggested wear monitoring method is simple and flexible enough for on-line implementation, which will allow more reliable hard turning in industry.

5. Using AI for tool wear monitoring during interrupted turning of Aluminium

5.1 Introduction

In the previous chapter it was shown how a new method of combining static and dynamic NNs can achieve reliable wear monitoring during hard turning. Although the wear monitoring method is novel, the laboratory setup is similar to other research efforts. The question remains if the method is applicable to industry in terms of its accuracy, reliability and cost-effectiveness. In this chapter, the focus will be on the implementation of the method in industry using cost-effective hardware and appropriate signal processing techniques.

Sick [226] proposed a generic sensor fusion architecture for TCM, consisting of various analogue and digital processing steps. Many research efforts follow more or less this approach, focusing on one, some, or all of the processing levels (described in Chapter 3). However, in recent surveys (see Appendices A and B) of commercial hardware for TCM, it was found that no products are available in the market utilising these techniques. Furthermore, it has been found that not one single commercial TCMS is operational in South Africa. Most manufacturers deem the TCMSs developed up to date too expensive and unreliable. One of the problems identified during the course of this research is the fact that TCM strategies suggested up to date are almost without exception developed and tested on laboratory data. There are two major problems with this approach:

- Laboratory equipment cannot be used on the shop floor (for various reasons).
- The noisy shop floor causes false alarms in a monitoring system developed in a laboratory.

The solution to successful TCM lies in the development of the TCMS with the shop floor conditions in mind. Suggesting a generic structure for the TCMS, and adopting the structure for the specific process, machine, sensors and environmental conditions could achieve this. This chapter describes the experiments and results of an AI monitoring method that was implemented on the shop floor of a piston manufacturer. The method is in principle the same as described in chapter 4, with slight differences in terms of hardware and signal processing. Hence, it will be showed that method described in Chapter 4 under laboratory conditions, can also be applied in industry with slight modifications. The process that is considered here is interrupted turning of an Aluminium alloy during piston manufacture. There are several reasons for this particular choice of operation, the main ones being:

- Industrial need
- It is a turning process with varying depth of cut
- Two different types of cuts are performed, namely facing and boring
- There are similarities between the mechanics of tool wear during interrupted turning and face milling, meaning that the approach could possibly be extended to milling.

5.2 Measurement system

5.2.1 General considerations

Apart from the fact that a TCMS must be designed with the shop floor environment in mind, a further solution to successful TCM lies in *"getting closer to the process"*. The initial selection and placement of sensors is of utmost importance. Tool wear induces very small changes in machining processes. There are also many other obstacles such as diverse failure modes, changing machining parameters and external disturbances that make the task of TCM very difficult.

A sensor that complies with all the TCM sensor requirements (refer to Chapter 3) does not exist. There will always be a trade-off: if the sensor is designed to be as technologically advanced as possible, it will be unaffordable. Manufacturers require a cost-effective solution in order to employ TCM on all their processes. It is universally recognised that force and vibration monitoring can be used for the TCM of turning operations. Often only the static components of the forces are used for TCM. However, using the dynamic components from the force and / or vibration signals can significantly enhance the performance of the TCMS [227]. It has also been shown that some dynamic components are less sensitive to changing machining conditions and external disturbances [6]. Force dynamometers are commonly used in machining experiments (*e.g.* Chapter 4). Although they can measure the static and dynamic forces very accurately, dynamometers are not suitable for the shop floor. Furthermore, their frequency ranges are usually very limited, with a natural frequency f_o (of the piezoelectric parts) at about 3kHz, and a resonant frequency f_r at about 1kHz when clamped. To avoid amplitude distortion, the usable frequency range of a piezoelectric transducer is about $0.6 \times f_o$. For dynamometers commonly used for turning experiments this amounts to about 1.8kHz.

Tramal and Opavsky [228] investigated the dynamics of a conventional force dynamometer for machining operations. It was found that the dynamometer has significant amplitude distortion in the frequency range that is quoted as the operating range by the manufacturer. Unfortunately the authors did not investigate the phase, but it can be assumed that there will be phase distortion as well. The authors suggest that the dynamic characteristics of the dynamometer (while clamped like it would be during measurements) be identified with a modal test and the effect of dynamometer dynamics be compensated for after measurements are made.

Schulz *et al.* [229] recently presented a development of sensor-integrated tool holders. Currently, systems are developed for drilling and milling applications. Forces are measured directly with the mechatronic tool holder with the advantage that the measurement is done close to the cutting edge. The instruments also make use of a flexible telemetric unit for wireless data transfer. Piezoresistive strain gauges are used for estimation of cutting forces. Unfortunately, it is unclear how the system is calibrated and how the clamping and high-speed rotation of the tool holder affects its accuracy. Furthermore, no information about the operational frequency range is given, and it is assumed that the system can only measure static forces. One advantage is an interface of the tool holder directly to the CNC control, which will assist in fast reaction if an undesirable condition is encountered.

It has been found by various authors that the frequencies useful for TCM can be as high as 8kHz, and usually above 1kHz, depending on the experimental setup. A quite simple method to estimate both the static and dynamic components of cutting forces without any distortion is to use resistance strain gauges. Strain gauges comply with most of the requirements for TCM sensors, and can accurately follow the static and dynamic response of a system up to 50kHz. It was therefore decided to develop a sensor-integrated tool holder using strain gauges for this research. Such a system would be cost-effective and could be used on a shop floor. The reader is referred to Appendix C for relevant formation on strain gauges.

5.2.2 Experimental setup

Li and Ulsoy [230] showed how the exact displacement at the tip of a boring bar can determined using strain gauges. Each vibration mode of a beam has its own characteristic gain for converting strain components into displacement, and these are taken into account to determine the beam displacement with submicron precision. The strain gauge measurements can thus be used to determine the exact tool tip displacement during flexible line boring, by using a simulation model.

Santochi *et al.* [103] from the University of Pisa showed how a tool holder could be modified to create a *"sensor-integrated tool"*. Strain gauges were adhered to a standard turning tool holder to estimate cutting forces. The tool holder also contained an amplifying and infrared data transmission unit for wireless data transfer. It was shown that the sensor-integrated tool could be used with relative ease and yields accurate measurements for a range of cutting conditions. A further development of the same tool appeared recently, where radio signals are used instead of infrared, and digital coding of the signal is included [232]. A schematic representation of the sensor-integrated tool developed at the University of Pisa is shown in Figure 5.1.

In this work it was also necessary to develop a similar sensor-integrated tool for use on the shop floor. The main requirements for such a sensor-integrated tool are:

- cost-effective
- robust
- accurate

It will be shown in this chapter that all these requirements were met. Furthermore, in order to obtain a large amount of shop floor data, it was decided to automate the system by making it a data logger that can be controlled and monitored via the Internet. The complete measurement system is schematically depicted in Figure 5.2. The system consists of the following:

- tool holder with strain gauges (3 half-bridges)
- strain gauge amplifiers
- anti-alias filters
- A/D conversion
- computer with data logging software

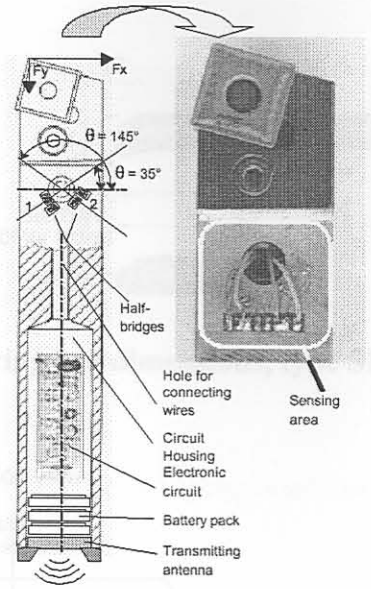


Figure 5.1: Sensor integrated tool [232]

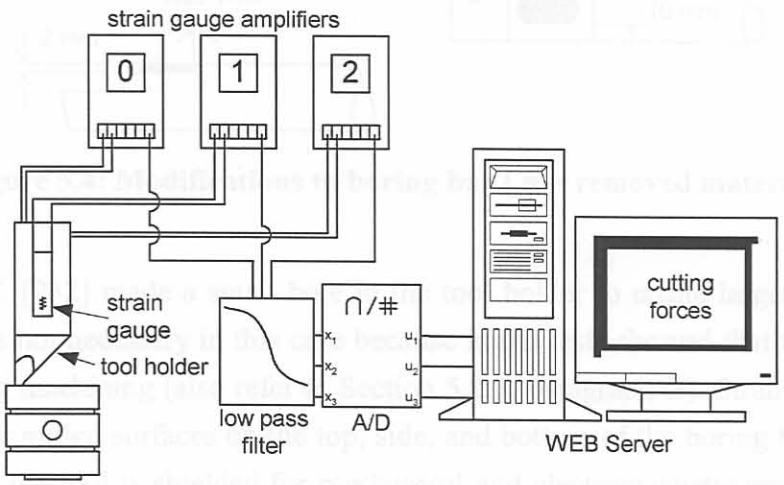


Figure 5.2: Schematic layout of measurement system

The components of the measurement system were either custom designed and built in-house, or were cost-effective commercial products. The complete development of the hardware forms a significant part of the thesis, and more details about the hardware development are given in Appendix E.

5.2.3 Tool holder with strain gauges

Three half-bridges (refer to Appendix C) were constructed on different parts of the tool holder. This creates three measurement channels that can be investigated for sensitivity towards tool wear. Furthermore, channel information should be fused as described in Section 4.6. The tool holder (in this specific case study a boring bar) had to be modified somewhat in order to create flat surfaces for the strain gauges. There was no significant loss of stiffness of the tool due to this modification. The dimensions of the original boring bar are shown in Figure 5.3. An example of the typical modifications is shown schematically in Figure 5.4.

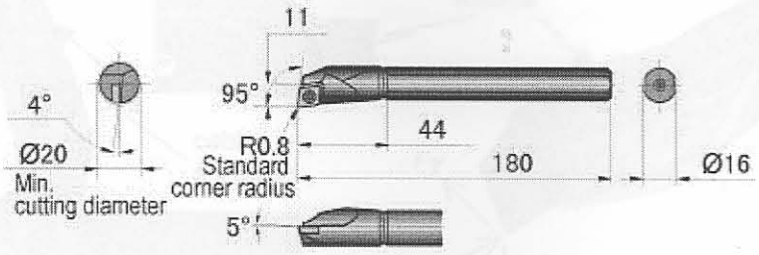


Figure 5.3: Boring bar dimensions, type S16QSCLPR09 [1]

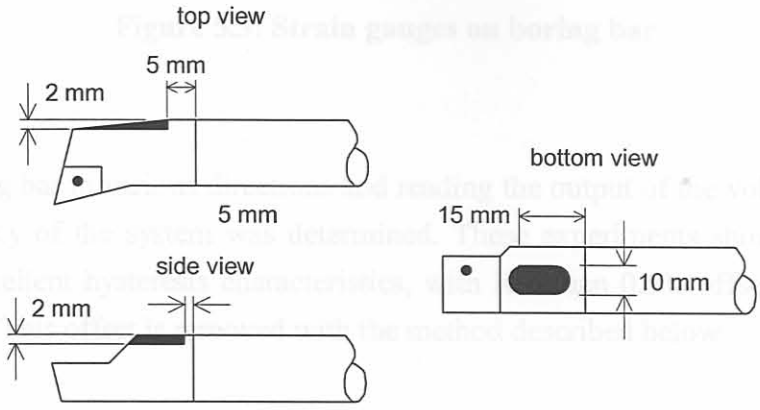


Figure 5.4: Modifications to boring bar (■ = removed material)

Although Failli *et al.* [232] made a small hole in the tool holder to create larger strains (mechanical amplification), it was not necessary in this case because initial tests showed that the strains in the tool are quite large during machining (also refer to Section 5.2.3, paragraph G). Strain gauge rosettes were simply adhered to the milled surfaces on the top, side, and bottom of the boring bar. After application of the strain gauges, the tool is shielded for mechanical and electromagnetic protection. Five sensor-integrated boring bars were manufactured during the course of this work. Pictures of some of the physical layouts of the gauges and shielding material are shown in Figure 5.5. Further information about the strain gauge circuits can be found in Appendices C and E. Some relevant issues with respect to the measurement system are discussed in the following sections.

A. Frequency range of strain gauges

A strain gauge tends to give an integrated average of the strains imposed over its length. The smaller the strain gradient across the element length, the closer the output will be to the true strain. To choose a strain gauge, the desired accuracy of peak strain and frequency extension must be considered. For the tool holder application, 5mm gauges were used, which offer very good accuracy up to 50kHz on steel. A detailed discussion on the electrical resistance strain gauge and how the gauge length should be selected can be found in Appendix C.

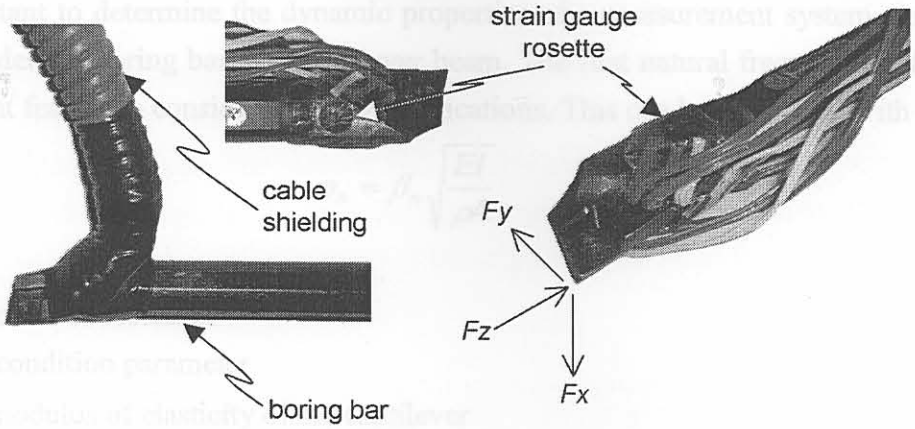


Figure 5.5: Strain gauges on boring bar

B. Hysteresis

By bending the boring bar in various directions and reading the output of the voltages on the channels, the hysteretic accuracy of the system was determined. These experiments showed that the measurement system had excellent hysteresis characteristics, with less than 0.5% offset due to hysteresis effects on all channels. This offset is removed with the method described below.

C. Temperature compensation and drift

Modern-day strain gauges are self-temperature compensated. The strain gauge amplifiers displayed some drift with variations in room temperature. The drift due to temperature effects causes a DC offset in the sensor signals. This offset can be digitally removed from the signal by subtracting the DC values collected when the tool is not engaged in cutting. Care must be taken not to let the system drift into the threshold value for triggering (refer to Section 5.2.7). Experiments were conducted to calculate the drift of the system with variation in room temperature, and the result is shown in Appendix E.

D. Clamping condition

The layout of a typical tool-holder setup is shown in Figure 5.6. The tightening of the clamping screws on the tool holder can cause the holder to bend and cause a DC offset on the measurement channels. The system is zeroed each time when a new tool is clamped, thus removing this effect. The tool holder is in actual fact very seldom removed.

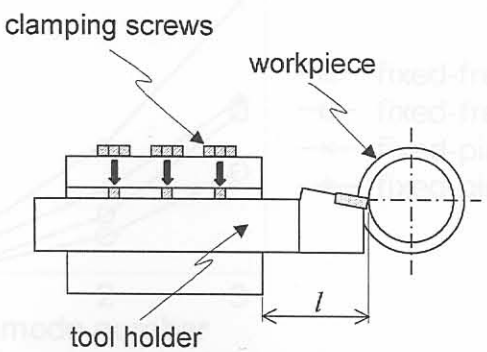


Figure 5.6: Clamping of tool holder

It is very important to determine the dynamic properties of a measurement system. An analytical approach is to model the boring bar as a cantilever beam. The first natural frequency of the cantilever is also an important feature to consider for TCM applications. This can be calculated with [233]:

$$\omega_n = \beta_n \sqrt{\frac{EI}{\rho A}} \tag{5.1}$$

where:

n = the mode of vibration
β_n = end condition parameter
E = the modulus of elasticity of the cantilever
I = the second moment of area
ρ = mass density of tool holder material

β_n is dependant on the length of the overhang of the cantilever (l), and for the first mode in a fixed-free configuration it is given by:

$$\beta_1 = 1.875104 / l \tag{5.2}$$

For a fixed-pinned configuration it is given by:

$$\beta_1 = 3.926602 / l \tag{5.3}$$

The physical difference between the two configurations is diagrammatically depicted in Figure 5.7.

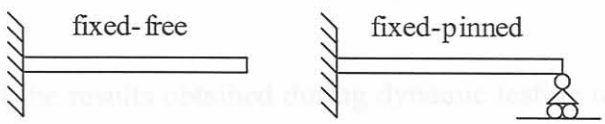


Figure 5.7: Cantilever models

Substituting $\rho = 7850 kg/m^3$, $E = 200 GPa$, $l = 46 mm$, and tool holder radius $r = 16$ mm, the natural frequencies of the cantilever model can be calculated (for the Fx direction) for each new value of l. The result of Equation 5.1 for different values of l and a comparison between a fixed-free and a pinned-free configuration is shown in Figure 5.8. According to this result, the first natural frequency of the boring bar can be expected to be below 10kHz for typical values of l. The higher modes are too high to have any significant influence on the system.

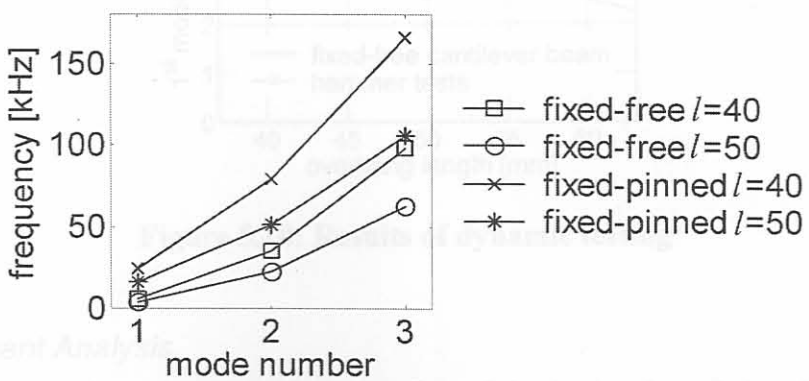


Figure 5.8: Cantilever models natural frequencies

E. Hammer tests

A simple method to experimentally determine the natural frequencies and verify the results described above is to conduct a modal hammer test. When the holder is clamped in position, a modal hammer can be used to dynamically excite the system, and the Frequency Response Functions (FRFs) from the strain gauges can be determined. In this case, the first mode is dominant when excited on the tip of the tool holder with the modal hammer. A picture of the experimental setup for hammer tests is shown in Figure 5.9.

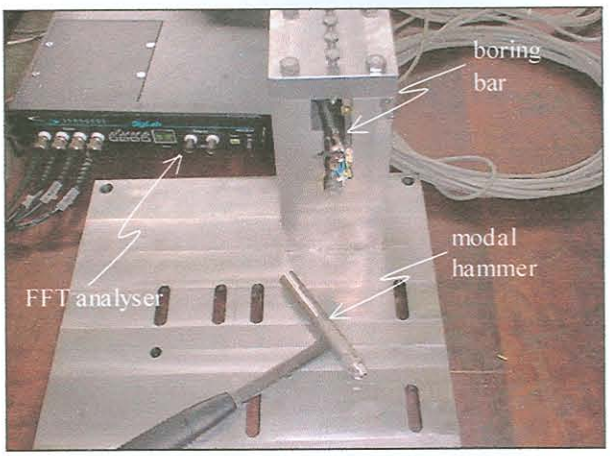


Figure 5.9: Modal hammer experimental setup

Figure 5.10 is a summary of the results obtained during dynamic testing of the sensor-integrated tool. The following conclusions can be made:

- the cantilever model is a rough yet valid estimation method
- the fixed-free cantilever model is the most accurate
- The first natural frequency can be expected to be below 5kHz, probably in the area of 3.5kHz

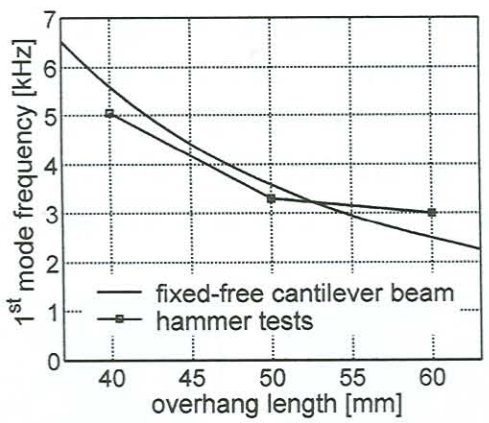


Figure 5.10: Results of dynamic testing

F. Finite Element Analysis

As a further investigation into the dynamic properties of the boring bar, finite element analysis with MSC Nastran was conducted. The automesh function of the software with tetrahedral 10 elements was used. The result of the automesh is shown in Figure 5.11. The aim of the analysis was to:

- verify dynamic tests results
- determine if machining the boring bar has a big influence on the stiffness properties
- determine optimal position to apply strain gauges
- determine the mode shapes

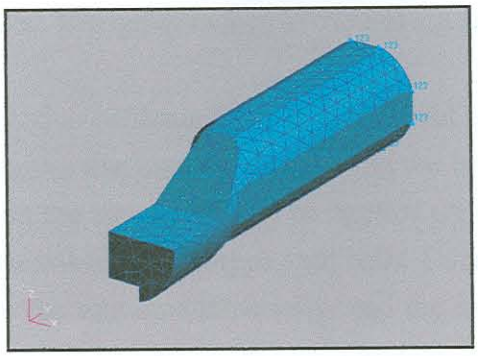


Figure 5.11: Finite element model of boring bar

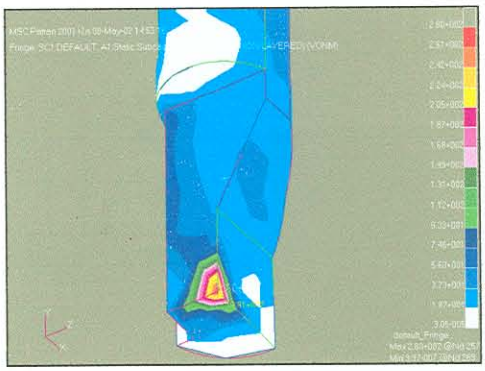


Figure 5.12: Von-Mises stress distribution with load at tool tip

A result of the Von-Mises stress distribution with a typical load on the tool tip is shown in Figure 5.12. Suitable positions for the strain gauges were identified in this way. The stress analysis showed that the machining of flat surfaces had little impact on the properties of the tool holder. Unfortunately, the natural frequencies could not be resolved very well with the model. This can be attributed to the boundary conditions of the clamping which is difficult to model properly. Some of the mode shapes from the finite element analysis are shown in Figure 5.13. The mode shapes also assisted in selection appropriate positions for the strain gauges (refer to Section 4.6.2 on natural frequencies).

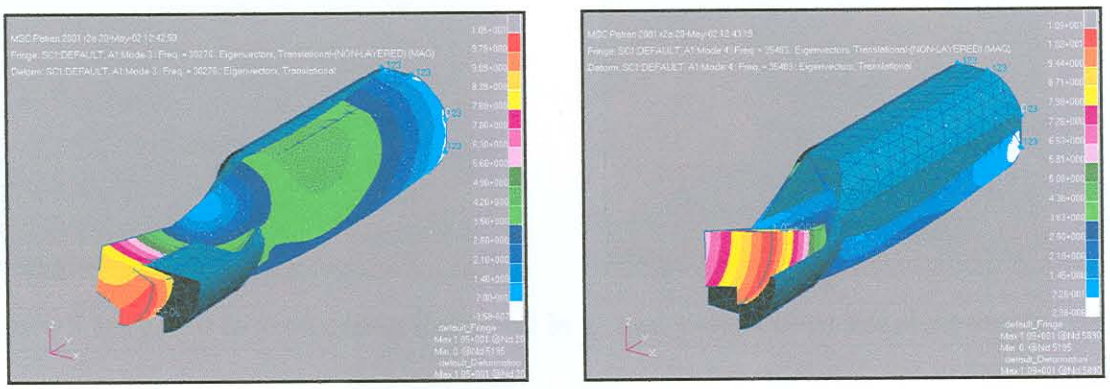


Figure 5.13: Modes three (left) and four (right)

G. Calibration

With the dynamic properties identified, the next step to ensure proper measurements is to perform a static calibration. A matrix of calibration factors for the system can be determined in order to convert voltage into force. Applying known forces to the system and reading the output of the strain gauges will accomplish calibration. In order to calibrate the sensor-integrated tool, a special static calibration test rig was designed and assembled. Design drawings of the test rig are shown in Figure 5.14.

The calibration test rig was designed to calibrate the tool at different overhang lengths. This is possible with an assembly of clamping devices that can also slide along the base plate. By means of a threaded socket on the three mini towers on the rig, forces can be applied on the tool tip in the *Fx*, *Fy* and *Fz* directions. The input forces are measured with Z-type load cells. Consequently, different forces can be applied in the three directions on the tool simultaneously, and the test can be done for any overhang length. The voltage output on the strain gauges is read and these are used to determine a matrix of calibration coefficients. Pictures of typical load cases are shown in Figure 5.15.

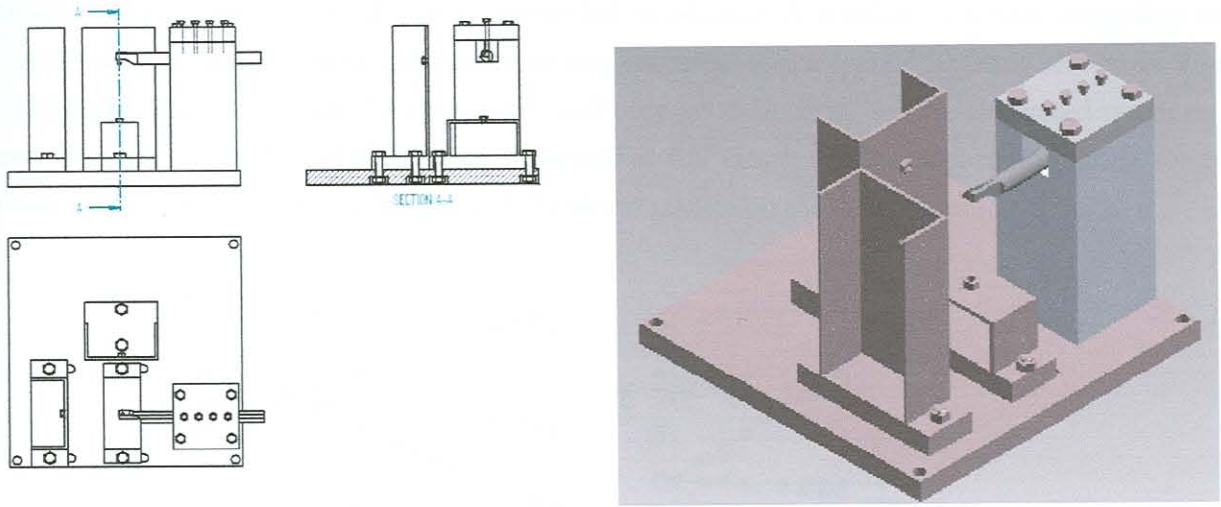


Figure 5.14: Clamped tool on calibration test rig

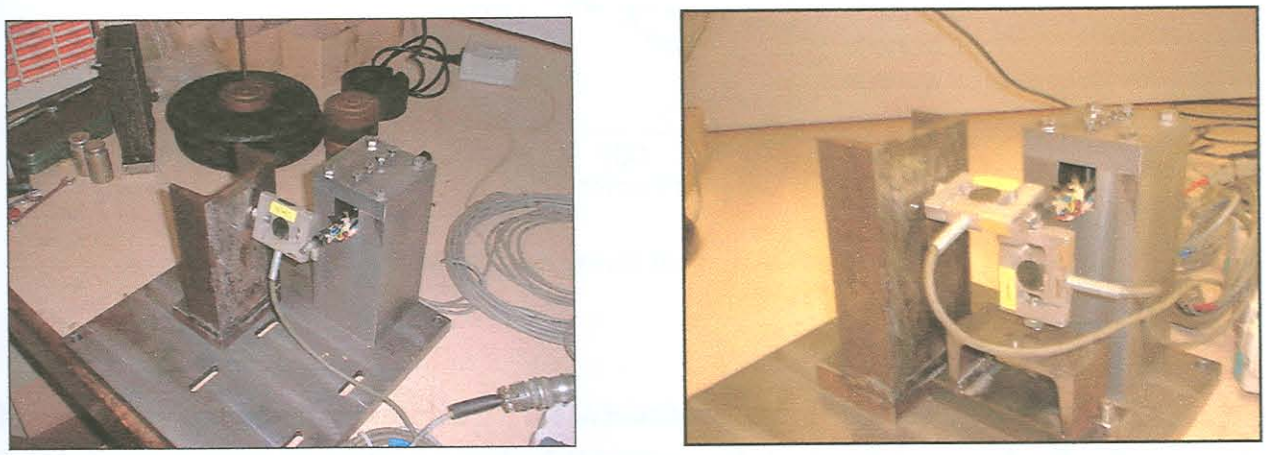


Figure 5.15: Load in y-direction (left) and in x- and y-directions simultaneously (right)

The strain on the tool holder can be related to the force through a linear system written as:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \mathbf{A} \begin{bmatrix} Fx \\ Fy \\ Fz \end{bmatrix} \quad \text{with } \mathbf{A} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \tag{5.4}$$

$$\text{or } \boldsymbol{v} = \boldsymbol{Af} \tag{5.5}$$

Matrix coefficients k_{ij} for calibration matrix $\boldsymbol{A}$ can be determined by setting Fx to a known value, and Fy and Fz to zero. For example, setting Fx to 200N:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} 200 \\ 0 \\ 0 \end{bmatrix} \text{ yields } \begin{matrix} v_1 = 200k_{11} \\ v_2 = 200k_{21} \\ v_3 = 200k_{31} \end{matrix} \tag{5.6}$$

The measured data can the be calibrated on-line with:

$$\boldsymbol{f} = \boldsymbol{A}^{-1}\,\boldsymbol{v} \tag{5.7}$$

The results of the calibration procedure showed that Fx and Fy could be determined with an accuracy of about 1% for the static calibration range. Fz can be determined with limited accuracy (approximately 15% for static forces) due to the fact that longitudinal stiffness of the tool holder is higher in this direction. The response of Fz is also somewhat non-linear. A number of sensor-integrated tools were calibrated and tested during the course of this research, and a typical result is shown in Figure 5.16. The calibration of Fx and Fy always showed very good linearity for the calibrated range on all channels. The sensor-integrated tools were calibrated at full range, and the calibration matrix for each too is unique due to the shape and location of the strain gauges on the tool holder.

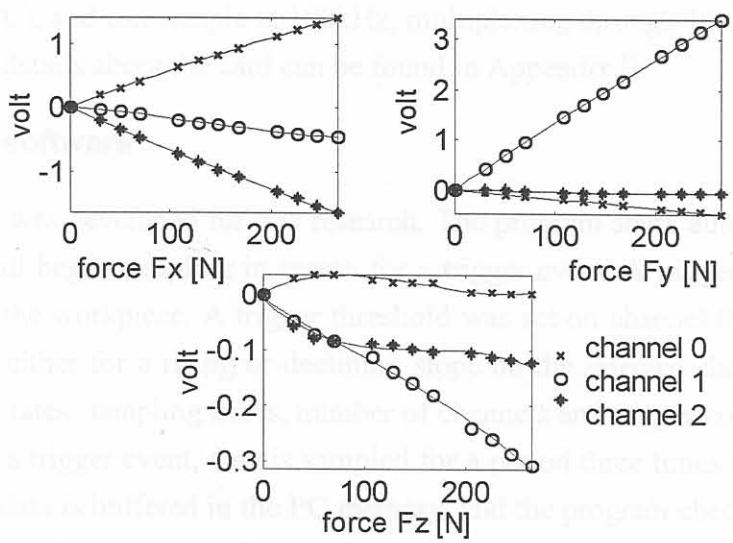


Figure 5.16: Calibration result for one of the developed tools

5.2.4 Strain gauge amplifiers

The purpose of the amplifiers is to excite the Wheatstone bridge used for strain gauge measurements. The amplifiers provide the voltage over the full-bridges and also have a zero-pod and a digital read-out unit. One consideration when choosing a strain gauge amplifier is the frequency range of the amplifier. The usable range of the amplifiers used for this application is DC – 6kHz. More information can be found in Appendix E.

5.2.5 Anti-alias filters

A very important consideration when sampling dynamic signals is the phenomenon of aliasing. Aliasing can only be prevented by:

- Using a sampling rate (F_s) of at least twice the bandwidth (BW) of interest.
- Using an analogue anti-alias filter with corner frequency at BW.

Versatile analogue filters are available commercially, though at a very high price. Furthermore, commercial filters are usually designed for laboratory work and it cannot be afforded to leave them permanently on the shop floor. Therefore it was decided to build an anti-alias filter in-house, with the help of the *Filterlab Low-pass* software from Microchip [234]. More details on the filter development are available in Appendix E.

Three identical channels were built to pass DC - 4000Hz, with the -3dB point at approximately 4350Hz. A sampling frequency of 20kHz is suggested when using the filter to prevent aliasing in the DC – 4kHz band. However, the filter has significant phase distortion on analogue signals. This characteristic can be rectified digitally by creating an identification model of the filter. An Autoregressive (AR) time series model was made from identification data of the filter. This model is then saved and all measured data are then reversed through the model to compensate for the phase distortion.

5.2.6 A/D conversion

A/D conversion was done with a 12-bit card from Eagle Technologies (PC-30). The card slots into a Personal Computer (PC), and can sample at 100kHz, multiplexing through 16 single-ended or 8 differential channels. More details about the card can be found in Appendix E.

5.2.7 Data logging software

Special C++ software was developed for this research. The program starts automatically when the PC is powered on, and will begin sampling in search for a trigger event. A trigger event will occur when the tool engages into the workpiece. A trigger threshold was set on channel 0 at +0.05V. The trigger can be selected to be either for a rising or declining slope on the selected channel. The user can initially set the sampling rates, sampling times, number of channels and trigger conditions on a web page. During the search for a trigger event, data is sampled for a period three times longer than the selected sampling length. The data is buffered in the PC memory, and the program checks it for a trigger event. If a trigger event occurs, the program will save the data for the user-selected period. The program can also be instructed to save only every 10^{th}, 20^{th} or N^{th} trigger event, as not to save the data for every workpiece. Each file is saved with a time stamp. As a last step, after a number of data files were collected, they are stored in a compressed format to save disk space. The compressed files are then available for download through the Internet. The program settings can also be changed remotely via the Internet. On the shop floor, a person is responsible for logging tool changes and safekeeping the tools for later investigation under a microscope. For some tool inserts regular wear measurements were made. Each machined part is also automatically logged and time stamped on the shop floor and can thus be correlated with the recorded data. A flow diagram of the program is shown in Figure 5.17. The

complete system was installed in a steel enclosure, shown standing on the shop floor in Figure 5.18. A screen capture of the web page where the program setting are submitted in shown in Figure 5.19, and a screen capture of the logger in operation is shown in Figure 5.20.

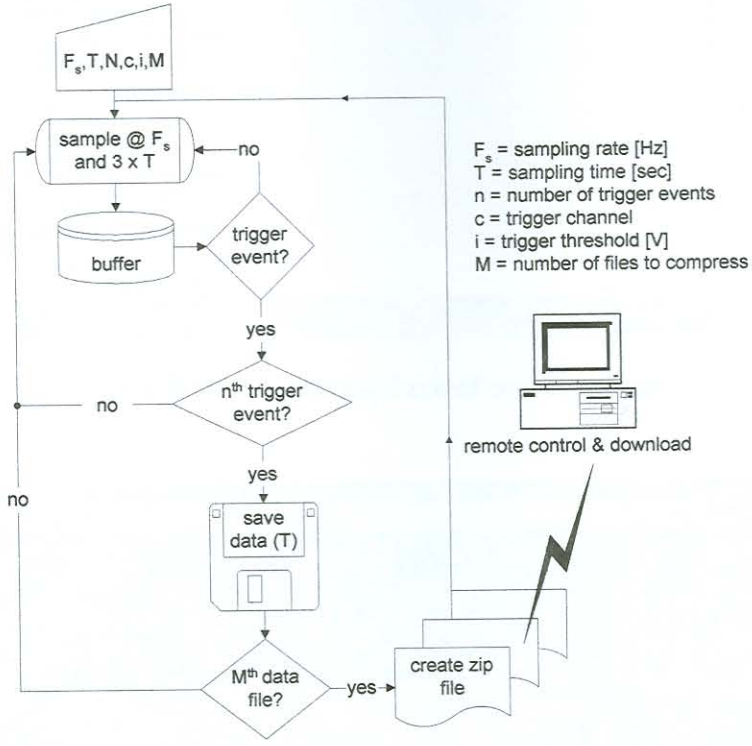


Figure 5.17: Flow diagram of data logging program

Figure 5.18: Steel enclosure for monitoring electronics

5.3 Cost breakdown

Cost-effectiveness was a very important consideration for the development, due to the various reasons stated in earlier chapters. The approximate cost of the individual components are summarised in Table 5.1. The total cost (hardware and software) of the system amounts to about €1000. This is significantly lower than any commercial measurement system available today.

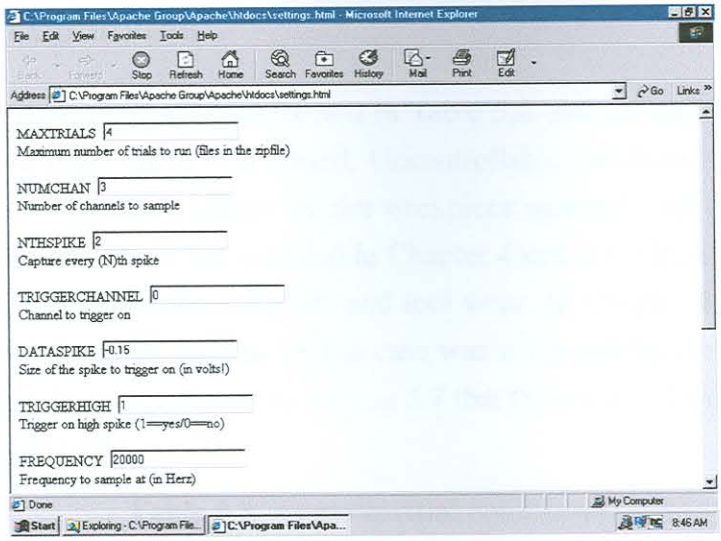


Figure 5.19: Screen capture of settings page

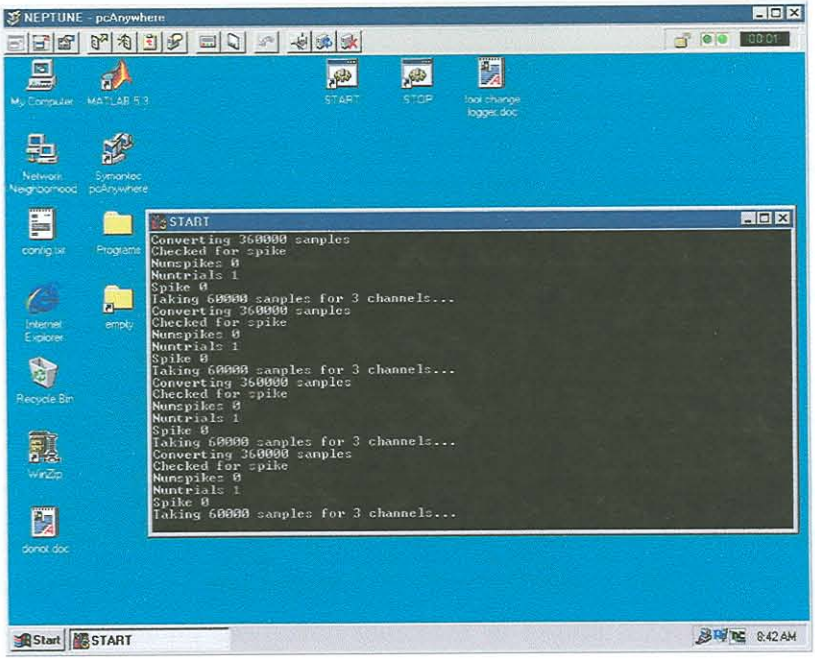


Figure 5.20: Logger in operation

Table 5.1: Cost breakdown of measurement system

boring bar	€ 80
strain gauges	€ 70
shielding material	€ 80
strain gauge amplifiers	€ 150
anti-alias filter	€ 20
A/D card	€ 100
computer (Pentium @ 200MHz)	€ 250
software	€ 100
miscellaneous costs	€ 150
Total	**€1000**

5.4 Experimental conditions

The basic experimental conditions are summarised in Table 5.2. All controllable conditions remained the same except the feed rate, which was varied. Uncontrollable conditions refer to external disturbances such as shop floor vibrations, softer / harder workpiece material, tool quality *etc*. These disturbances were more chaotic than those encountered in Chapter 4 and it is virtually impossible to investigate their individual effects on the force signals and tool wear. In Chapter 4 the effects of identified disturbances were filtered from the signals. In this case was not possible. However, despite the noisy nature of the environment, it will be shown in Section 5.7 that the proposed wear monitoring method is robust towards these conditions.

Table 5.2: Experimental conditions

machine	feed rate	0.3 mm/rev and 0.2 mm/rev
	cutting speed	390 m/min
	depth of cut	variable 1.7 mm max
	tool holder	S16 QSCLPR 09
	insert type	CCGT 09T304 FN-ALU (carbide)
	workpiece	Aluminium alloy (confidential)
strain gauges	3 half-bridge configurations	
	type	KYOWA KFG-1-120-D17-11
	amplifier	HBM clip® System
filter	custom made cut-off at 4 kHz	
A/D card	Eagle Technologies PC-30	
	sampling rate	20 kHz per channel
	sampling time	3 seconds per channel

The machine on which experiments were conducted is depicted in Figure 5.21, with a picture of the sensor cable exit from the machine in Figure 5.22. Some pictures of the sensor-integrated tool in operation are shown in Figure 5.23 and Figure 5.24. The operations that were considered consisted of a facing and boring on the areas on the piston indicated in Figure 5.25. It is a crucial section on the workpiece with respect to geometric accuracy and hence monitoring of the tool wear is essential. Due to the interrupted character of the cut, the tools also tend to wear fast. More details on the tool wear will follow in Section 5.5.

Figure 5.21: Turning machine

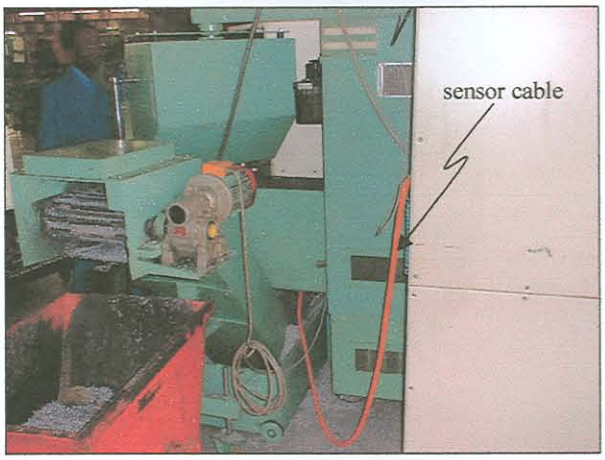


Figure 5.22: Cable exit

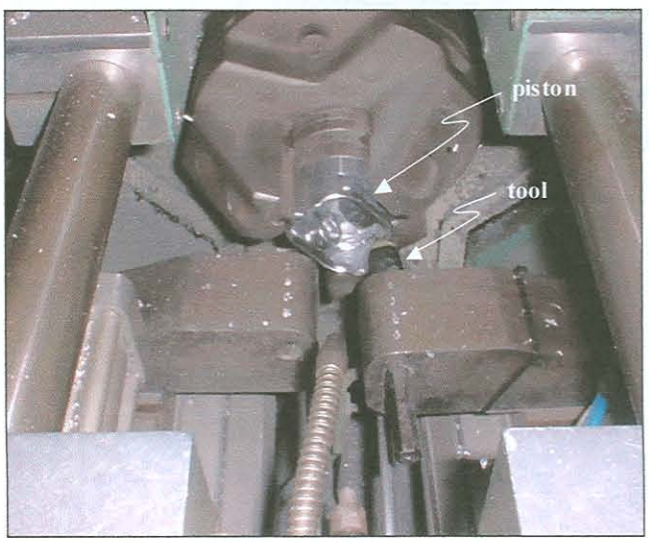


Figure 5.23: Sensor-integrated tool in operation (1)

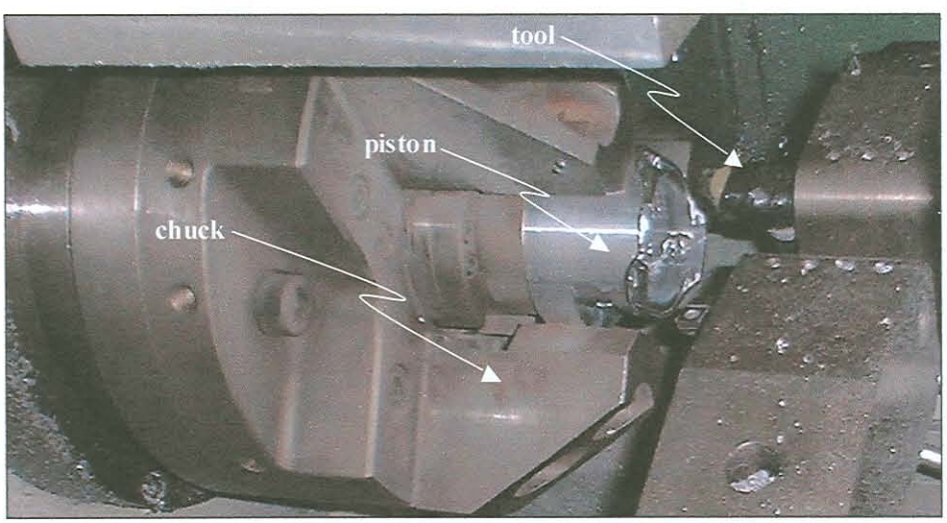


Figure 5.24: Sensor-integrated tool in operation (2)

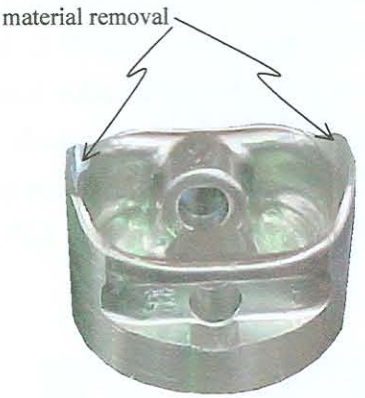


Figure 5.25: Workpiece

5.5 Tool wear

Before further analysis of the signals can be discussed, some remarks about the tool wear must be made. During the course of the research, data from almost 100 tool inserts were collected. This implies that force signals from the process were collected continuously from a new to a worn tool insert for many tools. The experience from the operators on the shop floor is that the tool wear is unpredictable. Sometimes a tool will last for thousands of components, and sometimes it will wear out after a few hundred. A conservative approach is used to eliminate the possibility of scrapping a part and tools are often replaced long before they have to be. The unpredictability of the tool wear was confirmed with wear measurements on the shop floor. A comparison of the flank wear of several tools with respect to the number of components that were machined with them is shown in Figure 5.26. It can be seen that the rate of tool wear is slightly different for every case, despite the fact that controllable conditions were kept constant.

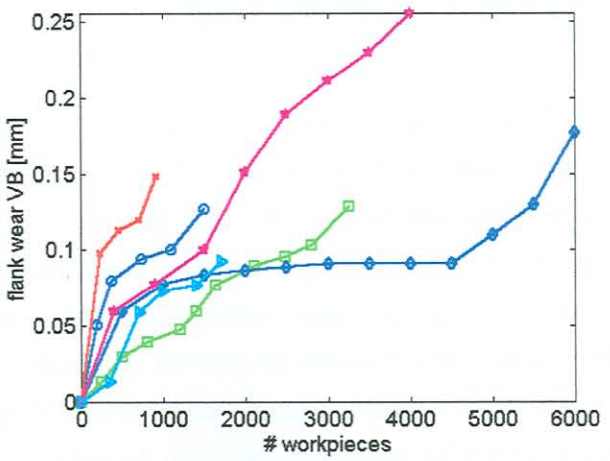


Figure 5.26: Tool wear with respect to number of machined workpieces

Reasons for this fluctuation in tool life can be attributed to conditions on the shop floor. More specifically, the rate at which components are manufactured plays a significant role. If the time allowed for the tool to cool down between runs is not kept constant, large variations in tool life can be expected. Fluctuations in the workpiece composition can also play a large role but is extremely difficult to quan-

tify these effects. A very significant conclusion from this is that tool life equations (*e.g.* modified Taylor equations [42]) cannot solve the problem in this case, hence justifying the need for an on-line TCMS. For all wear measurements and subsequent wear estimation, the average flank wear over a selected area of the cutting insert was chosen as a representative value of the tool condition. In order to obtain a closer understanding of the tool wear, Scanning Electron Microscope (SEM) pictures were taken of several used inserts. An SEM photo of a new insert is shown in Figure 5.27, compared to worn inserts in Figure 5.28.

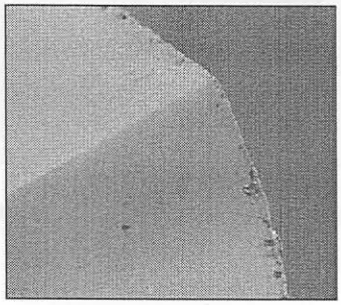


Figure 5.27: New insert

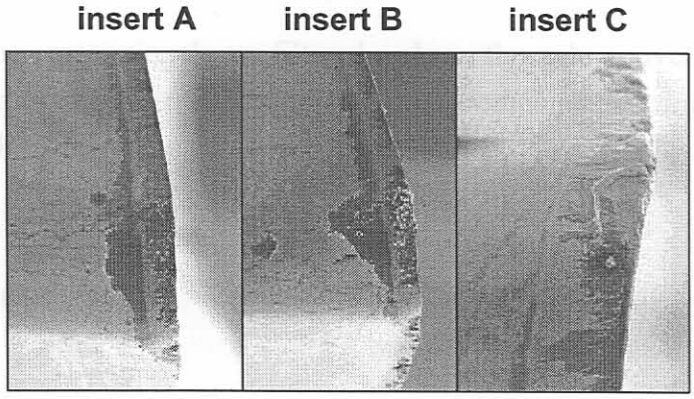


Figure 5.28: Worn inserts

Though the basic wear modes on the tools are similar, the geometry or shape of the wear land is different for each case. Similar observations were made with other worn tools under the SEM. From the SEM investigations it was concluded that the wear on the tools generally consisted of a plastic flow and an elastic contact region. There are also signs of micro-chipping and cracking at the edge, especially with severely worn tools. It was furthermore concluded that the chosen wear parameter of flank wear is adequate for describing the severity of wear on tool, because it has the most significant effect on the workpiece quality. Micro-chipping occurs at a late stage of the tool life, and it is virtually impossible to determine the effect of different severities of micro-chipping on part quality due to the large degree of flank wear also present. An example of the different wear modes on the tool is shown in Figure 5.29. The plastic flow and elastic contact zones form the total flank wear.

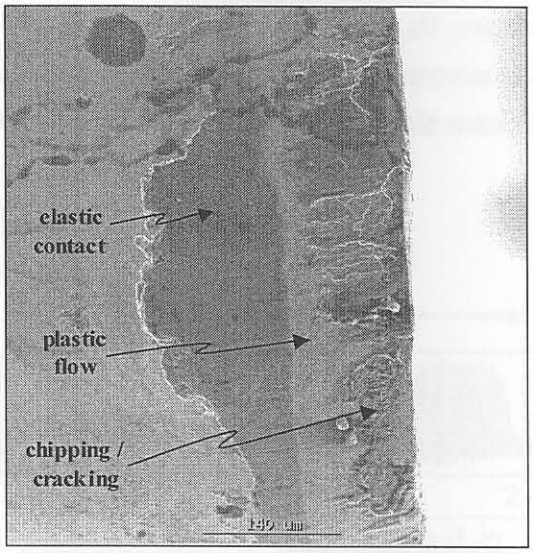


Figure 5.29: Different wear regions

5.6 Signal processing

5.6.1 Pre-processing

In the analogue form, the sensor signals are filtered and run through an overload protection unit. The aim of the digital pre-processing is to obtain reliable three-component cutting forces from the voltage values measured by the hardware. The following steps are taken:

- phase correction
- resampling
- DC offset compensation
- calibration

The phase correction is necessary due to the phase distortion caused by the analogue filter. Phase correction is done through an appropriate identification model of the filter, as discussed in Section 5.2.5. A high sampling rate must initially be used to prevent aliasing in the analogue format, but resampling makes the signals easier to process in digital format because of the lower file size. The process of resampling also involves digital low-pass filtering at half the new sampling rate. The signals were resampled to an effective sampling rate of 8kHz. The DC offset is removed with a reference signal collected when the tool is not engaged into the workpiece. A typical force measurement after the preprocessing step is shown in Figure 5.30. The cut essentially consists of two parts corresponding to a facing (A in the figure) and boring process (B in the figure). The feed direction is thus in the y-direction for A and in the z-direction for B. Due to the nature of the process, only section B of the signal contained information on the flank wear. The shape of the signals are due to the process, which is an interrupted cut with varying depth of cut and effective angles (resulting from the workpiece geometry).

The frequency of interruption is exactly 50Hz. It can be seen that this particular section actually creates positive and negative forces in the *Fy* direction, which is once again caused by convex shapes on the workpiece during the boring process. The cutting forces can now be analysed digitally to assist in the

design of a strategy to accurately monitor the process, with all possible internal and external disturbances. The system was set to store approximately 100 measurements during the life of a single tool. This would represent about 10% of the total number of possible measurements, but is still a sufficient amount of data.

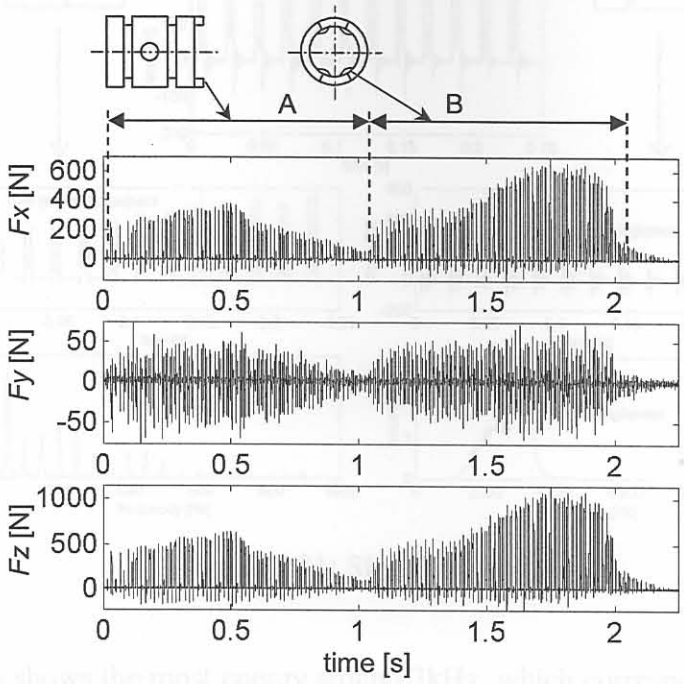


Figure 5.30: Typical cutting forces

5.6.2 Signal investigation

In order to identify possible signal features that might correlate with tool wear, several investigations were conducted. The signal basically consists of a low frequency region that is an indication of the static cutting forces. In the higher frequency range, the tool holder natural frequencies are observed because the cutting process excites them. Such a distinction of the two frequency ranges is shown in Figure 5.31. The spectrum in the lower range reveals many harmonics of the fundamental frequency of interruption, which is expected due to the square shape of the signal caused by the interrupted cut. Both these frequency ranges could contain information on the tool wear, but the type of analysis to be performed would differ for each range. Other initial signal investigation included time-frequency analysis to determine if there are certain frequencies increasing in amplitude with time. It is obvious that this will be the case due to the increasing and decreasing depth of cut. Hence, a particular section of the signal was selected for all subsequent analyses.

5.6.3 Feature extraction

The next step is to identify and extract features from the signals that are sensitive to tool wear. The static component of the force signal is represented in the 50Hz (and harmonics) component of the FFT. Thus, features representing the energy of the 50Hz component and its harmonics were generated as possible features. From initial investigations, it seemed that especially the 100Hz and 150Hz regions could be reliable features for TCM. A Cepstrum analysis was also carried out to see if the harmonics of 50Hz can be represented in a single parameter, but this did not yield a significant improvement.

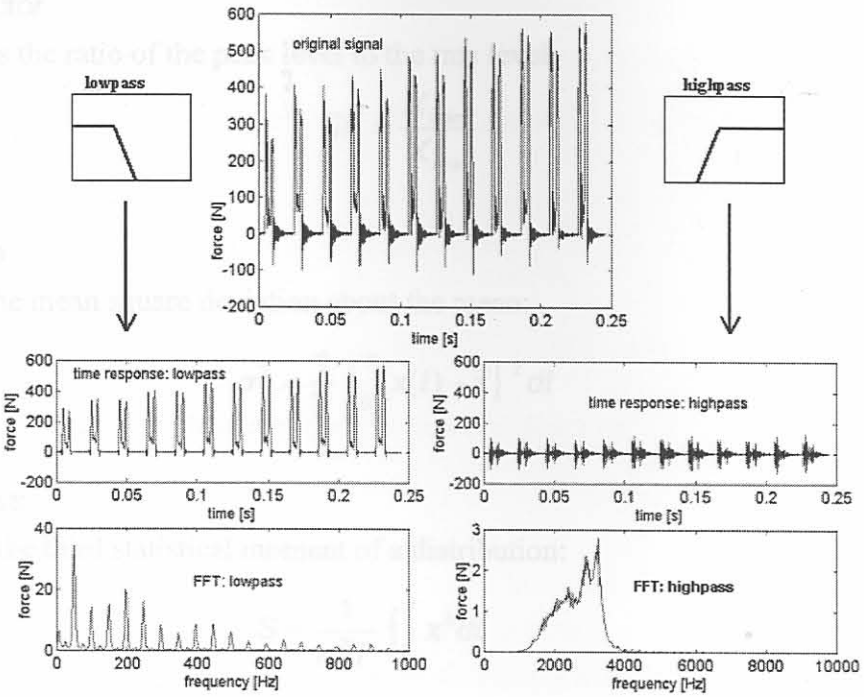


Figure 5.31: Signal character

The high-pass filtering shows the most energy around 3kHz, which corresponds to the tool holder cantilever natural frequency in the direction of *Fx*. These frequencies are often useful for TCM. The frequency spectrum was investigated in detail for frequency bands sensitive towards tool wear. This can be done by selecting certain frequency bands and calculating the energy contained within it. The energy in a frequency band can be expressed as:

$$\psi^2 = \int_{fl}^{fh} S_y df \tag{5.8}$$

with S_y the one-sided PSD function of the force signal and *fl* and *fh* chosen to reflect the energy in the regions of interest. Some simple time domain parameters were calculated as possible features, such as maximum and minimum values, peak-to-peak values, zero crossings and others. More common features often used for TCM applications in the literature were also calculated, which include:

A. Mean

The mean value of a function *x(t)* over an interval T is:

$$\bar{x} = \frac{\int_0^T x(t)dt}{T} \tag{5.9}$$

B. Root mean square (rms)

The rms value of a function *x(t)* over an interval of T is:

$$X_{rms} = \sqrt{\frac{\int_0^T x(t)^2 dt}{T}} \tag{5.10}$$

C. Crest factor

The crest factor is the ratio of the peak level to the rms level:

$$CF = \frac{X_{max}}{X_{rms}} \tag{5.11}$$

D. Variance

The variance is the mean square deviation about the mean:

$$\sigma^2 = \frac{1}{T}\int_0^T [x(t) - \bar{x}]^2 dt \tag{5.12}$$

E. Skewness

The skewness is the third statistical moment of a distribution:

$$S = \frac{1}{\sigma^3 T}\int_0^T x^3 dt \tag{5.13}$$

F. Kurtosis

The kurtosis is the fourth statistical moment of a distribution:

$$K = \frac{1}{\sigma^4 T}\int_0^T x^4 dt \tag{5.14}$$

G. Time series model features

Time series models can also be used to monitor a process, where the model coefficients are used as features. The model coefficients represent the characteristic behaviour of the signal. Depending on the order of the model, a number of model coefficients can be chosen. Normally only the first model coefficient, or sometimes the first three to four model coefficients are chosen because they are most descriptive of the signal [53,121,207]. Higher coefficients can actually become descriptive of noise within the signal, and therefore they are not preferred as wear monitoring features.

For the purpose of this research, the first coefficients from the Auto Regressive (AR) model, Moving Average (MA) model, and the Auto Regressive Moving Average (ARMA) model, were used as features for tool wear estimation. The models are calculated directly from the calibrated force signals. A brief discussion of each of the models follows [235-237]:

AR model

In a p-th order AR model for a time series $x(n)$, where n is the discrete time index, the current value of the measurement is expressed as a linear combination of p previous values:

$$x(n) = a_1 x(n-1) + a_2 x(n-2) + \ldots + a_p x(n-p) \tag{5.15}$$

where $a_1, a_2 \ldots a_p$ are the AR coefficients. The first AR coefficient was chosen as a feature.

MA model

In a q-th order MA model, the current measurement is expressed as a linear combination of q previous values from a sequence of Independent Identically Distributed (IID) random variables with a certain

probability density function.

$$x(n) = b_1 u(n-1) + b_2 u(n-2) + \ldots + b_q u(n-q) \tag{5.16}$$

where $b_1, b_2 \ldots b_q$ are the MA coefficients and $u(n)$ is assumed to be an IDD sequence. The first MA coefficient was chosen as a feature.

ARMA model

The ARMA model is a combination of the above two models:

$$x(n) = -\sum_{k=1}^{p} a_k (x(n-k)) + \sum_{k=1}^{q} b_k (u(n-k)) \tag{5.17}$$

The first two coefficients from this model were chosen as features.

5.6.4 Feature selection

A. Methods

With a list of possible features for TCM generated, a selection of the most reliable features must be made. This is one of the most important steps in designing a TCMS. There are various methods for feature selection and feature space reduction. In the case of feature space reduction, a method is applied that will reduce the dimensionality of the feature space. A popular way to achieve this is by means of Principal Component Analysis (PCA) [6]. The method of PCA calculates the eigenvectors of the feature matrix and re-aligns the feature matrix into its orthogonal dimensions. However, in the case of TCM, the input space (or the original feature matrix) is often one-dimensional because the best features tend to increase monotonically with tool wear. The PCA method was subject to further investigation and is discussed in Chapter 6.

Another method is to reduce the number of features by selecting the features that correlate the best with the objective (tool wear). There are a number of methods to achieve this. A very simple method is the use of the correlation function between the feature and the objective. If a particular feature displays a repeatable tendency towards high correlation with tool wear it can be selected as a feature for TCM. The correlation coefficient (expressed as a percentage) between the selected feature q and tool wear V can be calculated as follows:

$$\rho = \left| \frac{\sum_i (q_i - \bar{q})(V_i - \bar{V})}{\sum_i (q_i - \bar{q})^2 \sum_i (V_i - \bar{V})^2} \right| \times 100 \tag{5.18}$$

where $\bar{q}$ and $\bar{V}$ are the means of q and V, respectively; ρ is the correlation coefficient of which the value indicates linearity between q and V. When ρ is approaching 100%, there exist a relationship between q and V. The lower the value of ρ, the lesser the chance for the selected feature to show any trend towards tool wear.

Another method that was investigated as a possible method for feature selection is by means of a Genetic Algorithm (GA). The GA is an algorithm often used for discrete optimisation problems. Its formulation is based on the principles of natural selection. A detailed discussion of GAs is beyond the scope of this text, but more information can be found in [238]. For the purpose of feature selection, an optimisation problem can be formulated with the objective to select the optimal population of features

to estimate tool wear. Furthermore, the features can be automatically ordered from the best to the worst choice.

Both methods for feature selection were used during the course of this research. The methods always yielded similar results, but the GA is computationally much slower than the simple correlation coefficient approach. If feature selection must be done on-line, the GA approach is probably not the best choice. Furthermore, a GA must always be run more than once. In fact, it must be run at least three times to verify the result. The simple correlation coefficient approach is a simple alternative that is also feasible for on-line implementation (refer to Chapter 6).

As a last step of feature selection, some engineering judgement is required. The reason for this is that the automatic methods will often select features that are too similar or dependant on one another, and thus not achieving the goal of proper sensor fusion. In this case, the rules for selecting features based on engineering judgement can be stated as:

- Select features from the static and dynamic parts of the signal
- Select features from the different force directions
- Use time and frequency domain features
- Features based on simple signal processing methods are preferred
- There should be a reasonable physical explanation for the behaviour of a feature with respect to tool wear

It should be mentioned here that features were not investigated for insensitivity towards changing machining parameters. It has been shown that machining parameters can be included in the method and hence the behaviour of a feature with respect to machining parameters is irrelevant, the only requirement being that the feature remains sensitive towards tool wear when a machining parameter varies. The chosen features remained sensitive towards tool wear irrespective of machining parameters.

B. Selected features

It was found that the data from the shop floor is extremely noisy and that most features do not display any correlation with tool wear. However, a list of possible features was identified using the methods described above. As a last step engineering judgement was applied in selecting the following features:

Table 5.3: Description of features

feature	*description*
Fxs	standard deviation of *Fx*
Fxd	spectral energy of *Fx* at mode one (approx 2kHz)
Fym	mean of *Fy*
Fy50	energy around 50Hz for *Fy*

A plot of the chosen features with respect to tool wear (interpolated) is shown in Figure 5.32. Note that the features are always normalised for modelling purposes. It is clear from the figure that although the features tend to increase with increasing tool wear, the trend is far from consistent. There is a large degree of variance in the feature trend and the correlation with the true tool wear is rather poor.

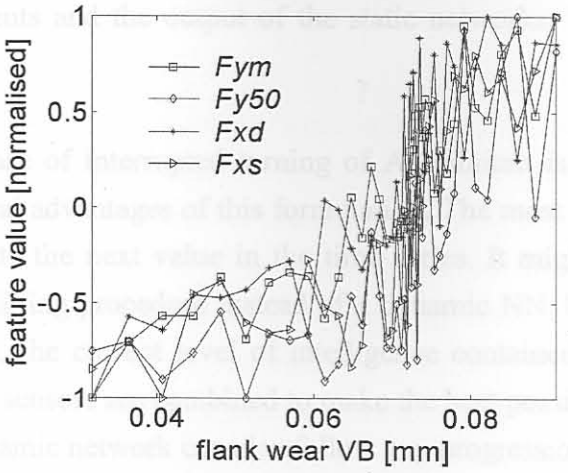


Figure 5.32: Normalised features

Some physical explanations for the increasing trend in the features are necessary. The standard deviation of *Fx* represents the vibration energy contained within the signal. In this case, the static cutting force will dominate this value, and one can conclude that the standard deviation increases due to an increase of the static force and to a lesser extent an increase in higher frequency vibration. The forces increase due to an increase of friction between the tool tip and workpiece when the tool flank is worn. The dynamic parameters in this case are governed by the harmonics of the interruption frequency and also by the dynamic behaviour of the tool holder in the *Fx* direction. The excitation from the cutting process causes the holder to vibrate freely at its natural frequency, and the vibration amplitudes in this range normally increase with increasing tool wear. The increase is once more caused by the larger impacts on the tool tip due an increase of friction between the tool and the workpiece. The parameters generated from the *Fy* force are both more representative of the static cutting forces, and the same arguments apply. Some information about the vibration energy in the *Fy* direction is contained in the *Fym* feature.

Due to the high degree of variance in the increasing trends of the features, a monitoring strategy based on any one of the features alone will never yield an accurate estimation of the tool wear. For this reason, an AI approach is used to estimate these values. The approach is exactly the same as that applied to hard turning in Chapter 4. The selected four parameters are sufficient to monitor tool wear because there are no other changes to the process, *e.g.* workpiece or tool type. If such disturbances were present, it would in all probability be necessary to include other features as well. During the experiments, two different feed rates were used, but the selected features applied for both feed rates.

5.7 AI method for TCM

The tool wear can be classified by using the selected features as inputs to an AI model. The principle of the AI approach proposed in this work relies on the use of a series of NNs, one of which is trained on-line and others that are trained off-line. The off-line networks are static networks that are trained to model the feature values for known values of tool wear and cutting conditions. The on-line network is a dynamic network that attempts to estimate the current wear on the cutting edge by using the previous estimations of tool wear as inputs. The training goal for the dynamic network is to minimize the error

between on-line measurements and the output of the static networks. This is the same method presented in Section 4.6.3.

The AI approach for the case of interrupted turning of Aluminium is diagrammatically depicted in Figure 5.33. There are several advantages of this formulation. The most important is in the use of temporal information to estimate the next value in the time series. It might seem possible to achieve a similar result with a curve fitting procedure instead of a dynamic NN, but the problem is much more complex. With this method, the current level of intelligence contained within the dynamic NN and knowledge from the on-line sensors are combined to make the best possible decision about the severity of wear on the tool. The dynamic network can also follow any progression of tool wear.

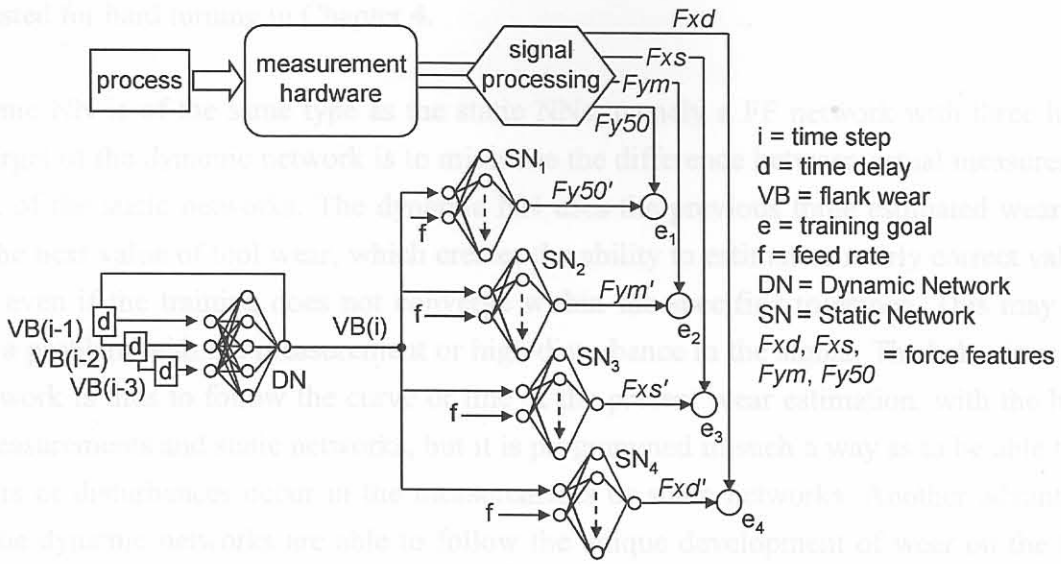


Figure 5.33: Monitoring strategy

5.7.1 Static networks

The static networks were trained, validated and tested for each of the chosen features. The data from eight cases of tool wear from a new to a worn tool insert was used as the training set. Two different cases of feed rate were considered and hence the feed rate was included in the static NN. The feed rate was excluded from the dynamic NN, with the assumption that the feed rate will only be changed with a new insert. Four relatively small static networks are used, all of which are FF networks with three layers. The middle layer consists of five 'tansig' neurons, and the output neuron has a linear activation function. The 'tansig' function is a combined tan and sine function. The static networks were trained with Levenberg-Marquardt backpropagation.

One of the most important considerations when training NNs is to prevent overtraining. This will cause the networks to memorise the training data and as a result they not be able to generalise when they are presented with new data. In this case, a combination of using small networks and early stopping were used to prevent this effect.

5.7.2 Dynamic network

With the static networks trained, the dynamic network is trained on-line to estimate the on-line value of the tool wear *VB(i)*. One of the main reasons why this approach is so efficient for TCM applications is the fact that tool wear almost never follows the same geometry and growth rate. If the static networks are trained appropriately, the dynamic network can follow any growth and geometry of tool wear. The dynamic network estimates the unknown parameter, namely tool wear, by using recent information obtained from the process. For this application, machining parameters did not change during the life of a tool and were consequently not included. However, any changing parameter that has a significant influence on the features can be included, for instance cutting speed, workpiece material or working angles. Another wear mode can also be included, for example flank wear and crater wear, as was suggested for hard turning in Chapter 4.

The dynamic NN is of the same type as the static NNs, namely a FF network with three layers. The training target of the dynamic network is to minimise the difference between actual measurements and the output of the static networks. The dynamic NN uses the previous three estimated wear values to estimate the next value of tool wear, which creates the ability to estimate a nearly correct value for the tool wear even if the training does not converge within the specified tolerance. This may happen if there was a problem with the measurement or high disturbance in the signal. The behaviour of the dynamic network is thus to follow the curve or line of the present wear estimation, with the help of the on-line measurements and static networks, but it is programmed in such a way as to be able to generalise if errors or disturbances occur in the measurements or static networks. Another advantage is the fact that the dynamic networks are able to follow the unique development of wear on the tool insert through on-line training (also refer to Section 4.6.3).

The method utilises sets of inner and outer steps or time-increments. The inner steps are training steps of the dynamic NN to achieve a specified convergence. Hence, during the inner steps, the tool wear is assumed constant and the NNs attempt to estimate this value. When this is achieved, an outer step is taken, and in this case it is an incremental step in the tool wear.

The problem can be described by considering a vector $\mathbf{x}$ containing the network bias and weight values of the Dynamic Network (DN):

$$\mathbf{x} = [x_1\ x_2 \ldots x_n] \tag{5.19}$$

In order to increment an outer step, the following optimisation problem must be solved, which is the training goal of the DN:

$$\text{minimise } f(\mathbf{x}) = \sum_{j=1}^{4} e_j \text{ such that } f(\mathbf{x}) \le tol \tag{5.20}$$

with the initialisation space for a new tool starting at:

$$D = \left\{ (x_1 \ldots x_n) \in \Re^n : -1 \le x_i \le 1,\ i = 1..n \right\} \tag{5.21}$$

and *tol* a suitable convergence tolerance on the function value. The initialisation space for a worn tool is obtained from the solution of the previous outer step. The error functions in equation 5.20 are de-

fined as:

$$e_1 = \sqrt{(Fy50' - Fy50)^2}$$
$$e_2 = \sqrt{(Fym' - Fym)^2}$$
$$e_3 = \sqrt{(Fxs' - Fxs)^2} \qquad (5.22)$$
$$e_4 = \sqrt{(Fxd' - Fxd)^2}$$

*Fy*50', *Fym*', *Fxs*' and *Fxd*' are the outputs of the static networks $SN_{1...4}$. When the DN reaches its training goal, an outer step can be taken and new values for *Fy50, Fym, Fxs* and *Fxd* can be measured using the on-line hardware and software. Note that all variables are normalised before they are entered into the NNs, and again de-normalised at the network output for interpretation of the results.

Similar to Chapter 4, it was found that the Particle Swarming Optimisation Algorithm (PSOA) yielded the best result for training the DN. Training NNs is in essence an unconstrained global optimisation problem, and recent literature also states that the PSOA outperforms methods like Genetic Algorithms (GAs) in unconstrained global optimisation. Furthermore the method is easy to implement and computationally efficient. The PSOA simulates the physical movement of social creatures, for instance the movement of a flock of birds. The algorithm does not utilise gradient evaluations and its efficient use of random information is advantageous for training NNs. A detailed formulation of the training algorithm is included in Appendix D. The algorithm usually achieved convergence within less than about 10 inner steps in a matter of seconds on a PIII computer. The convergence criterion was set on the function value. If convergence is not achieved within 10 steps, the training also terminates.

5.7.3 Discussion

It should be kept in mind that the aim of NN modelling should never be to make estimations outside the range of training. A NN should always be properly trained for its range of application. The most important part of the training is to ensure proper generalisation capabilities of the network. The network structure should be optimised to ensure that the network size and functions are chosen correctly for the application. In the case of the network size, iterations with different network sizes were used to choose the optimal network size. The result was quite small networks with adequate generalisation capabilities with respect to the problem. The network type and functions were selected by using proper judgement and experience with different types of NNs. The same arguments apply to the static and the dynamic networks. Further discussions can be found in Chapter 6.

5.8 Tool wear estimation results

5.8.1 From a new to worn tool insert

Results from the AI method with on-line estimations with previously unseen data for four tools is shown in Figure 5.34. In each case, the estimation starts with a new tool and ends when the operator removed the tool from the machine. The estimations are also compared with flank wear measurements taken under a microscope. It can be seen that there is a very good agreement between the estimation and the actual value of tool wear. Disturbances on the features also cause the estimated growth of tool

wear to appear somewhat noisy, but much less than that compared to the features. In two of the cases the flank wear exhibits the typical *initial fast* and *regular* wear stages. The tool is normally removed before the *accelerated final* wear stage and subsequent tool breakage is reached. However, it can be seen that the progression of tool wear, or the wear growth rate, is different for every case.

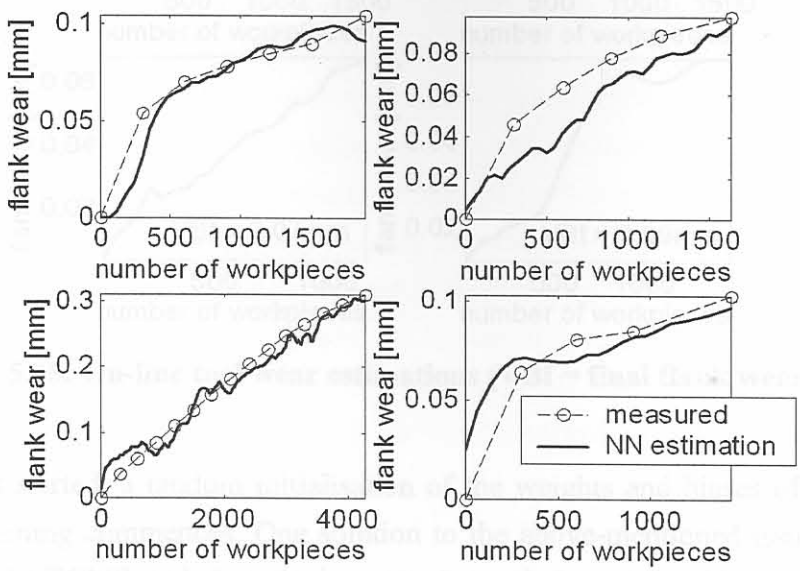


Figure 5.34: Tool wear estimations for four tools

The output of the dynamic NN is reported in terms of a "sliding output window" which means that an average value is taken that moves along with the outer steps. It was found that taking three averages is adequate. The usage of a sliding output window also helps to eliminate the possibility of making a decision based on only a single measurement that may contain errors. Many other examples of tool wear estimations are available, but these can only be compared to the final value of tool wear, because regular tool wear measurements could only be made for a limited number of tools. Four examples of online tool wear estimations are Figure 5.35. These can be compared to the final value of tool wear, and again a very good agreement is visible. The same observations regarding the growth rate of the flank wear can be made from these graphs.

5.8.2 Previously worn insert

In the previous section it was shown that the dynamic NN could estimate the tool wear accurately for a new insert that is allowed to wear normally. It should be kept in mind that the dynamic NN would tend to estimate the next value in the series of the previous three values of tool wear. The question then arise how the dynamic NN would react in the following two cases:

- If the tool is changed from a severely worn insert to a new insert
- If a worn insert is used from the start

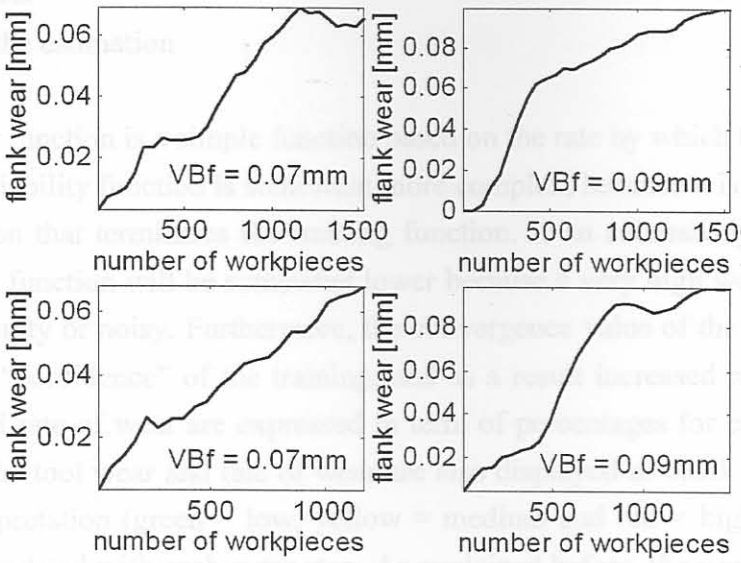


Figure 5.35: On-line tool wear estimations (VBf = final flank wear value)

When the TCMS is started, a random initialisation of the weights and biases of the dynamic NN is used after which training commences. One solution to the above-mentioned issues could be that the operator must reset the TCMS each time the insert is changed. A reset would cause the random initialisation of the dynamic network. Another solution is automatic detection of a new insert. Checking the output of the dynamic NN to see if it estimates a much lower value than before, towards 0 mm flank wear, could do this. The automatic detection proved to work very well. However, a manual reset is also included in the final implementation. An example of an estimation with the automatic detection and reset during a tool change with a new insert is shown in Figure 5.36a). When a worn insert is changed with a worn insert, the system must also react to quickly converge to the correct value of the tool wear. An example of a tool change with a worn insert is shown in Figure 5.36b). It is clear from the figure that the dynamic NN only requires a few steps to converge to the correct value of tool wear.

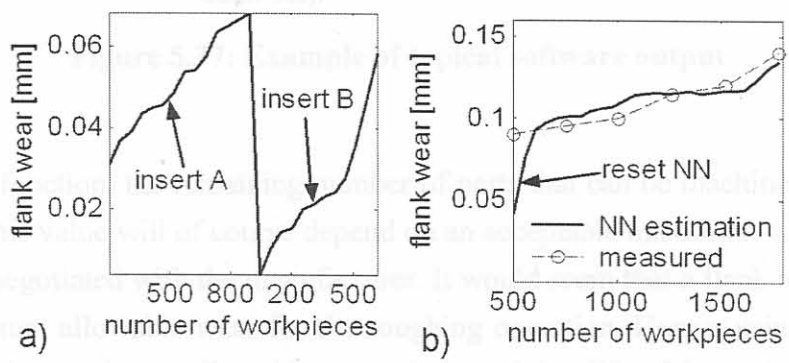


Figure 5.36: Dynamic NN response

5.8.3 Accuracy and reliability

A very important aspect of this development is the accuracy and reliability of the complete system. In terms of the reliability of the software, it was decided to also implement two additional functions that the end user can easily interpret when making a decision about the tool wear. These two functions are:

- rate of tool wear
- reliability of the estimation

The rate of tool wear function is a simple function based on the rate by which the estimated wear value is increasing. The reliability function is somewhat more complex, because it is based on the rate of tool wear and the criterion that terminates the training function. If an excessively high wear rate is estimated, the reliability function will be somewhat lower because a very high wear rate can be estimated if the input data is faulty or noisy. Furthermore, the convergence value of the training algorithm gives an indication of the "confidence" of the training, and as a result increased reliability. The reliability value, tool wear, and rate of wear are expressed in term of percentages for easy interpretation by the machine operator. The tool wear and rate of wear are also displayed as block diagrams and are colour coded for easy interpretation (green = low, yellow = medium and red = high). The percentages and block diagrams are updated with each outer step. As explained before, the user can select the tempo of outer steps with one outer step per workpiece the maximum. An example of the on-line software output is shown in Figure 5.37. Typically, when "flank wear" reaches 80% deviations in product quality can be expected.

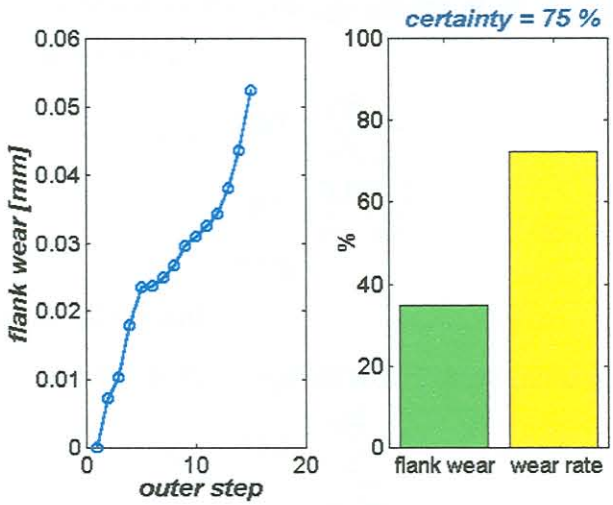


Figure 5.37: Example of typical software output

Using the wear rate function, the remaining number of parts that can be machined with the tool insert can be calculated. This value will of course depend on an acceptable maximum value of tool wear, and this will have to be negotiated with the manufacturer. It would seem that a flank wear of up to 0.25mm would be the maximum allowable wear for the roughing operation. Hence, using the estimated wear rate and 0.25mm as the maximum allowable wear, the remaining life of the tool can be calculated and updated with each outer step.

The accuracy of the tool wear estimations can be compared to the full range of observed flank wear, or to the current range of the particular tool. Furthermore, the average accuracy of the system can be compared to the minimum accuracy, thus the average deviation from the true tool wear and the maximum deviation. The following definitions can be made to describe the accuracy of the TCMS:

$$\text{Full range of flank wear } VB_{fr} = 0.31\text{mm} \tag{5.23}$$

$$\text{Maximum deviation from true wear } VB_{md} = 0.052\text{mm} \tag{5.24}$$

$$\text{rms deviation from true wear } VB_{rms} = 0.085\text{mm} \tag{5.25}$$

$$\text{1\% of full range } VB_{fr1\%} = {VB_{fr}}/{100} = 0.0031\text{mm} \tag{5.26}$$

These values were obtained from the results of all the data.

The rms deviation can be described as:

$$VB_{rms} = \frac{\sum_{i=1}^{n}\sqrt{\left(VB_{TCMS} - VB_{true}\right)_i^2}}{n} \tag{5.27}$$

where

n = the total number of tool wear measurements

VB_{true} = the measured tool wear

VB_{TCMS} = the wear estimated by the monitoring system

The following calculations are based on the average results of all the experiments. The average accuracy with respect to full range of wear is:

$$\begin{aligned}\eta_{Afr} &= 100 - \frac{VB_{rms}}{VB_{fr1\%}} \\ &= 100 - \frac{0.0085}{0.0031} \\ &= 97\%\end{aligned} \tag{5.28}$$

where $VB_{fr1\%}$ is one hundredth of the full range.

The minimum accuracy with respect to full range of wear can be calculated as:

$$\begin{aligned}\eta_{Mfr} &= 100 - \frac{VB_{md}}{VB_{fr1\%}} \\ &= 100 - \frac{0.052}{0.0031} \\ &= 83\%\end{aligned} \tag{5.29}$$

The average accuracy with respect to current range of wear is somewhat more complex but can be calculated with:

$$\begin{aligned}\eta_{Acr} &= 100 - \frac{1}{n}\left(\sum_{i=1}^{n}\frac{\sqrt{\left(VB_{TCMS} - VB_{true}\right)_i^2}\times 100}{\left(VB_{true}\right)_i}\right) \\ &= 92\%\end{aligned} \tag{5.30}$$

The minimum accuracy with respect to current range of wear can be written as:

$$\begin{aligned}\eta_{Acr} &= 100 - \frac{VB_{md}\times 100}{VB_{true}} \\ &= 100 - \frac{0.052\times 100}{0.16} \\ &= 100 - 32.5 \\ &= 68\%\end{aligned} \tag{5.31}$$

From the accuracy calculations the conclusion can be made that the system will estimate the true tool wear with an accuracy of above 92% as an average, and in extreme cases with accuracy not lower than 68%. The extreme cases rarely happen and in these cases the reliability index on the computer implementation will advise the user that the estimation could be inaccurate. It is important to mention here that the approach of measuring and monitoring was tested under shop floor conditions with different clamping conditions of the tool and workpieces over several months. Furthermore, the machines was completely stripped and serviced during tests. Different sensor-integrated tools were tested and none of these changes had a significant influence on the system. It can thus be concluded that the method is insensitive towards such changes and can be expected to always operate within the calculated range of accuracy.

5.9 Conclusion

This chapter described the implementation of a TCMS on the shop floor using AI. It was shown that the technique proposed is very effective for estimating the flank wear of tool inserts using features derived from strain gauge measurements. Features representative of tool wear are generated from both static and dynamic parts of the force signals. A combined static and dynamic NN technique was used with the PSOA method to train the DN. The method proved to be a very effective algorithm for training the DN. The system was developed and tested in a production environment, proving that such systems could run effectively on shop floor situations, despite the many disturbances present on the shop floor. An advantage of the system is its cost-effectiveness, due to the use of simple sensors and electronics. Another advantage of this approach is the fact that the tool wear is not accelerated, but it is allowed to run through its normal life, and hence it describes a normal wear pattern. It was also shown that the system is exceptionally accurate in estimating the true wear on the cutting tool. This approach proposed here can be modified to include any other parameters that might influence the features or growth of tool wear, and can also be applied to other machining operations if appropriate features can be generated. The method is not sensitive towards the clamping condition of the tool or other vibrational effects of the machine tool.

CHAPTER 6

6. Further investigations towards improvement of AI method

6.1 Introduction

In the previous chapters it was shown how a tool wear monitoring system can be designed using AI. The formulation of the AI method is such that it is not sensitive to the noisy shop floor conditions and other typical influences. The method was shown to be applicable to different turning operations and with different machining parameters. An important component of the research was that it was shown that a TCMS using AI could be implemented on the shop floor in a cost-effective way. The purpose of this chapter is to discuss some further results into the improvement of the TCMS. The investigations in this chapter are mainly concerned with improving the AI, signal processing and feature selection methods.

6.2 Signal processing

6.2.1 Coherence function

The coherence function is commonly used as a measure of the integrity of frequency response functions in vibration analysis. It can also be used to detect non-linearities in a system with an input-output relationship. Coherence is a function of frequency with values between 0 and 1 that indicate how well input corresponds to the output at each frequency. The coherence between an input X and an output Y is defined as:

$$\gamma^2 = \frac{\|P_{xy}\|^2}{P_{xx}P_{yy}} \qquad (6.1)$$

where:

P_{xx} = the power spectral density of x

P_{yy} = the power spectral density of y

P_{xy} = the cross spectral density of x and y

It follows that if x and y are completely correlated at a particular frequency the coherence will be 1 at that frequency. In practice, a value near one is reached for correlation and a value near zero is reached when there is no correlation. The coherence is typically calculated for a force input with a vibration output. However, the coherence can be calculated for any two sensors, and the coherence function can assist as a feature for condition monitoring. The coherence function between two acceleration signals was used by Li *et al.* [109] to detect tool wear and chatter during turning. They established that the value of coherence at certain frequencies could be used to trend tool wear or to detect the onset of chatter. Because the coherence is easy to calculate and interpret, it might be useful for on-line implementation, and it was decided to investigate the coherence function as a possible feature for TCM.

If force and vibration measurements are present, the coherence function is particularly useful to investigate the dynamics of the system. In Chapter 4, an accelerometer was attached to the tool holder dynamometer, as shown in Figure 6.1. The coherence function can now be used to determine the coherence in the response of the dynamometer over a range of frequencies. The coherence functions between force and acceleration in the three principal directions for various experiments with increasing tool wear was investigated over the range 8Hz – 2kHz.

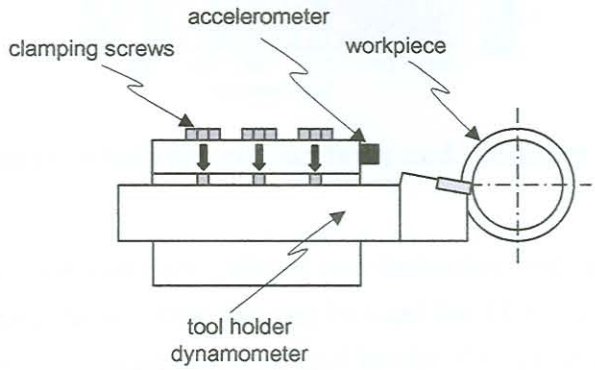


Figure 6.1: Sensor configuration (Chapter 4)

Examples of the coherence functions for experiment 2 are shown in Figure 6.2 and Figure 6.3. The vertical red lines on the figure indicate a high level of coherence. It is interesting to see that there are in fact quite a few areas in both figures where the coherence is very low, especially in the low frequency range. There are several possible reasons why coherence could be low:

- noise in the measurements
- resolution bias errors in the spectral estimates
- the output is due to other inputs besides the measured input
- non-linearity of the system

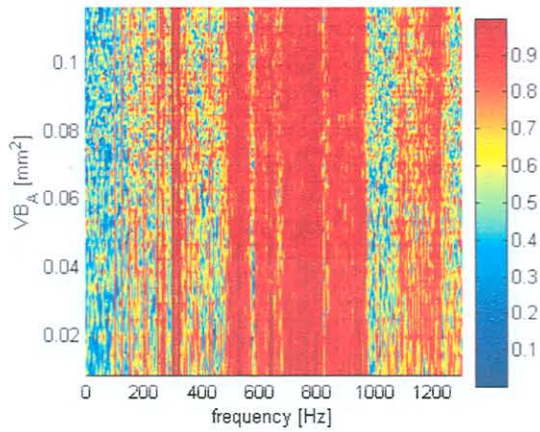


Figure 6.2: Coherence between thrust force and vibration (hard turning)

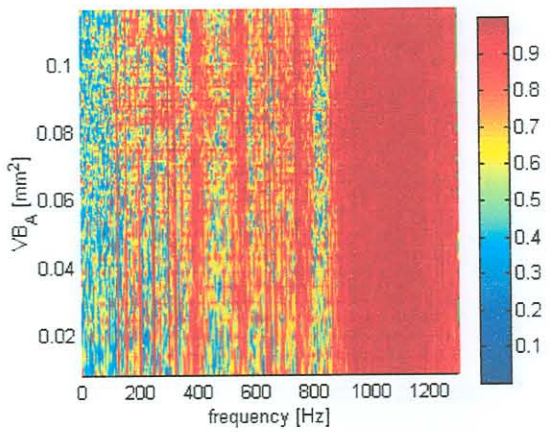


Figure 6.3: Coherence between cutting force and vibration (hard turning)

The areas where coherence is low probably indicate non-linearities and non-measured inputs. The areas with high coherence indicate frequencies that can be used for TCM because they are due to the actual dynamic force inputs and the dynamics of the tool holder. Closer investigations revealed that the **average** coherence function in the cutting force direction shows a slight increase due to tool wear, but would not suffice as a feature for TCM. There were also no particular frequencies that showed a significant correlation with tool wear.

The investigation was then moved to the coherence between two acceleration signals in the different directions, hence the method proposed by Li *et al.* [109]. Some of the results from experiment 2 are shown in Figure 6.4 and Figure 6.5. Again, there are certain regions where the coherence is very low, but in this case it is expected because different directions are used. However, in contrast with the findings in [109], there were no frequencies in the different coherence functions with a useful agreement with tool wear. From this it can be concluded that the usefulness of the coherence function is limited to the particular experimental setup, and is thus not a general solution for TCM.

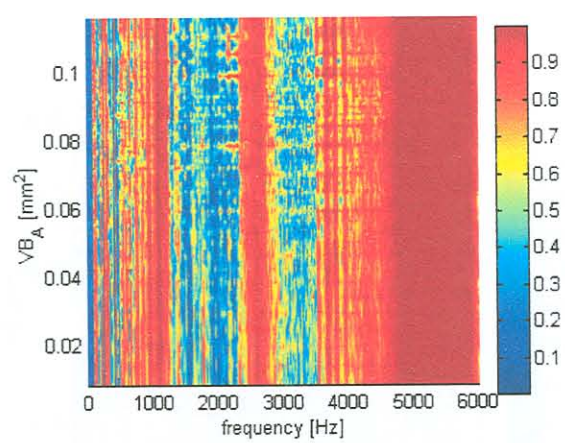


Figure 6.4: Coherence between feed and thrust vibration (hard turning)

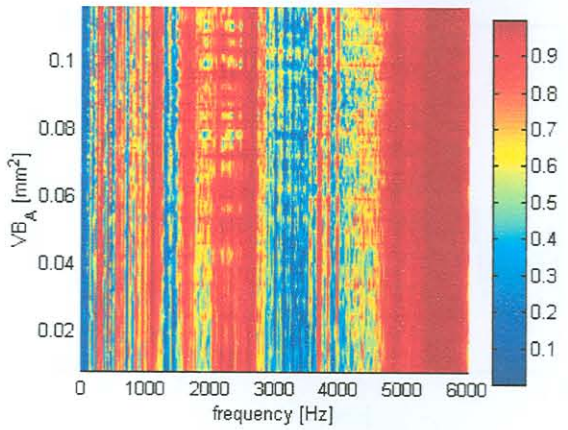


Figure 6.5: Coherence between cutting and thrust vibration (hard turning)

The results from the coherence function investigations with the data from Chapter 4 also raised an important question – does the dynamometer measure the actual dynamic force inputs or is it measuring its response to these inputs, which may contain non-linearities within the measuring range? This can only be answered if proper dynamic tests are conducted. However, the use of resistance strain gauges (Chapter 5) offered a cost-effective alternative that had no effect on the dynamic characteristics of the system. The dynamic properties of the sensor-integrated tool holder were characterised in Section 5.2.3. Consequently, coherence functions were also calculated on the force data from the strain gauges for the Aluminium turning experiments. A result for the coherence between forces *Fx* and *Fy* (refer to Chapter 5) is shown in Figure 6.6. It can be seen that in this case, the coherence is high for most of the frequencies. However, the coherence is influenced by the tool wear in the lower frequency range up to about 800Hz, and also at the tool holder natural frequency near 3kHz. In this case, increasing tool wear generally cause lower coherence values, except for the very low frequencies near 0 Hz where an increase is present. The mean value of the coherence over the whole frequency with respect to tool wear is shown in Figure 6.7. In this case the coherence could be used as a feature for TCM. Further investigations revealed that the features chosen for wear monitoring in Chapter 5 are more reliable than the coherence function.

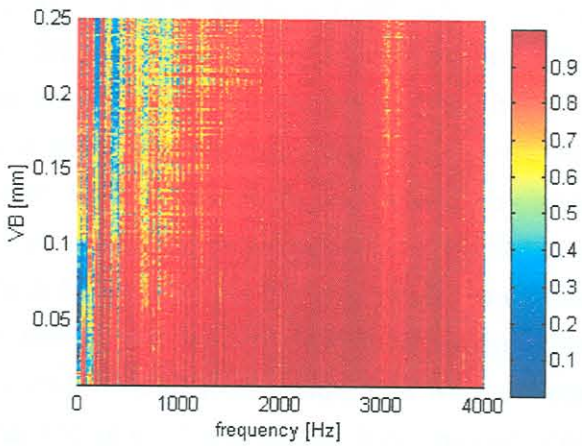


Figure 6.6: Coherence between Fx and Fy (Aluminium turning)

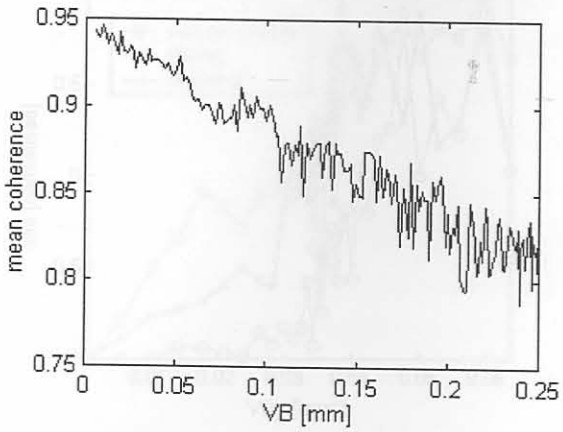


Figure 6.7: Mean of the coherence function between *Fx* and *Fy* (Aluminium turning)

6.2.2 Wavelet analysis

As a further step towards generating more reliable features for wear monitoring, wavelet analyses were conducted on the data sets. Some authors state that wavelet analysis is the key to successful TCM (refer to Section 3.4.2). However, the usefulness of wavelet analysis in TCM applications is debatable. Wavelet analysis cannot be used effectively to detect temporal information in data because they are not time invariant. In the case of TCM, this is not always important, depending on the aim of monitoring system. They only way that wavelet analysis can really assist in TCM applications is to act as a filter to enhance the signal-to-noise problem common to tool wear data.

The reader is referred to Appendix F for a more detailed discussion on wavelet analysis. During this research, wavelet packet analyses were performed with various types of wavelets to act as digital filters. With the wavelet packets calculated, certain packets may contain useful information on tool wear with little noise present. A simple method is to select the packets containing the most energy by means of an energy function (*e.g.* rms or entropy) for a second phase of processing. The second phase is generally a repeat of the initial signal processing in the time and frequency domains, such as rms, variance and crest factor values.

The purpose of the wavelet analysis is to act as a filter, which automatically identifies certain frequency ranges and bandpass filter the signal through these ranges. When the signal is reconstructed with the wavelet coefficients from the selected wavelet packets, it can be compared with the original signal. During this research, it was often observed that the reconstructed signal is just a filtered representation of the original. Hence, the same result can be obtained by using an appropriate digital filter. In fact, the wavelet packet type of filtering is a "black box" type filter because there is no indication beforehand or afterward which frequencies are attenuated. In the case of digital filtering, this can be controlled and explained. A result from the wavelet packet analysis from experiment 4 is plotted in Figure 6.8. In this figure, the rms trend with increasing tool wear are compared for a selected wavelet packet, a digital filtered signal and an unfiltered signal. It can be seen from this figure that although the wavelet packet rms trend is more consistent than the unfiltered signal, the digital filter yielded the best result. A similar result can be observed in Figure 6.9, where the trends of the variance with increasing tool wear are compared.

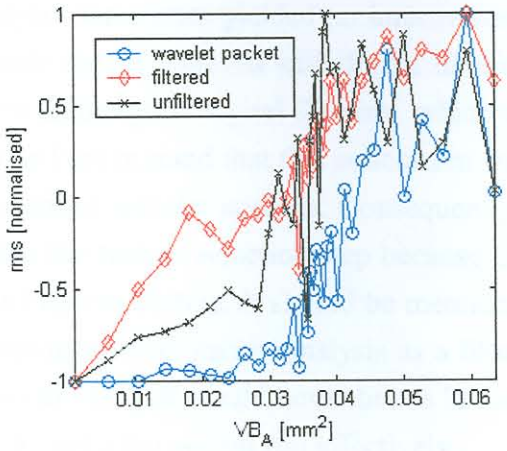


Figure 6.8: Comparison of rms trend with tool wear (hard turning)

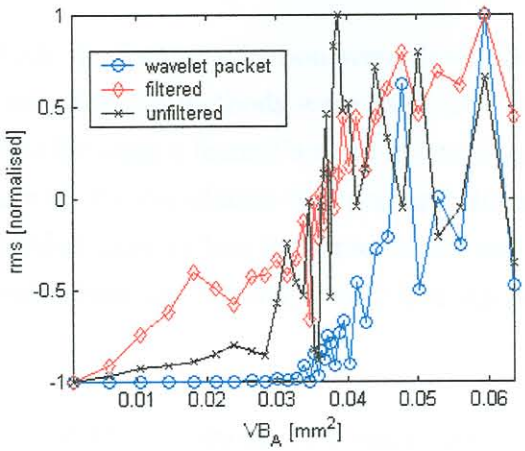


Figure 6.9: Comparison of variance trend with tool wear (hard turning)

The wavelet analysis was also conducted on the TCM data for Aluminium turning. Exhaustive investigations were conducted with different types of wavelets and threshold values. One of the results is shown in Figure 6.10, where the rms values of the best wavelet packets are compared to the rms of the digital highpass filtered signal and unfiltered signals. In this case, it is again clear that the digital filter yields the best results.

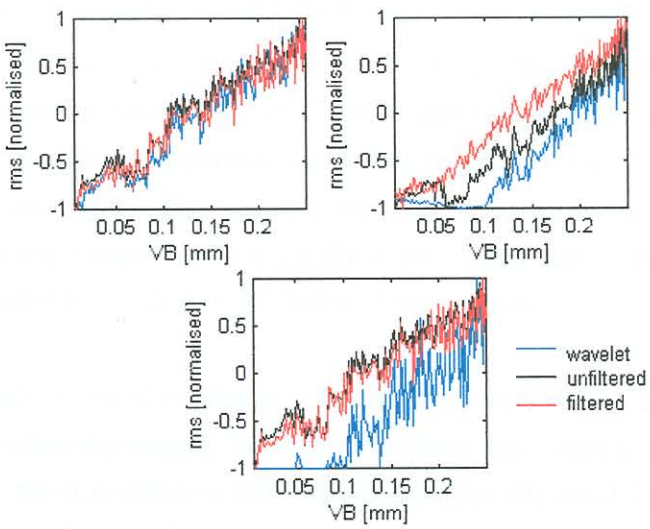


Figure 6.10: Wavelet packet rms comparison (Aluminium turning)

It was found that wavelet analysis sometimes yielded an improvement in the trend, but the improvement is small and the same result can be achieved with digital filtering. Furthermore, wavelet analysis is computationally much slower than regular digital filtering, which hampers it as a possibility for on-line implementation. It should be kept in mind that this conclusion was only reached for TCM applications and not for other applications of wavelet analysis. Consequently, the features generated by wavelet analysis were discarded from the feature selection step because the improvement was too small to make it worthwhile for on-line implementation. It should be mentioned here that there exists a TCMS for ball-end milling which is using wavelet packet analysis as a filter, and the method is successfully running as a demonstration system [239]. It would nevertheless be interesting to investigate the necessity of wavelet packet analysis to make the system run effectively.

6.2.3 Feature selection and feature space reduction

In Chapters 4 and 5 the methods for feature selection were described. A combination of engineering judgment and automated feature selection methods were used. The automated feature selection methods were based on the agreement between a feature's trend and tool wear. The methods were employed to quantify this agreement, namely the correlation function and difference minimisation by means of GAs. Both methods usually yielded more or less the same results and as a last step engineering judgement was used to select particular features. Normally, only four signal features were chosen for monitoring purposes.

Another method of verifying that the best features have indeed been selected is by calculating the Statistical Overlap Factor (SOF). The SOF determines the degree of separation of a feature between the new and worn tool conditions. Ideally, a feature should show a high degree of separation due to the worn condition and a low degree of variance due to noise. The SOF is one method that can assist in investigating the behaviour of a feature. The SOF is defined by:

$$SOF = \left| \frac{\bar{x}_1 - \bar{x}_2}{(\sigma_1 + \sigma_2)/2} \right| \tag{6.2}$$

where $\bar{x}_1$ is the mean and σ_1 is the standard deviation of vector x_1. Vector x_1 should be data collected from new tools and x_2 should be data collected from worn tools. As a consequence the SOF will yield a value that is an indication of a feature's ability to separate between new and worn conditions. Comparisons between the correlation coefficient approach and the SOF were made for several experiments (correlation coefficient approach described in earlier chapters: The correlation is calculated between tool wear and the feature vectors). One result is shown as a bar graph comparison in Figure 6.11 for 30 different features calculated from the Aluminium turning data. Ideally, features should be chosen that exhibit a high degree of SOF and correlation.

One problem with the SOF method is that it will not separate between linearly dependant features. Thus, there is the risk of selecting linearly dependant features when only the SOF method is employed. Another method that can assist to remove linear dependency is Principal Component Analysis (PCA). However, it is known that TCM problems are often one-dimensional (all the signal features increase when tool wear increase), and thus features that are linearly dependant to some degree will have to be

selected. Consequently, as a last step, engineering judgement must be used to select features that are not completely linearly dependant and also have high SOF and correlation values. Referring to Figure 6.11, features 7, 14, 17 and 25 were selected.

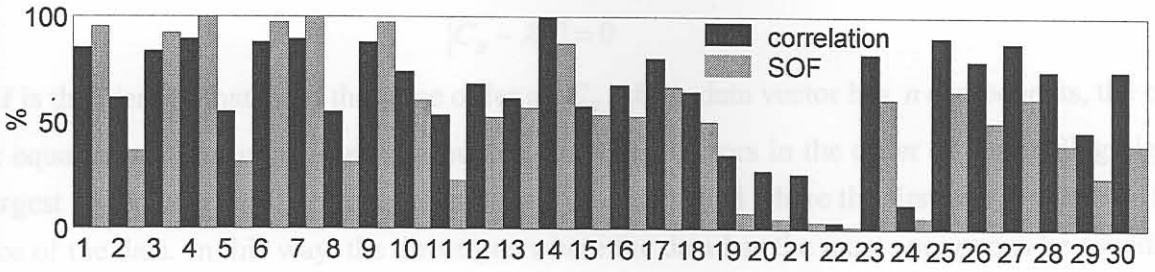


Figure 6.11: Comparison of correlation coefficient and SOF

In many cases, choosing many features and then using a feature space reduction method such as PCA to reduce the feature space can achieve a better result. Fortunately in data sets with many features, groups of features often move together. One reason for this is that more than one variable may be measuring the same driving principle governing the behaviour of the system. In many systems there are only a few such driving forces. But an abundance of instrumentation allows us to measure dozens of system variables. When this happens, the problem can be simplified by replacing a group of variables with a single new variable. PCA is a one method for achieving this simplification. The method generates a new set of variables, called principal components. Each principal component is a linear combination of the original variables. All the principal components are orthogonal to each other so there is no redundant information. The principal component as a whole forms an orthogonal basis for the space of the data, and is essentially based on the statistical representation of a random variable. Suppose there is a random vector population $\boldsymbol{x}$ where:

$$\boldsymbol{x} = (x_1, \ldots, x_n)^T \tag{6.3}$$

and the mean of that population is denoted by

$$\mu_x = E\{\boldsymbol{x}\} \tag{6.4}$$

and the covariance of the same data set is

$$\boldsymbol{C}_x = E\left\{(\boldsymbol{x} - \mu_x)(\boldsymbol{x} - \mu_x)^T\right\} \tag{6.5}$$

The components of $\boldsymbol{C}_x$, denoted by c_{ij}, represent the covariances between the random variable components x_i and x_j. The component c_{ii} is the variance of the component x_i. The variance of a component indicates the spread of the component around its mean. If two components x_i and x_j of the data are uncorrelated, their covariance is zero $\left(c_{ij} = c_{ji} = 0\right)$. The covariance matrix will always be symmetric. From a sample of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$, the mean and covariance matrices can be calculated. From the covariance matrix, an orthogonal basis can be calculated by determining eigenvalues and eigenvectors. The eigenvectors $\boldsymbol{e}_i$ and the corresponding eigenvalues λ_i are the solutions of the following equation:

$$C_x e_i = \lambda_i e_i \text{ with } i = 1,...,n \tag{6.6}$$

For simplicity it is assumed that λ_i is distinct. These can be determined by solving the characteristic equation:

$$|C_x - \lambda_i I| = 0 \tag{6.7}$$

where I is the identity matrix of the same order as C_x. If the data vector has n components, the characteristic equation will have order n. By ordering the eigenvectors in the order of descending eigenvalues (largest first), an ordered orthogonal basis can be established where the first eigenvector has largest variance of the data. In this way, the directions where the data has the most energy can be found. If the data is transformed with these principal components, it reduces the amount of data but retains the information containing the most energy.

Principal component analyses were performed on the hard turning and Aluminium turning experiments. The results of the 1st two principal components, together with the features chosen for the AI monitoring method in the previous chapters, are plotted in Figure 6.12 and Figure 6.13. It can be seen that the 1st principal component follows the increasing trend of features with less noise than features themselves. The 2nd principal component is unfortunately only noise, because the TCM problem is one-dimensional. Features generally increase (or sometimes decrease) with increasing tool wear. The tool wear is also a monotonically increasing variable. As a result, the 1st principal component contains the information along the increasing axis, and the 2nd principal component is generated by the inherent noise in the data. An advantage of PCA is that the 1st principal component is less noisy than the feature values themselves. Care must be taken not to contaminate the principal components with noise from redundant data.

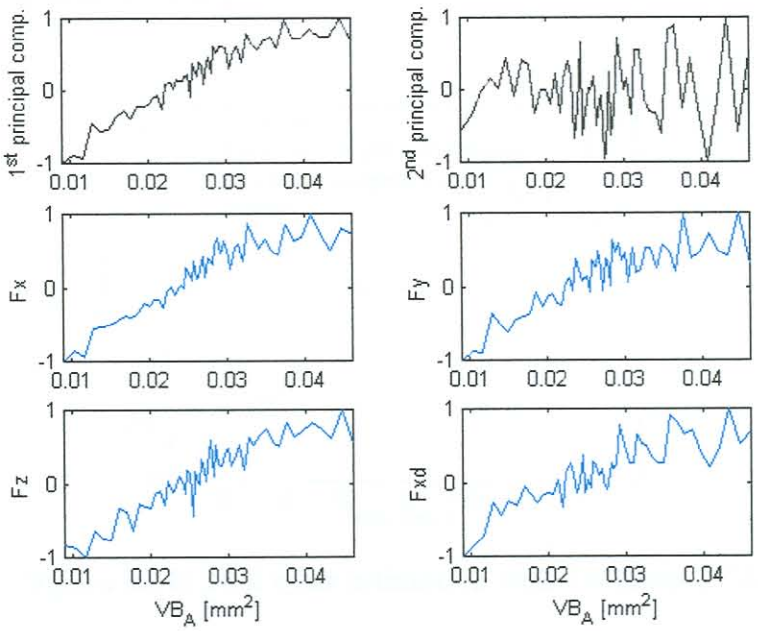


Figure 6.12: Principal component analysis Hard turning (data normalised)

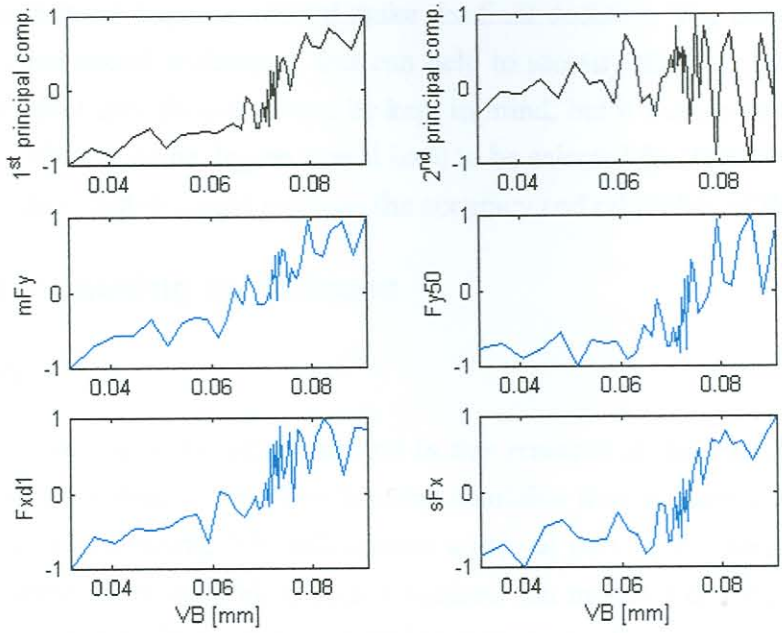


Figure 6.13: Principal component analysis Aluminium turning (data normalised)

It is clear from the examples from the two different turning processes that a PCA can actually increase the reliability of the TCMS, despite the fact that the problem is one-dimensional. One possibility would be to keep the existing four features and just add a fifth, namely the 1st principal component of the data. Another possibility would be to base the TCMS only on the 1st principal component. This was done with data from Aluminium turning and the results are compared in Figure 6.14. It can be seen from the figure that adding the 1st principal component achieved a slight increase in the reliability of the TCMS. However, using the 1st principal component only did not work very well. This is an important result, because it shows that using data that are linearly correlated to some degree yields better results than using only the principal component.

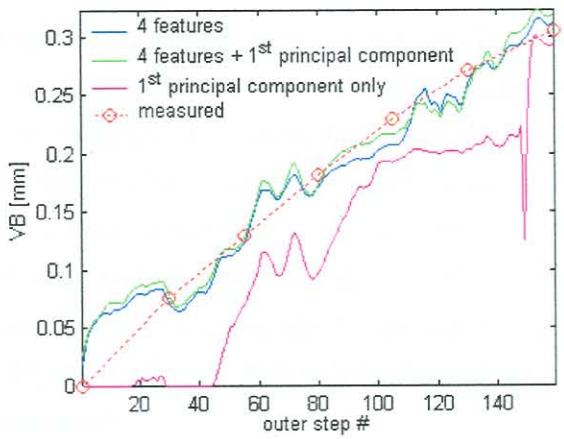


Figure 6.14: Tool wear estimation with / without PCA

The conclusion from the investigations into feature selection is that available techniques can be used to assist in the decision but the engineer should make the final decision. The correlation coefficient, the SOF and PCA are three useful techniques that can help to identify the best features. The problem of selecting linear dependant data should always be kept in mind, but it was shown that selecting features that are linearly dependant to some degree would have to be selected for TCM. Adding the 1st principal component as an addition feature could increase the accuracy and reliability of the TCMS.

6.3 Alternative modelling techniques

6.3.1 Introduction

One of the advantages of the proposed AI method in this research is that certain parts of it can be replaced by other methods - should a method become available that is more accurate and fit for shop floor implementation. The dynamic NN will remain a crucial part of the method, but the static NNs can be replaced by some other method, if such a method can model a chosen signal feature. Models that can calculate static cutting forces for sharp tools are common. However, due to the complex nature of tool wear, it is difficult to estimate the worn tool forces with most other methods. Features such as those derived from the frequency spectrum would be very difficult if not impossible to determine by means of theoretical models only.

One possibility would be to combine available theoretical models with the AI approach. An advantage of theoretical approaches is that they can handle changing machining conditions with more ease and accuracy than AI models. The reason for this is the *a priori* knowledge of the theoretical models of what the effect of changing conditions would be, whereas the AI model needs appropriate training samples to obtain this knowledge. In this section, some promising theoretical methods are discussed that may enhance the accuracy of a TCMS.

6.3.2 Finite Element Method (numerical models)

The use of the FEM to model machining operations was discussed in Section 2.5. Besides the FEM, various other numerical computer simulation methods are available or are under development. This approach, though still in the development stage for many machining operations, seems to be one of the most promising to assist is sensor-based TCM. A numerical model can be used instead of the static NN for TCM. This is what makes the new formulation proposed in this work particularly useful: Instead of modelling the tool wear as an output of the model it is actually used as input to the static NN. Hence, with tool wear as an input, a numerical method such as the FEM can estimate the static cutting forces for many combinations of tools, workpieces and machining parameters. The dynamic NN will still be included to follow the development of wear, but will be trained with error between the FEM model simulation and on-line measurements.

One problem is that this approach will be computationally very slow with available computing power. In future years computers might become fast enough to use the FEM model iteratively to train the dynamic NN. In fact, with available computing speed, it might not be extremely slow because the dynamic NN usually requires less than ten iterations. The use of response surfaces could also prove very

useful to lessen the number of FEM simulations. Unfortunately, the FEM approach could not be investigated for this work but is suggested for further research. FEM models could also be used to normalise data with respect to machining conditions. To achieve this will require intensive research and collaboration between various research groups.

6.3.3 Theoretical models

The basics of theoretical models to predict cutting forces were discussed in Chapter 1. A pure theoretical model that can accurately predict worn tool forces does not exist and establishing such a model would be virtually impossible. The only possible use of these models is to assist another method, such as the AI approach. For instance, an accurate theoretical model can be used to:

- Predict the sharp tool cutting forces
- Normalise the AI approach for cutting conditions
- Assist as a validation procedure

In this research, the analytical method described in Section 2.2.3 was evaluated as a possible method of calculating the sharp tool forces and then adding the worn tool forces by means of the AI approach. The analytical procedure involves determining the oblique cutting constants through orthogonal cutting tests. When this is found, an oblique cutting transformation is applied to two regions along the radius of the tool insert. The complete model relies on the accuracy in determining certain constants (some of which will be available from a database for some tool and workpiece combinations [240]) and the validity of a few assumptions. Because this and other theoretical models rely on cutting tests and underlying assumptions, it raises the question if it is worthwhile to implement the method if it can only assist with TCM in part, namely predicting sharp tool cutting forces. Also, the dynamic behaviour of the tool is not included in the model. After careful consideration it was decided that it is not worthwhile to attempt a shop floor implementation of the method because it will raise the complexity of a problem that the AI approach can already handle to satisfaction.

6.3.4 Mechanistic models

The underlying assumption of mechanistic cutting models is that the cutting forces are proportional to the uncut chip area. The constants of proportionality depend on the cutting conditions and geometry and material properties. Kapoor *et al.* [241,242] also described a method that can be incorporated into a basic mechanistic approach that can predict dynamic cutting forces. This could be achieved by impact tests on the machine structure. They also describe a worn tool force model [223,243] for turning operations. This is the first model based on theoretical foundations for predicting worn tool forces. Of course, other experimental models exist that rely on the accurate determination of certain empirical constants.

Because the mechanistic worn tool force model is a possible complete replacement of the static NNs, it was decided to attempt an implementation of the model. The complete mechanistic approach for predicting worn tool forces is described in Appendix G. The method was applied to the hard turning data after personal correspondence with Kapoor and DeVor from UIUC indicated that the method would

after personal correspondence with Kapoor and DeVor from UIUC indicated that the method would possibly be applicable to the type of processes investigated in this research. To determine the sharp tool forces is basically a matter of calculating empirical constants from cutting experiments. With this a calibration procedure is carried out that will yield the mechanistic constants. The part of model that deals with determining the worn tool forces relies on a number of assumptions, the most important being:

- There is a linear growth of the plastic flow region on the tool flank
- There exists a critical value of flank wear after which plastic flow will be observed
- The worn tool forces are governed only by flank wear
- An accurate calculation of the maximum effective stress in the workpiece is possible

It was found that for most of the hard turning experiments, the critical value of flank wear could not be determined. Even after extensive testing, it was not always observed. Attempts were made to calculate the maximum contact stress in the workpiece by the method described in Appendix G, and a result is shown in Figure 6.15. When this is accurately established and verified, the 3-D worn tool forces can be calculated by the equations described in Appendix G. However, the maximum contact stress could not be verified because the assumptions made for the mechanistic model does not seem to apply for hard turning. It was mentioned several times in Chapter 4 that the stability of hard turning is governed largely by crater wear. During hard turning, the crater wear also has an influence on the cutting forces, but crater wear is not included in the mechanistic approach. In fact, a theoretical inclusion of the crater wear is rather difficult because the mechanics of cutting with crater wear is not as clear as the case with flank wear.

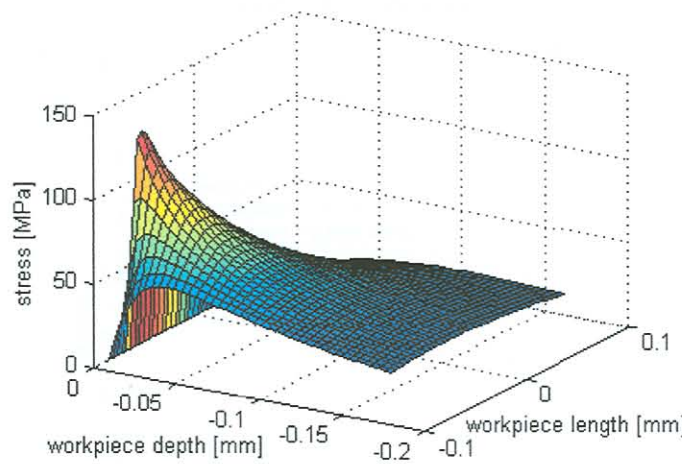


Figure 6.15: Calculating maximum contact stress in workpiece (hard turning)

The linear growth of the plastic flow and the critical flank wear was also not so clearly observed with the CBN tools, and as a result the maximum contact stress could not be determined. However, the plastic and elastic zones were better observed in the Aluminium turning experiments. Because the aim of this work is to determine a method of TCM that can be treated as a general methodology for turning, it was decided not to attempt an implementation of the model on Aluminium turning because the method did not apply to hard turning as well.

6.3 AI approach improvement

6.3.1 Introduction

In this section, some aspects regarding the improvement of the AI methodology are investigated. In the opinion of the author, the improvement that can be obtained by optimising the network type, structure and activation functions are relatively small compared to selecting the correct measurement procedure, signal processing and feature selection steps. In this case, same aspects of the NN structure were investigated in order determine if it yields a worthwhile improvement or not.

6.3.2 Type of network

Many different types of network were compared for the best results. This included FF networks, FF networks with time delays, radial basis function networks, perceptrons, recurrent networks and unsupervised networks. In the case of the FF backpropagation networks, different activation functions and networks sizes were also compared. A comparison of the FF network as formulated in Chapter 5 and FF networks with time delays are shown Figure 6.16. The time delay networks required more training steps and when the convergence criteria were kept the same, the time delay networks did not yield very good results. Investigations only apply to the static NNs. If more training steps were taken, the time delay networks improved but did not yield better results than the initial formulation of the FF networks. Sick [171] has shown that time delay NNs should be used for TCM, because the TCM problem requires temporal information for accurate estimation. In the case of a the AI implemented in this research, the temporal information is already built into the time delay of the dynamic networks, and is not required for the static NNs. It was decided not to attempt to optimise time delays in the static NNs because it requires more training and also slows the training of the dynamic NNs down.

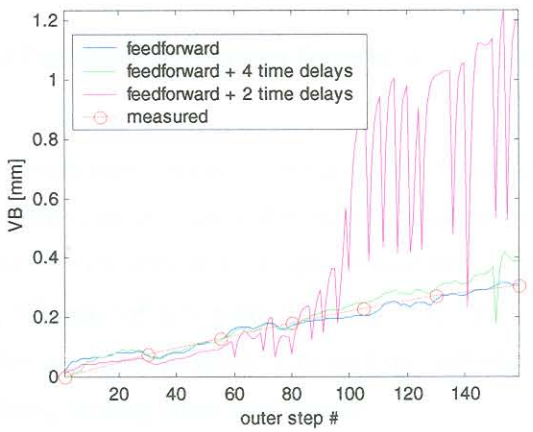


Figure 6.16: Time delay comparison (Aluminium turning)

In Figure 6.17 the result of using different network types is shown. The FF and the cascade forward networks yielded similar results. The Elman network requires longer training but the smooth response of the Elman network is a very nice attribute. The Elman network utilise feedback connections, and thus doubles the use of temporal information in the AI approach. The last investigation into network

formulation was to compare different activation functions. The results of using three different activation functions are shown in Figure 6.18. The tansigmoid and radial basis function yielded the best result.

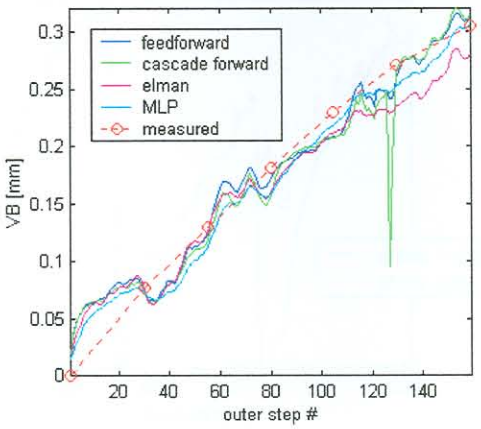


Figure 6.17: Different network types (Aluminium turning)

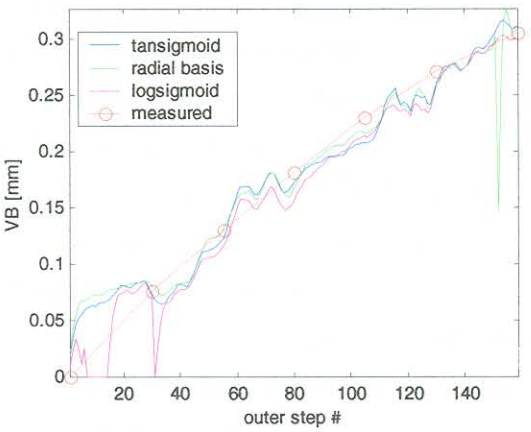


Figure 6.18: Different activation functions (Aluminium turning)

The results presented in this sub-section are only a fraction of the many investigations into formulating the best method for TCM. Unfortunately, only a few results can be shown and discussed here. In conclusion it can be stated that the FF network with a tansigmoid and linear layers yielded the best results and requires the least amount of time for adequate training. The Multilayer Perceptron (MLP) also exhibits good results, but requires a larger network. Elman type networks can also be considered due to their smooth response but training is very time consuming.

An important contribution of this research lies in the use of the dynamic NN for on-line monitoring. However, using only static NNs can also model the tool wear with the chosen features. For this reason a comparison of the two methods are included here to show that the proposed formulation of using static and dynamic networks is indeed the better one for practical applications. An example is shown in Figure 6.19, where simulations of a regular FF network is compared with that of the new formulation proposed in this study (without a sliding window output). The data is from three tools that wear from

new to approximately 0.1mm flank wear. It can be seen that the performance of the combined static and dynamic formulation is better than the FF network (static only). A much smoother network response is noted and the network has no difficulty to return to zero (automatic re-initialisation was used).

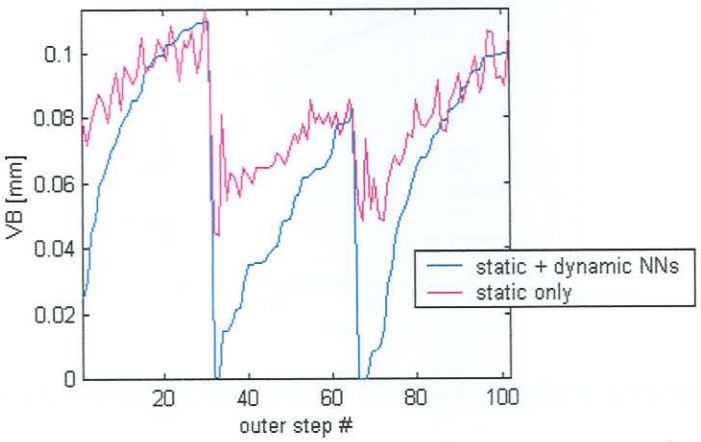


Figure 6.19: Comparison of formulations (Aluminium turning)

The new formulation was also compared to other conventional formulations, some of which performed well but do not hold the advantages of the combined formulation. No other NN formulation that was investigated outperformed the proposed combined formulation on the noisy shop floor data. Some of the results are reported in the figures that follow. The results are based on networks trained on the same data and tested on a previously unseen set of data. Where possible, the same network training tolerances were used. As a result, the sizes of the networks are not exactly the same for all the cases. The results of using of a Single Layer Perceptron (SLP), Adaptive Linear Neuron Networks (ADELINE) and adaptive linear filter network (linear layer with input delays) are plotted in Figure 6.20. The SLP can only be used for classification and is not recommend for continuous estimation. The ADELINE and linear adaptive filter display more or less the same result.

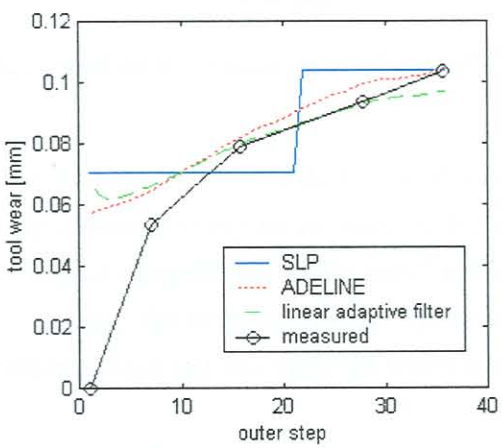


Figure 6.20: SLP, ADELINE and linear adaptive filter result

A Radial Basis Function (RBF) network is compared with a Generalised Regression Neural Network (GRNN), which is a RBF with an added linear layer at the output), and a two-layer FF network trained

with backpropagation in Figure 6.21. Again, the results are very similar.

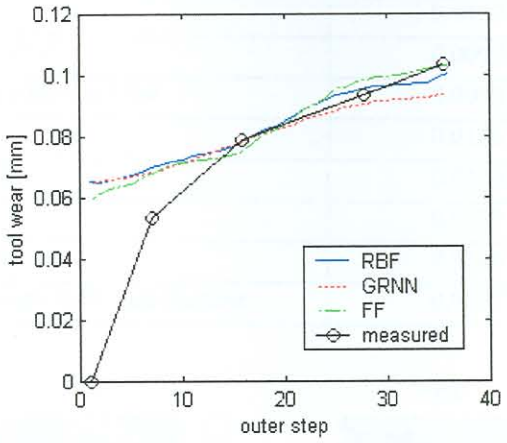


Figure 6.21: RBF, GRNN and FF backpropagation network result

A recurrent network (Elman formulation) is compared with the new formulation of combining the static and dynamic networks in Figure 6.22. The Elman network performs well but the new combined network still outperforms the Elman network.

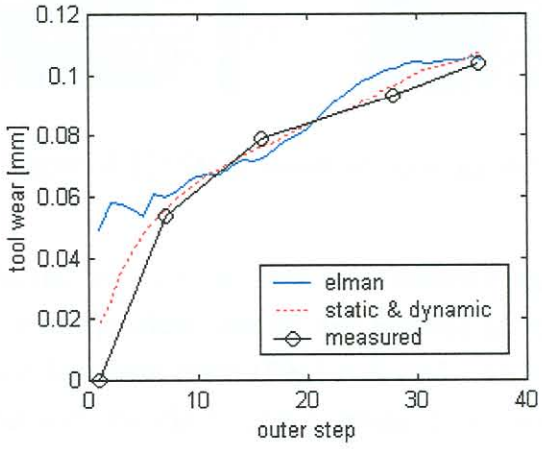


Figure 6.22: Recurrent (Elman) and new formulation with static and dynamic networks result

The results for this case study are summarised in Table 6.1. It should be kept in mind that the result reported here is for one case study although many more were performed to ensure that the newly proposed dynamic formulation does indeed outperform "static only" networks. From the graphs it can also be seen that besides increased accuracy, the response of the dynamic network is smoother and more stable. Training and stability problems were not encountered when using the PSOA.

As a last step the use of an unsupervised NN were investigated, one again in the form of the SOM (refer to Appendix H). The normalised training features were subjected to SOM training, and the result is shown in Figure 6.23. From the figure it can be seen that the features are automatically arranged in low and high regions.

Table 6.1: rms errors on training data

network type	*rms error [mm]*
SLP	0.0167
ADELINE	0.0097
linear adaptive filter	0.0092
RBF	0.0106
GRNN	0.0106
FF	0.0109
Elman	0.0106
combined static and dynamic	0.0057

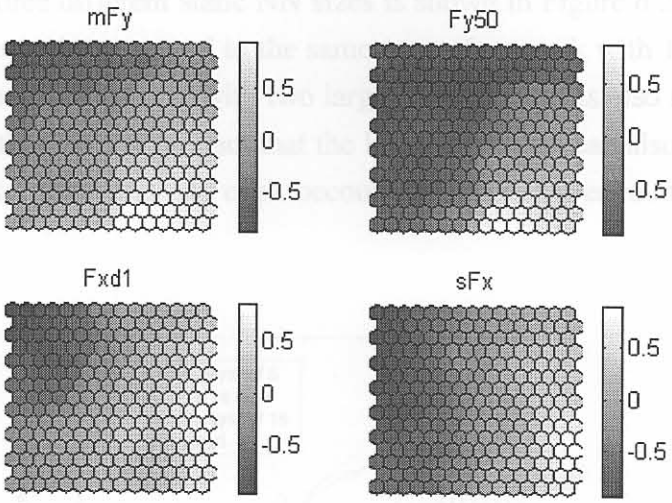


Figure 6.23: SOM result of training data

The labels of the training data, (in this case the flank wear values in mm) are plotted on the left hand side on Figure 6.24. A separate independent data set was labelled as "brand new", "new", "medium", "worn" and "replace". The Best Matching Units (BMUs) for this data were calculated the trajectory is plotted in Figure 6.25, together with the classification labels. From this it can be established that the SOM is very useful for interpreting the multi-dimensional data, but should rather be used for discrete classification (*e.g.* "new" or "worn") than continuous estimation. In this example, the SOM yielded very good results in terms of classification.

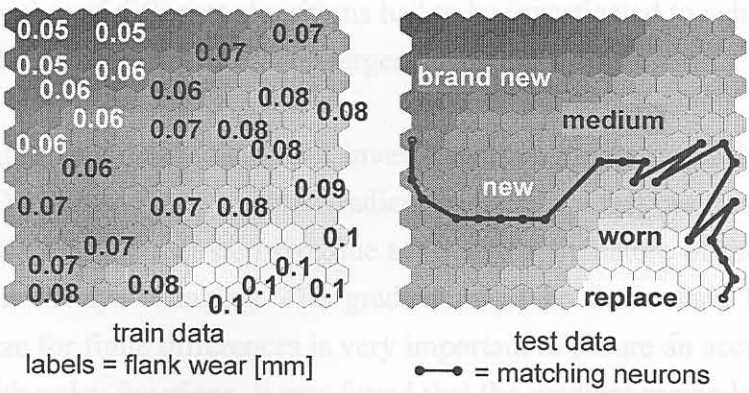


Figure 6.24: Training labels and testing classification

6.3.3 Size of networks

Although no mathematical optimisation procedure was performed, the size of the static and dynamic networks was optimised with manual iterations. To improve network generalisation, the networks were kept as small as possible to avoid overtraining. This also helps to improve training speed. If larger network sizes were used, the training was slow and the network could not generalise properly. If feed rate or cutting speed was included in the networks, a larger network size had to be used. In general, a middle layer of five neurons sufficed for a network with no machining parameters. If machining parameters were included, a middle layer of approximately ten neurons had to be used.

As an example, a result of three different static NN sizes is shown in Figure 6.25. A FF network with five neurons in the middle layer is compared to the same type of network with 15 neurons in the middle layer. The result from another network, with two large hidden layers is also shown in the figure. It can be seen from the figure that despite the fact that the larger networks can also follow the tool wear, they are much more prone to noise and can even become unstable if it encounters a noisy measurement.

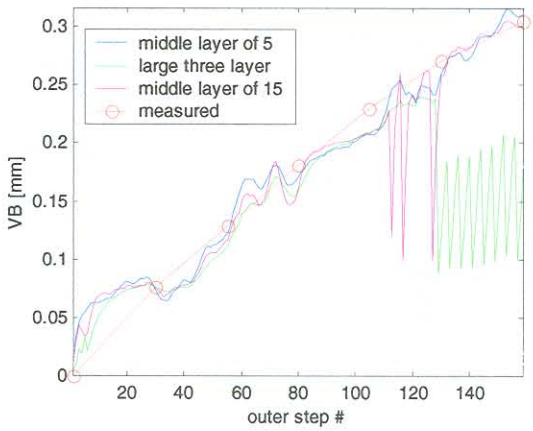


Figure 6.25: Different network sizes (Aluminium turning)

6.3.4 Training algorithm

The static networks were FF networks with backpropagation as a training procedure. The method was fast and accurate enough for the application without any convergence problems. In the case of the dynamic algorithms, a number of different algorithms had to be investigated to achieve fast on-line training. With conventional training procedures, convergence was slow or not at all.

The different optimisation algorithms that were investigated are discussed in Appendix D. These are ETOP, SQSD and LFOP. It was found that the gradient methods are fast but it is difficult to determine the gradient function with the correct step size due to the dynamic nature of the network. This is because the input data is always changing. The gradient must be determined by a finite difference method and the step size for finite differences is very important to ensure an accurate estimation of the gradient, especially with noisy functions. It was found that the gradient methods improve for dynamic training when the step size is decreased linearly when the objective function approach zero.

It was then decided to investigate another method that does not utilise a gradient calculation, and this was found in the PSOA. It was found that the PSOA provided quick and accurate training and rarely fails to converge. Furthermore, the PSOA does not have the problem of calculating gradient functions, and the random nature of the algorithm is ideal for this application. It was concluded that the PSOA is the best choice for on-line training of dynamic NNs.

6.3.5 Repeatability of simulations

Another important test for any implementation of NNs is to determine the repeatability of simulation results. This involves re-initialising and re-training all the networks and repeating the simulation on the same data. This was repeated for several data sets, and an example is plotted in Figure 6.26. Due to the nature of the PSOA and various convergence criterions to enforce generalisation, the simulations are not expected to be exactly the same every time, but should at least be very similar. This can be observed from the various simulations in Figure 6.26. Each follows the same progression of tool wear but they are not numerically the same. It can thus be concluded that the simulations are repeatable with newly trained networks.

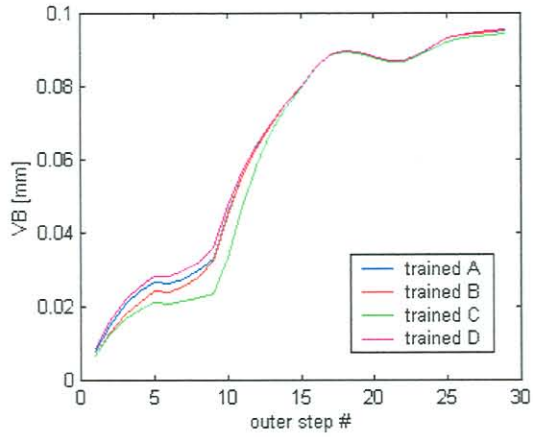


Figure 6.26: Repeated simulations with newly trained static NN

6.4 Conclusion

In this chapter, it was shown that:

- The coherence function is a method that can assist in TCM but is not universally applicable.
- Wavelet analysis can sometimes improve the accuracy of a TCMS but the same result (if not better) can be reached by simple digital filtering procedures.
- Feature selection could be done by the correlation function method and the SOF. PCA is another useful method and including the 1st principal component as another data feature can improve the accuracy of a TCMS.
- Engineering judgement remains the most important step of feature selection and care must be taken to avoid 100% linearly correlated data.
- Different modelling methods can assist in the AI method of TCM, but not all methods are universally applicable and often increase the complexity of required experiments and

mathematical formulation unnecessarily, making these methods insufficient for on-line implementation.

- Attempts were made to improve the formulation of the NNs in the AI approach for TCM by investigating network type, activation functions, network size and training algorithms.
- Comparisons between formulations were made and it was also shown that the result of the new formulation is repeatable.

7. Conclusion and recommendations

7.1 Preface

In this chapter, the final conclusions of the research are represented in a bulleted manner to make the conclusions and contributions clear and concise. The conclusions are grouped into sections dealing with measurement, signal analysis and modelling issues. The conclusions are measured against the specific objectives, listed in Section 1.5. After the conclusions, some recommendations for future research in this area are made. The recommendations are specifically aimed to continue the current success with employing NNs for TCM in industry. Recommendations are also made on the broad scope of hardware and software issues related to this work.

7.2 Conclusions

7.2.1 Summary of conclusions

A new AI approach for TCM is proposed. It was shown that the method:

- can monitor two wear modes accurately during hard turning with inclusion of machining parameters,
- monitor flank wear in interrupted cutting of Aluminium on a shop floor with varying feed rate.
- utilises the advantages of AI, using a combination of NNs that estimates the wear values based on basic knowledge (static networks), past knowledge (dynamic networks) and present knowledge (on-line sensors) and can be used with cost-effective hardware instead of expensive laboratory equipment
- is the first industrial implementation of an AI approach to TCM, and provides a useful solution to industry,
- provides significant new knowledge as to how to solve the problem of TCM.

When measured against the general objectives in Section 1.5.1, it can be stated that the objectives were met adequately.

7.2.2 Signal measurement

In this research, several sensor approaches were investigated. An exhaustive survey of research and industrial developments was also included and the following conclusions are made:

- Sensors for TCM in industry must be cost-effective, robust and reliable. A measurement as close as possible to the point of metal removal is absolutely essential for continuous tool wear estimation.
- AE sensors can be used on a shop floor situation because they are robust, small and easy to install. However, they are not reliable for continuous wear estimation due to a lack of physical in-

terpretation of the obtained signal and their sensitivity towards unrelated process noise. AE sensors are however very efficient for tool breakage detection and other applications in machining process monitoring.

- Vibration monitoring assist in collecting the dynamic behaviour of the machine tool and if vibration signals are processed by an expert a continuous wear estimation can be achieved. A difficulty still lies with appropriate interpretation of vibration signals and some basic ways of analysing the vibration signals were presented in this work. An advantage is that wideband accelerometers specifically for machine tools are available although they are somewhat expensive.
- It was shown that measuring a parameter related to the cutting force can achieve cost-effective and reliable TCM in industry. This was accomplished using strain gauges. The strain gauge approach has several advantages:
    - low cost
    - robust
    - small
    - measurement close to point of metal removal
    - wireless signal transmission possible
    - does not change dynamic properties of machine
    - 3-D cutting force can be determined with reasonable accuracy
    - wideband frequency analysis possible
- Thus, the strain gauge combines the advantages of force and vibration sensors in one simple package. The optimal position of strain gauges on the tool holder was determined using the FEM. Furthermore, static and dynamic testing procedures were presented to calibrate the strain gauge system to reconstruct the cutting forces. The measurements were insensitive to clamping conditions, materials and other vibrational effects.
- It was shown how an automated data logger can be built cost-effectively. The data logger utilises an overload protection device, amplifiers, filters, A/D conversion and a computer with C++ software. Drift (due to temperature effects) compensation is provided and the system could be monitored though the Internet.
- Another advantage of the data logger approach is that the tool wear is recorded under realistic conditions (hence not an accelerated tool life test).
- Care must be taken to avoid electrical disturbances when applying sensors on machine tools. Proper earthing and cable shielding are essential.
- Other methods besides force and vibration based approaches have yet to show that they can provide an acceptable solution for continuous wear estimation in industry.

Measured against the signal measurement objectives in Section 1.5.2, these conclusions exceed the expectations set in the objectives. All objectives were met adequately.

7.2.3 Signal processing

Another important step for successful TCM is the signal processing that is employed. Appropriate signal processing methods must be used to generate signal features that correlate or indicate tool wear.

The best features must then be selected to be used as an input to a wear model or decision-making technique. The following conclusions regarding signal processing are made:

- The type of signal processing that will generate reliable features for TCM depends on the type of operation, machine and sensors. As a general rule, the following analyses should be included (when available):
  - Static forces as features
  - Frequency analysis of force and / or vibration to identify wear sensitive frequency ranges as features
  - AErms as a feature (usefulness of AE frequency analysis is debatable)
  - Time domain features
  - Statistical features
  - Time-frequency investigation for non-stationary behaviour
- Using one or several of the techniques listed above, wear sensitive signal features can be generated. The signal features do not have to be insensitive to controllable machining conditions.
- The usefulness of wavelet analysis for TCM is debatable. Using wavelets adds to the complexity of the TCMS and using an appropriate selection of digital filters can achieve better results.
- A combination of features derived from time, frequency and statistical analyses will yield the best results. A TCMS should not be based on features from only one of the domains.
- Identifying internal and external disturbances to the signal features can assist in interpretation of the signals and features. Disturbances should be avoided, or their effect removed from the signals. It was shown how a SOM analysis could assist in data mining for data collected from machining processes. The result can be used to identify and avoid disturbances in experimental data.
- Feature selection or feature space reduction can be achieved through various automated methods. Methods that were used in this study was:
  - Correlation Coefficient Approach (CCA)
  - Statistical Overlap Factor (SOF)
  - Genetic Algorithm (GA)
  - Principal Component Analysis (PCA)
- Combining the CCA and SOF for feature selection yielded the best results and is most suitable if a quick and simple method is required. The GA is somewhat slower and requires more trial runs (a GA procedure must always be repeated several times to ensure that the global optimum was reached).
- Using only the Principal Components as features is generally not reliable for TCM. It was shown that including the 1st principal component as an additional feature can improve the result somewhat. However, the adaptability of the PCA toward small changes in the process conditions must be investigated before industrial application.
- The most important aspect of feature selection is engineering judgement. Care must be taken not take 100% linearly correlated data. However, using features that show some degree of correlation is not wrong, if the features were generated from different sensors and also using different signal processing methods. Hence, an appropriate combination representing different sensors and processing methods yields the best results. However, care must be taken not to choose too many fea-

tures. If a single sensor single feature approach works very well, adding more features will probably worsen the situation. In the case of TCM, it was found that the single feature approach is generally not reliable due to noisy conditions and therefore sensor / feature fusion is required. As a general rule 4-10 features could be used.

The objectives listed in Section 1.5.3 were met adequately.

7.2.4 Modelling and monitoring

Different techniques of modelling were investigated during the course of this research. The following conclusions are made:

- Analytical / theoretical modelling is of limited use to TCM. The models suffer from many limitations and might never reach the level of sophistication where a model is available for any process without experimentation. Furthermore, tool wear cannot be described by these methods.
- Numerical / simulation models are reaching a level of sophistication where most processes can be modelled with any combination of materials and geometry. Because the methods require a lot of expertise and time, they are not feasible for on-line implementation but can be considered as an additional tool for either treating the effect of new cutting conditions or as verification of the experimental models.
- Empirical / experimental modelling is the only remaining option. Many different approaches exist, for instance parametric or non-parametric approaches. AI modelling also has a non-parametric empirical nature. In this research, it was shown for the first time that the AI method can work effectively on shop floor conditions.
- The use of a mechanistic model proved unsuccessful due to some basic assumptions within the model that do not apply to a diverse range of turning operations.
- In this research, it was shown that the growth of tool wear is always unique and unpredictable. As a result, the sensorless approaches to the tool wear problem will not be effective enough if optimal tool use is required. The sensorless approaches yield a tool life equation that will only be an estimation of the real tool life. With sensor-based methods, the tool life is monitored.
- An AI method of monitoring is proposed based on Neural Networks (NNs). The method utilises combinations of static and dynamic NNs. The method has several advantages:
    - Excellent generalisation capabilities
    - Effective use of temporal information
    - Insensitive to noisy data
    - Machining conditions (*e.g.* feed rate and speed) can be included
    - Insensitive towards clamping conditions and other external disturbances
    - Combines the use of current and historical data as well as a knowledge basis
    - Same architecture applies to different turning operations
    - The method can follow any geometrical development of tool wear
    - The method can follow more than one wear mode (*e.g.* flank wear and crater wear)
- Several NN architectures were compared for use on the static NN level. It was found that the inclusion of time delays requires a slightly larger network with more training. The inclusion of the

time delays did not yield a further improvement. Radial Basis Function (RBF) networks, Multi-layer Perceptron (MLP), Elman type networks and FF networks with different activation functions were compared. Most yielded acceptable results. It is suggested to use the Elman or FF network with the "tansigmiod" activation function for best results.

- The methods employed in this research to ensure generalisation capabilities of the static networks were early stopping and using a small network. It was shown that a too large network or over-trained networks cause instabilities.
- Several new optimisation algorithms were investigated for training the dynamic NNs, because it was found that conventional methods are too slow for on-line implementation and do not always converge. The Particle Swarm Optimisation Algorithm (PSOA) was found to be best algorithm for on-line training. The PSOA has the following advantages:
    - Fast and reliable
    - Simple formulation and implementation
    - No gradient function evaluation required
    - Random nature of optimising ideal for NNs
    - The method outperforms other methods in unconstrained global optimisation
- The AI monitoring method was trained, validated and tested on separate data sets. It was also shown that the results are repeatable after re-initialisation and training of the static NNs. The dynamic NNs converge within a few outer steps after re-initialisation.

Measured against the signal measurement objectives in Section 1.5.4, most conclusions exceed the expectations set. In the area of numerical models more research is required, and this is explored in Section 7.2.3.

7.3 Recommendations

7.2.1 Measurement

The following suggestions can be made for future research with respect to signal measurement:

- optimise the number, size and position of strain gauges on tool holder.
- investigate the possibility of an on-board strain gauge amplifier on tool holder
- develop better mechanical protection for strain gauges
- investigate the industrial implementation of wireless data transfer
- attempt constructing a sensor-integrated tool for larger tool holders
- use mechanical amplification on tool holder (*e.g.* holes that cause stress concentrations)
- extend the Internet monitoring capabilities of the system

7.2.2 Signal processing

Many signal processing methods were investigated in this study. Future work should be directed towards feature selection or feature space reduction. Other techniques that have been mentioned in the literature but not considered in this work are for example octave analysis and bispectrum analysis. These might prove useful for future research. It is however suggested that other types of machining

operations be investigated. The type of operation also determines which signal processing technique will be most effective. It might require the development of a custom technique.

7.2.3 Modelling

A significant improvement in the use of the on-line AI method would be to minimise the amount of training data required for successful implementation. Future work should be mainly directed towards this topic. The disadvantage of NNs (and many other experimental) models is that they require training / calibration data for the range of conditions they are expected to operate on. A NN cannot be expected to yield accurate results for previously unseen machining conditions (although it might perform to satisfaction, it cannot be expected *a priori*). Thus, methods of normalising data with respect to machining conditions should be the main focus of future work. Consequently the NNs do not need to be trained for every condition. The use of numerical models to achieve this is one attractive option and should be investigated in future, such as a simulation model described by Weinert and Zabel [244]. The use of any kind of experimental model will basically have the same influence as including the machining conditions in the AI approach, as was done in this research. The advantage of a numerical model would be that no experimental data is required, and therefore no adjustments to the TCMS would be required when a machining condition change. Present analytical models will not provide a solution to this problem.

Future work can of course also include the integration of the current system into the CNC machine, instead of a separate stand-alone device. Additionally, the exact machining parameters and machining profile could be provided to the TCMS in this way, and the machine can be automatically shut down when excessive wear is detected.

7.4 Contribution

Despite exhaustive investigations, conversations and communications with researchers and industrial representatives worldwide, the author could not find a single example of a TCMS using AI that was proved to work on industrial data (also refer to Section 1.5.6). Thus, this research has overcome the difficulties involved with a real implementation of the AI method for TCM. This was achieved by designing, developing, and building the system from start to finish (hardware and software). This required insight into many different disciplines, *e.g.* electronics, structural dynamics, NNs, mathematics (optimisation and statistics), manufacturing, signal processing, data acquisition and computer programming. A suitable combination of knowledge from the different disciplines enabled a unique solution to the TCM problem, and is claimed to be the first practical implementation of a TCMS using AI. The formulation and application of the AI tool wear monitoring method proposed is unique in terms of its formulation and application, and was shown to outperform other AI approaches. This contribution provides a significant improvement towards more cost-effective, reliable and accurate tool wear monitoring.

REFERENCES

[1] Mitsubishi Materials Corporation, *Internet web site*, http://www.mmc.co.jp, **2002**.

[2] Promess Inc., *Keep an eye on process with tool monitoring*, Tooling and Production, pp. 63-64, February, **1994**.

[3] J. Kopac, *Influence of cutting material and coating on tool quality and tool life*, Journal of Materials Processing Technology, vol. 78, pp. 95-103, **1998**.

[4] A.K.M.N. Amin, M.R.A. Sarker. M. Ahmed and A.N.M. Karim, *Selection of cemented carbide turning tools using EMF and optimization criteria*, Journal of Materials Processing Technology, vol. 77, pp. 59-63, **1998**.

[5] G. Byrne, D. Dornfeld, I. Inasaki, G. Ketteler, W. König and R. Teti, *Tool Condition Monitoring (TCM) - The Status of Research and Industrial Application*, Annals of the CIRP, vol. 44, pp. 541-567, **1995**.

[6] B. Sick, *Online and indirect tool wear monitoring in turning with artificial neural networks: A review of more than a decade of research*, Mechanical Systems and Signal Processing, vol. 16, pp. 487-546, **2002**.

[7] C.S. Leem and D.A. Dornfeld, *Design and implementation of sensor-based tool-wear monitoring systems*, Mechanical Systems and Signal Processing, vol. 10, pp. 439-458, **1996**.

[8] A. Ghasempoor, J. Jeswiet and T.N. Moore, *On-line wear estimation using neural networks*, Proceedings of the Institution of Mechanical Engineers - B, vol. 212, pp. 105-112, **1998**.

[9] R.J. Silva, K.J. Baker, S.J. Wilcox and R.L. Reuben, *The adaptability of a tool wear monitoring system under changing cutting conditions*, Mechanical Systems and Signal Processing, vol. 14, pp. 287-298, **2000**.

[10] G.V. Stabler, *The chip flow law and its consequences*, Advances in Machine Tool Design and Research, pp. 243-251, **1964**.

[11] M.E. Merchant, *Mechanics of the metal cutting process, 2: Plasticity Conditions in Orthogonal Cutting*, Journal of Applied Physics, vol. 16, pp. 318-324, **1945**.

[12] E.H. Lee and B.W. Shaffer, *Theory of Plasticity Applied to the Problem of Machining*, Journal of Applied Mechanics, vol. 18, pp. 405-413, **1951**.

[13] W.B. Palmer and P.L.B. Oxley, *Mechanics of Orthogonal Machining*, Proceedings of the Institution of Mechanical Engineers, vol. 172, pp. 623-654, **1959**.

[14] Y. Altintas, *Manufacturing Automation: Metal Cutting Mechanics, Machine Tool Vibration and CNC Design*, Cambridge University Press, Cambridge, U.K., **2000**.

[15] J.A. Arsecularatne, P. Mathew and P.L.B. Oxley, *Prediction of chip flow direction and cutting forces in oblique machining with nose radius tools*, Proceeding of the Institute for Mechanical Engineers, vol. 209, pp. 305-315, **1995**.

[16] J. Krystof, *Berichte uber Betriebswissenschaftliche Arbeiten, Bd.*, VDI Verlag, vol. 12, **1939**.

[17] M.E. Merchant, *Mechanics of the cutting process*, Journal of Applied Physics, vol. 16, pp. 318-324, **1945**.

[18] International Standard (ISO) 3002/1, *Basic quantities in cutting and grinding - Part 1: Geometry of the active part of cutting tools - General terms, reference systems, tool and working angles, chip breakers*, Second Edition, **1982**.

[19] Mitsubishi Carbide Tools, *Internet web site*, http://www.mitsubishicarbide.com, **2002**.

[20] I.S. Jawahir, R. Ghosh, X.D. Fang and P.X. Li, *An investigation of the effects of chip flow on tool-wear in machining with complex grooved tools*, Wear, vol. 184, pp. 145-154, **1995**.

[21] I.S. Jawahir and X.D. Fang, *A knowledge-based approach for designing effective grooved chip-breakers - 2D and 3D chip flow, chip curl and chip breaking*, International Journal of Advanced Manufacturing Technology, vol. 10, pp. 225-239, **1995**.

[22] N. Fang, I.S. Jawahir and P.L.B. Oxley, *A universal slip-line model with non-unique solutions for machining with curled chip formation and a restricted contact tool*, International Journal of Mechanical Sciences, vol. 43, pp. 557-580, **2001**.

[23] R. Gosh, O.W. Dillon and I.S. Jawahir, *An investigation of 3-D curled chip in machining - Part 1: A mechanics-based analytical model*, Machining Science and Technology, vol. 2, pp. 91-116, **1998**.

[24] R. Gosh, O.W. Dillon and I.S. Jawahir, *An investigation of 3-D curled chip in machining - Part 2: Simulation and validation using FE techniques*, Machining Science and Technology, vol. 2, pp. 177-135, **1998**.

[25] B.K. Ganapathy and I.S. Jawahir, *Modeling the chip-work contact force for chip breaking in orthogonal machining with a flat-faced tool*, Journal of Manufacturing Science and Engineering, vol. 120, pp. 49-56, **1998**.

[26] A.K. Balaji, G. Sreeram, I.S. Jawahir and E. Lenz, *The effects of cutting tool thermal conductivity on tool-chip contact length and cyclic chip formation in machining with grooved tools*, Annals of the CIRP, vol. 48, pp. 33-38, **1999**.

[27] X.D. Fang and I.S. Jawahir, *The effects of progressive tool wear and tool restricted contact on chip breakability in machining*, Wear, vol. 160, pp. 243-252, **1993**.

[28] I.S. Jawahir and J. Fei, *A comprehensive evaluation of tool inserts for chip control using fuzzy modeling of machinability parameters*, Transactions of NAMRI/SME, vol. 21, pp. 205-213, **1993**.

[29] I.S. Jawahir and J.P. Zhang, *An analysis of chip curl development, chip deformation and chip breaking in orthogonal machining*, Transactions of NAMRI/SME, vol. 23, pp. 109-114, **1995**.

[30] Y. Morimoto, Y. Ichida and R. Sata, *Excitation Technique by 2-axes shaker of an CNC Late*, Proceedings of the 18th International Conference on Modal Analysis, San Antonio, Texas, U.S.A., pp. 1643-1648, **2000**.

[31] T. Koizumi, N. Tsujiuchi and Y. Matsumura, *Diagnosis with the correlation integral in the time domain*, Mechanical Systems and Signal Processing, vol. 14, pp. 1003-1010, **2000**.

[32] L. Lago, S. Olsson, L. Hakansson and I. Claesson, *Design of an efficient chatter control system for turning and boring applications*, Proceedings of the 20th International Modal Analysis Conference (IMAC XX), Los Angeles, California, U.S.A., **4-7 February 2002**.

[33] R. Du, *Signal understanding and tool condition monitoring*, Engineering Application of Artificial Intelligence, vol. 12, pp. 585-597, **1999**.

[34] A. Ghasempoor, J. Jeswiet and T.N. Moore, *Real time implementation of on-line tool condition monitoring system*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1883-1902, **1999**.

[35] D.Y. Jang and A. Seireg, *Machining parameter optimization for specified surface conditions*, Journal of Engineering for Industry, vol. 114, pp. 254-257, **1992**.

[36] S.B. Rao, *Tool wear monitoring through the dynamics of stable turning*, Journal of Engineering for Industry, vol. 108, pp. 184-189, **1986**.

[37] C. Scheffer and P.S. Heyns, *Synthetic diamond tool wear monitoring using vibration measurements*, Proceedings of the 18th International Modal Analysis Conference, San Antonio, Texas, U.S.A., pp. 245-251, **7-10 February 2000**.

[38] M.E.R. Bonifacio and A.E. Diniz, *Correlating tool wear, tool life, surface roughness and tool vibration in finish turning with coated carbide tools*, Wear, vol. 173, pp. 137-144, **1994**.

[39] P.J.A. Lever, M.M. Marefat and T. Ruwani, *A Machine Learning Approach to Tool Wear Behavior Operational Zones*, IEEE Transactions on Industry Applications, vol. 33, pp. 264-273, **1997**.

[40] D.W. Yen and P.K. Wright, *Adaptive Control in Machining - a new approach based on the physical constraints of tool wear mechanisms*, Journal of Engineering for Industry, vol. 105, pp. 31-38, **1983**.

[41] D.E. Dimla, *Sensor signals for tool-wear monitoring in metal cutting operations - a review of methods*, International Journal of Machine Tools and Manufacture, vol. 40, pp. 1073-1098, **2000**.

[42] I.S. Jawahir, P.X. Li, R. Gosh and E.L. Exner, *A new parametric approach for the assessment of comprehensive tool wear in coated grooved tools*, Annals of the CIRP, vol. 44, pp. 49-54, **1995**.

[43] P.X. Li, I.S. Jawahir, X.D. Fang and E.L. Exner, *Chip-groove effects on multiple tool-wear parameters in machining*, Transactions of NAMRI/SME, vol. 24, pp. 33-38, **1996**.

[44] I.S. Jawahir, R. Gosh, A.K. Balaji and P.X. Li, *Predictability of tool failure modes in turning with complex grooved tools using the equivalent toolface (ET) model*, Wear, vol. 224, pp. 94-103, **2000**.

[45] G. Parakkal, R. Zhu, S.G. Kapoor and R.E. DeVor, *Modeling of turning process cutting forces for grooved tools*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 179-191, **2002**.

[46] K.H.W. Seah, X. Li and K.S. Lee, *The effect of applying coolant on tool wear in metal cutting*, Journal of Materials Processing Technology, vol. 48, pp. 495-501, **1995**.

[47] Q. Zhou, G.S. Hong and M. Rahman, *A new tool life criterion for tool condition monitoring using a neural network*, Engineering Applications of Artificial Intelligence, vol. 8, pp. 579-588, **1995**.

[48] SECO Tools Corporation, *Internet web site*, http://www.secotools.com, **2002**.

[49] G.H. Lim, *Tool-wear monitoring in machine turning*, Journal of Materials Processing Technology, vol. 51, pp. 25-36, **1995**.

[50] T. Lundholm and B. Lindstrom, *A flexible real-time control system for turning*, Annals of the CIRP, vol. 40, pp. 441-444, **1991**.

[51] S.C. Lim, *Recent developments in wear-mechanism maps*, Tribology International, vol. 31, pp. 87-97, **1998**.

[52] S.C. Lim, C.Y.H. Lim and K.S. Lee, *The effects of machining conditions on the flank wear of TiN-coated high speed steel tool inserts*, Wear, vol. 181-183, pp. 901-912, **1995**.

[53] T. Obikawa, C. Kaseda, T. Matsumura, W.G. Gong and T. Shirakashi, *Tool wear monitoring for optimizing cutting conditions*, Journal of Materials Processing Technology, vol. 62, pp. 374-379, **1996**.

[54] Z. J. Da, J.P. Sadler and I.S. Jawahir, *Predicting optimum cutting conditions for turning operations at varying tool-wear states*, Transactions of the North American Manufacturing Research Institution of SME, vol. 25, pp. 75-80, **1997**.

[55] X.P. Li, H.H. Ng and S.C. Lim, *A predictive mapping system for tool wear in metal cutting*, Journal of Materials Processing Technology, vol. 89-90, pp. 279-286, **1999**.

[56] R. G. Silva, R. L. Rueben, K. J. Baker and S. J. Wilcox, *Tool wear monitoring of turning operations by neural network and expert system classification of a feature set generated form multiple sensors*, Mechanical Systems and Signal Processing, vol. 12, pp. 319-332, **1998**.

[57] I. Grabec, E. Govekar, E. Susic and B. Antolovic, *Monitoring manufacturing processes by utilizing empirical modeling*, Ultrasonics, vol. 36, pp. 263-271, **1998**.

[58] A. Ruiz, D. Guinea, L.J. Barrios and F. Betancourt, *An empirical multi-sensor estimation of tool wear*, Mechanical Systems and Signal Processing, vol. 7, pp. 105-199, **1993**.

[59] W.J. Braun, M.H. Miller and J.F. Schultze, *The development of machine-tool force reconstruction for wear identification*, Proceedings of the 17th International Modal Analysis Conference, Kissimee, Florida, U.S.A., pp. 94-98, **1999**.

[60] H.V. Ravindra, Y.G. Srinivasa and R. Krishnamurthy, *Modelling of tool wear based on cutting forces in turning*, Wear, vol. 169, pp. 25-32, **1993**.

[61] W.S. Lin, B.Y. Lee and C.L. Wu, *Modeling the surface roughness and cutting force for turning*, Journal of Materials Processing Technology, vol. 108, pp. 286-293, **2001**.

[62] G. Boothroyd and W.A. Knight, *Fundamentals of Machining and Machine Tools*, Marcel Dekker, New York, pp. 178, **1989**.

[63] I.S Jawahir, X.D. Fang, P.X. Li and R. Ghosh, *Method of assessing tool-life in grooved tools*, United States Patent no. 5,689,062, **1997**.

[64] Q. Meng, J.A. Arsecularatne and P. Mathew, *Calculation of optimum cutting conditions for turning operations using a machining theory*, International Journal of Machine Tools and Manufacture, vol. 40, pp. 1709-1733, **2000**.

[65] K.S. Park and S.H. Kim, *Artificial intelligence approaches to determination of CNC machining parameters in manufacturing: a review*, Artificial Intelligence in Engineering, vol. 12, pp. 127-134, **1998**.

[66] D.A. Axinte, W. Belluco and L. De Chiffre, *Reliable tool life measurements in turning - an application to cutting fluid efficiency evaluation*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 1003-1014, **2001**.

[67] S.M. Athavale and J.S. Strenkowski, *Finite Element Modeling of Machining: From proof-of-concept to engineering applications*, Proceedings of the CIRP International Workshop on Modeling of Machining Operations, Atlanta, Georgia, U.S.A., pp. 203-216, **May 19, 1998**.

[68] D.R. Sandstrom, *Modeling the physics of metal cutting in high-speed machining*, Proceedings of the CIRP International Workshop on Modeling of Machining Operations, Atlanta, Georgia, U.S.A., pp. 217-224, **May 19, 1998**.

[69] M.R. Lovell, S. Bhattacharya and R. Zeng, *Modeling of orthogonal machining processes for variable tool-chip interfacial friction using explicit dynamic finite element methods*, Proceedings of the CIRP International Workshop on Modeling of Machining Operations, vol. Atlanta, Georgia, U.S.A., pp. 265-276, **May 19, 1998**.

[70] A. Marty, P. Lorong and G. Coffignal, *Including workpiece vibrations in numerical simulation in machining*, Proceedings of the 3rd International Conference on Metal Cutting and High Speed Machining, Metz, France, pp. 107-115, **27-29 June 2001**.

[71] L. C. Lee, K.Y. Lam and X.D. Liu, *Characterisation of tool wear and failure*, Journal of Materials Processing Technology, vol. 40, pp. 143-153, **1994**.

[72] T.D. Marusich and E. Askari, *Modeling residual stress and workpiece quality in machined surfaces*, Proceedings of the 3rd International Conference on Metal Cutting and High Speed Machining, Metz, France, pp. 95-105, **27-29 June 2001**.

[73] J. Mackerle, *Finite-element analysis and simulation of machining: a bibliography*, Journal of Materials Processing Technology, vol. 86, pp. 17-44, **1999**.

[74] I.S. Jawahir, O.W. Dillon, A.K. Balaji, M. Redetzky and N. Fang, *Predictive modelling of machining performance in turning operations*, Machining science and technology, vol. 2, pp. 253-276, **1998**.

[75] Y. Altintas, *Modeling Approaches and Software for Predicting the Performance of Milling Operations at MAL - UBC*, Manufacturing Automation Laboratory (MAL), http://www.mech.ubc.ca/~mal, **2001**.

[76] D.S. Ermer, *Optimization of the constrained machining economics problem by geometric programming*, Journal of Engineering for Industry, vol. 93, pp. 1067-1072, **1971**.

[77] X.D. Fang and I.S. Jawahir, *Predicting total machining performance in finish turning using integrated fuzzy-set models of the machinability parameters*, International Journal of Production Research, vol. 32, pp. 833-849, **1994**.

[78] I.S. Jawahir, A.K. Balaji, R. Stevenson and C.A. van Luttervelt, *Towards predictive modeling and optimization of machining operations*, Manufacturing Science and Technology, vol. 2, pp. 3-12, **1997**.

[79] X.D. Fang, J. Fei and I.S. Jawahir, *A hybrid algorithm for predicting chip form/chip breakability in machining*, International Journal of Machine Tools and Manufacture, vol. 16, pp. 1093-1107, **1996**.

[80] S.K. Choudhury, E. Kumar and A. Ghosh, *A scheme of adaptive turning operations*, Journal of Materials Processing Technology, vol. 87, pp. 119-127, **1999**.

[81] C. Zhou and R.A. Wysk, *Tool status recording and its use in probalistic optimization*, Journal of Engineering for Industry, vol. 114, pp. 494-499, **1992**.

[82] Y.S. Tarng, S.T. Chen and Y.S. Wang, *Adaptive Control of Machining Operations*, Key Engineering Materials, vol. 138-140, pp. 263-287, **1998**.

[83] Y. Hatamura, T. Nagao and M. Mituishi, *A fundamental structure for intelligent manufacturing*, Precision Engineering, vol. 15, pp. 266-273, **1993**.

[84] A. Davies, *The intelligent machine*, Manufacturing Engineer, vol. 73, pp. 182-185, **1994**.

[85] A. Jeang, *Reliable tool replacement policy for quality and cost*, European Journal of Operational Research, vol. 108, pp. 334-344, **1998**.

[86] M.S. Akturk and S. Avci, *Tool allocation and machining conditions optimization for CNC machines*, European Journal of Operational Research, vol. 94, pp. 335-348, **1996**.

[87] B. Gopalakrishnan and F. Al-Khayyal, *Machine parameter selection for turning operations with constraints: an analytical approach based on geometric programming*, International Journal of Production Research, vol. 29, pp. 1897-1908, **1991**.

[88] C.M. Nicolescu, *On-line Identification and Control of Dynamic Characteristics of Slender Workpieces in Turning*, Journal of Material Processing Technology, vol. 58, pp. 374-378, **1996**.

[89] D.R. Martinez, T.D. Hinnerichs and J.M. Redmond, *Vibration Control for Precision Manufacturing Using Piezoelectric Actuators*, **1999**.

[90] S.B. Billatos and P. Tseng, *Knowledge-based optimization for intelligent manufacturing*, Journal of Manufacturing Systems, vol. 10, pp. 464-475, **1991**.

[91] C.A. van Luttervelt, T.H.C. Childs, I.S. Jawahir, F. Klocke and P.K. Venuvinod, *Present situation and future trends in modelling of machining operations. Progress report of the CIRP working group "Modelling of Machining Operations"*, Annals of the CIRP, vol. 47, pp. 587-626, **1998**.

[92] L. Dan and J. Mathew, *Tool wear and failure monitoring techniques for turning - a review*, International Journal of Machine Tools and Manufacture, vol. 30, pp. 579-598, **1990**.

[93] C. Scheffer and P.S. Heyns, *Tool condition monitoring systems - an overview*, International Conference on Competitive Manufacturing (COMA '01), Stellenbosch, South Africa, pp. 316-323, **31 Jan - 2 Feb 2001**.

[94] R. Teti, *A review of tool condition monitoring literature database*, Annals of the CIRP, vol. 44, pp. 659-667, **1995**.

[95] T. Pfeifer and H. Thrum, *Open systems link sensors and measurement applications to machine tool control units*, Measurement, vol. 19, pp. 113-121, **1996**.

[96] K.F. Martin, *A review by discussion of condition monitoring and fault diagnosis in machine tools*, International Journal of Machine Tools and Manufacture, vol. 34, pp. 527-551, **1994**.

[97] D. Cho, S.J. Lee and C.N. Chu, *The state of machining process monitoring research in Korea*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1697-1715, **1999**.

[98] I. Kinghorn, *Smart sensor technology - the next generation*, Paper Technology, vol. 35, pp. 39-41, **1994**.

[99] N. Tanaka, S.D. Snyder and C.H. Hansen, *Distributed parameter modal filtering using smart sensors*, Journal of Vibration and Acoustics, Transactions of the ASME, vol. 118, pp. 630-640, **1996**.

[100] J. Bryzek, *Smarter, less costly sensors are on the way*, Process and Control Engineering, vol. 47, pp. 40-43, **1994**.

[101] Tamas Szecsi, *Automatic cutting-tool condition monitoring on CNC lathes*, Journal of Materials Processing Technology, vol. 77, pp. 64-69, **1998**.

[102] L.J. Barrios, A. Ruiz, D. Guinea, A. Ibanez and P. Bustos, *Experimental Comparison of sensors for tool-wear monitoring on milling*, Sensors and Actuators A, vol. 37-38, pp. 589-595, **1993**.

[103] M. Santochi, G. Dini and G. Tantussi, *A Sensor-Integrated Tool for Cutting Force Monitoring*, Annals of the CIRP, vol. 46, pp. 49-52, **1996**.

[104] L.F. Puerta and J. Madl, *Tool condition monitoring in drilling*, Manufacturing Technology, vol. 1, pp. 33-37, **2001**.

[105] P. Novak and J. Madl, *Effective evaluation of measured dynamics values of cutting forces and torques*, Manufacturing Technology, vol. 1, pp. 56-62, **2001**.

[106] W. Li, J. Ni, D. Li and S.J. Hu, *Diagnosis of tapping process by two-stage pair-wise feature selection and classification using spindle motor current*, Personal correspondence, University of Michigan, **2000**.

[107] P.C. Tseng and A. Chou, *The intelligent on-line monitoring of end milling*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 89-97, **2002**.

[108] T.I. El-Wardany, D. Gao and M.A. Elbestawi, *Tool condition monitoring in drilling using vibration signature analysis*, International Journal of Machine Tools and Manufacture, vol. 36, pp. 687-711, **1996**.

[109] X.Q. Li, Y.S. Wong and A.Y.C. Nee, *Tool wear and chatter detection using the coherence function of two crossed accelerations*, International Journal of Machine Tools and Manufacture, vol. 37, pp. 425-435, **1997**.

[110] C.Y. Jiang, Y.Z. Zhang and H.J. Xu, *In-Process monitoring of tool wear stage by the frequency band energy method*, Annals of the CIRP, vol. 36, pp. 45-48, **1987**.

[111] D. Meredith, *Practical tool condition monitoring*, Manufacturing Engineering, pp. 34-39, **1998**.

[112] D. Bähre, M. Müller and G. Warnecke, *Basic characteristics on cutting effects in correlation to dynamic effects*, 1997 Technical Papers of the North American Manufacturing Research Institution of SME, pp. 21-26, **1997**.

[113] S. Kim, B.E. Klamecki, *Milling Cutter Wear Monitoring Using Spindle Shaft Vibration*, Journal of Manufacturing Science and Engineering, vol. 119, pp. 118-119, **1997**.

[114] K. Jemielniak, *Some aspects of AE application in tool condition monitoring*, Ultrasonics, vol. 38, pp. 604-608, **2000**.

[115] A.J.M.M. Araujo, S.J. Wilcox and R.L. Reuben, *Sliding friction as a possible source of Acoustic Emission in metal cutting*, Proceedings of the 13th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2000), Houston, Texas, U.S.A., pp. 381-387, **3-8 Dec, 2000**.

[117] J. Kim, M. Kang, B. Ryu and Y. Ji, *Development of an on-line tool-life monitoring system using acoustic emission signals in gear shaping*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1761-1777, **1999**.

[118] S. Dolinsek and J. Kopac, *Acoustic emission signals for tool wear identification*, Wear, vol. 225-229, pp. 295-303, **1999**.

[119] X. Li, *A brief review: acoustic emission method for tool wear monitoring during turning*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 157-165, **2002**.

[120] K. Jemielnak and O. Otman, *Tool failure detection based on analysis of acoustic emission signals*, Journal of Materials Processing Technology, vol. 76, pp. 192-197, **1998**.

[121] H.V. Ravindra, Y.G. Srinivasa and R. Krishnamurthy, *Acoustic emission for tool condition monitoring in metal cutting*, Wear, vol. 212, pp. 78-84, **1997**.

[122] T.A. Carolan, S.R. Kidd, D.P. Hand, S.J. Wilco, P. Wilkonson, J.S. Barton, J.D.C. Jones and R.L. Reuben, *Acoustic emission monitoring of tool wear during the face milling of steels and aluminium alloys using a fibre optic sensor Part 1: Energy Analysis*, Proceedings of the Institute of Mechanical Engineers, vol. 211, pp. 299-309, **1997**.

[123] T.A. Carolan, S.R. Kidd, D.P. Hand, S.J. Wilco, P. Wilkonson, J.S. Barton, J.D.C. Jones and R.L. Reuben, *Acoustic emission monitoring of tool wear during the face milling of steels and aluminium alloys using a fibre optic sensor Part 2: Frequency Analysis*, Proceedings of the Institute of Mechanical Engineers, vol. 211, pp. 311-319, **1997**.

[124] I.N. Tansel, M.E. Trujillo, W. Bao and T.T. Arkan, *Detection of tool breakage in micro-end-milling operations by monitoring acoustic emission*, 1997 Technical Papers of the North American Manufacturing Research Institution of SME, pp. 69-74, **1997**.

[125] C. Chungchoo and D. Saini, *A computer algorithm for flank and crater wear estimation in CNC turning operations*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 1465-1477, **2002**.

[126] S. Tani and I. Inasaki, *Development of monitoring system for milling processes*, Proceedings of the 3rd International Conference on Intelligent Computation in Manufacturing Engineering, Ischia, Italy, pp. 283-288, **3-5 July 2002**.

[127] J. Lin, *Inverse estimation of the tool-work interface temperature in end milling*, International Journal of Machine Tools and Manufacture, vol. 35, pp. 751-760, **1995**.

[128] J.G. Chow and P.K. Wright, *On-line estimation of tool/chip interface temperatures for a turning operation*, Transactions of the ASME: Journal of Engineering for Industry, vol. 110, pp. 56-65, **1988**.

[129] L. Wang, K. Saito and I.S. Jawahir, *Infrared temperature measurement of curled chip formation in metal machining*, Transactions of NAMRI/SME, vol. 14, pp. 87-92, **1996**.

[130] F. Klocke and S. Hoppe, *FEM modelling of the high-speed cutting process and its experimental verification*, Proceedings of the 3rd International Conference on Metal Cutting and High Speed Machining, Metz, France, pp. 65-72, **27-29 June 2001**.

[131] N. H. Abu-Zahra and T. H. Nayfeh, *Calibrated method for ultrasonic on-line monitoring of gradual wear during turning operations*, International Journal of Machine Tools and Manufacture, vol. 37, pp. 1475-1484, **1997**.

[132] T.H. Nayfeh, O.S. Eyeda and J.C. Duke, *An integrated ultrasonic sensor for monitoring gradual wear on-line during turning operations*, International Journal of Machine Tools and Manufacture, vol. 35, pp. 1385-1395, **1995**.

[133] S. Kurada and C. Bradley, *A machine vision system for tool wear assessment*, Tribology International, vol. 30, pp. 295-304, **1997**.

[134] A. Novak and H. Wiklund, *On-line prediction of the tool life*, Annals of the CIRP, vol. 45, pp. 93-96, **1996**.

[135] A. Karthick, S. Chandra, B. Ramamoorthy and S. Das, *3D Tool wear measurement and visualisation using stereo imaging*, International Journal of Machine Tools and Manufacture, vol. 37, pp. 1573-1581, **1997**.

[136] P. Wilkinson, R.L. Reuben, J.D.C. Jones, J.S. Barton, D.P. Hand, T.A. Carolan and S.R. Kidd, *Surface finish parameters as diagnostics of tool wear in face milling*, Wear, vol. 205, pp. 47-54, **1997**.

[137] S. A. Coker and Y. C. Shin, *In-process control of surface roughness due to tool wear using a new ultrasonic system*, International Journal of Machine Tools and Manufacture, vol. 36, pp. 411-422, **1996**.

[138] S. Vajpayee, *Analytical study of surface roughness in turning*, Wear, vol. 70, pp. 165-175, **1981**.

[139] D. Y. Jang, Y. Choi, H. Kim and A. Hsiao, *Study of the correlation between surface roughness and cutting vibrations to develop an on-line roughness measuring technique in hard turning*, International Journal of Machine Tools and Manufacture, vol. 36, pp. 453-464, **1996**.

[140] Y. K. Chou and C.J. Evans, *Tool wear mechanism in continuous cutting of hardened tool steels*, Wear, vol. 212, pp. 59-65, **1997**.

[141] M. Thomas, Y. Beauchamp, A.Y. Youssef and J. Masounave, *Effect of tool vibration on surface roughness during lathe dry turning process*, Computers and Industrial Engineering, vol. 31, pp. 637-644, **1996**.

[142] F. Caiazzo, G.S. Palazzo and R. Pasquino, *The influence of working parameters on the response of a capacitive sensor used in-process for the measurement of tool wear*, International Journal of Machine Tools and Manufacture, vol. 38, pp. 871-879, **1998**.

[143] Y.S. Wong, A.Y.C. Nee, X.Q. Li and C. Reisdorf, *Tool condition monitoring using laser scatter pattern*, Journal of Materials Processing Technology, vol. 63, pp. 205-210, **1997**.

[144] S.K. Choudhury, V.K. Jain and Ch. V. V. Rama Rao, *On-line monitoring of tool wear in turning using a neural network*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 489-504, **1999**.

[145] D. Choi, W.T. Kwon and C.N. Chu, *Real-Time Monitoring of Tool Fracture in Turning Using Sensor Fusion*, International Journal of Advanced Manufacturing Technology, vol. 15, pp. 305-310, **1999**.

[146] E. Govekar, J. Gradisek and I. Grabec, *Analysis of acoustic emission signals and monitoring of machining processes*, Ultrasonics, vol. 38, pp. 598-603, **2000**.

[147] D.E. Dimla and P.M. Lister, *On-line metal cutting tool condition monitoring I: force and vibration analyses*, International Journal of Machine Tools and Manufacture, vol. 40, pp. 739-768, **2000**.

[148] D.E. Dimla and P.M. Lister, *On-line metal cutting tool condition monitoring II: tool-state classification using multi-layer perceptron neural networks*, International Journal of Machine Tools and Manufacture, vol. 40, pp. 769-781, **2000**.

[149] S. A. Kumar, H.V. Ravindra and Y.G. Srinivasa, *In-process tool wear monitoring through time series modeling and pattern recognition*, International Journal of Production Research, vol. 35, pp. 739-751, **1997**.

[150] M. Bayramoglu and Ü. Düngel, *A systematic investigation on the force ratios in tool condition monitoring for turning operations*, Transactions of the Institute of Measurement and Control, vol. 20, pp. 92-97, **1998**.

[151] G. Warnecke and S. Siems, *Dynamics in high speed machining*, Metal Cutting and High Speed Machining, Kluwe Academic / Plenum Publishers, New York, pp. 21-30, **2002**.

[152] D.A. Axinte, W. Belluco, L. De Chiffre, *Evaluation of cutting force uncertainty components in turning*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 719-730, **2001**.

[153] I. Choi and J. Kim, *Development of monitoring system on the diamond tool wear*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 505-515, **1999**.

[154] L.C. Lee, K.S. Lee and C.S. Gan, *On the correlation between dynamic cutting force and tool wear*, International Journal of Machine Tools and Manufacture, vol. 29, pp. 295-303, **1989**.

[155] D.K. Baek, T.J. Ko and H.S. Kim, *Real time monitoring of tool breakage in a milling operation using a digital signal processor*, Journal of Materials Processing Technology, vol. 100, pp. 266-272, **2000**.

[156] Y. Yao, X.D. Fang and G. Arndt, *Comprehensive Tool Wear Estimation in Finish-Machining via Multivariate Time-Series Analysis of 3-D Cutting Forces*, Annals of the CIRP, vol. 39, pp. 57-60, **1990**.

[157] Y. Yao and X.D. Fang, *Modelling of multivariate time series for tool wear estimation in finish turning*, International Journal of Machine Tools and Manufacture, vol. 32, pp. 495-508, **1992**.

[158] C. Jun and S. Suh, *Statistical tool breakage detection schemes based on vibration signals in NC milling*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1733-1746, **1999**.

[159] A.D. Jennings and P.R. Drake, *Machine tool condition monitoring using statistical quality control charts*, International Journal of Machine Tools and Manufacture, vol. 37, pp. 1243-1249, **1997**.

[160] B.Y. Lee and Y.S. Tarng, *Milling cutter breakage detection by the discrete wavelet transform*, Mechatronics, vol. 9, pp. 225-234, **1999**.

[161] G.S. Hong, M. Rahman and Q. Zhou, *Using neural network for tool condition monitoring based on wavelet decomposition*, International Journal of Machine Tools and Manufacture, vol. 36, pp. 551-566, **1996**.

[162] L. Xiaoli, Y. Yingxue and Y. Zhejun, *On-line tool condition monitoring system with wavelet fuzzy neural network*, Journal of Intelligent Manufacturing, vol. 8, pp. 271-276, **1997**.

[163] Y. Wu and R. Du, *Feature extraction and assessment using wavelet packets for monitoring of machining processes*, Mechanical Systems and Signal Processing, vol. 10, pp. 29-53, **1996**.

[164] C. Scheffer, *Monitoring of tool wear in turning operations using vibration measurements*, Masters dissertation (MEng), Department of Mechanical and Aeronautical Engineering, University of Pretoria, South Africa, **1999**.

[165] C. Scheffer and P.S. Heyns, *Wear monitoring in turning operations using vibration and strain measurements*, Mechanical Systems and Signal Processing, vol. 15, pp. 1185-1202, **2001**.

[166] G. Luo, D. Osypiw and M. Irle, *Tool wear monitoring by on-line vibration analysis with wavelet algorithm*, Metal Cutting and High Speed Machining, Kluwe Academic / Plenum Publishers, New York, pp. 393-405, **2002**.

[167] C.J. Li and T. Tzeng, *Multimilling-insert wear assessment using non-linear virtual sensor, time-frequency distribution and neural networks*, Mechanical Systems and Signal Processing, vol. 14, pp. 945-957, **2000**.

[168] S. Gu, J. Ni and J. Yuan, *Non-stationary signal analysis and transient machining process condition monitoring*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 41-51, **2002**.

[169] J.C. Fu, K. Mori, and M. Yokomichi, *Application of entropy functions in on-line vibration classification for cylindrical plunge grinding*, International Journal of Production Research, vol. 32, pp. 1477-1487, **1994**.

[170] C. Chungchoo and D. Saini, *The total energy and the total entropy of force signals - new parameters for monitoring oblique turning operations*, International Journal of Machine Tools and Manufacture, vol. 40, pp. 1879-1897, **2000**.

[171] B. Sick, *Tool wear estimation in turning with process-specific pre-processing and time-delay neural networks*, Submitted to International Journal of Smart Engineering System Design, **2000**.

[172] B. Sick, *On-line tool wear monitoring in turning using neural networks*, Neural Computing and Applications, vol. 7, pp. 356-366, **1998**.

[173] B. Sick, A. Sicheneder and H. Lindinger, *A comparative evaluation of different neural network paradigms for tool wear classification in turning*, Proceedings of the 3rd International Workshop "Neural Networks in Applications (NN '98)", University of Magdeburg, pp. 139-146, **12-13 February 1998**.

[174] B. Sick and A. Sicheneder, *Time-delay neural networks for on-line tool wear classification and estimation in turning*, Proceedings of the 3rd Conference on Neural Networks and Their Applications, Kule, pp. 461-466, **14-18 October 1997**.

[175] A. Al-Habaibeh, N. Gindy and N. Radwan, *An automated approach for monitoring gradual tool wear in high speed milling of titanium*, Proceedings of the 13th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2000), Houston, Texas, U.S.A, pp. 371-380, **3-8 Dec, 2000**.

[176] Y. Quan, M. Zhou and Z. Luo, *On-line robust identification of tool-wear via multi-sensor neural network fusion*, Engineering Applications of Artificial Intelligence, vol. 11, pp. 717-722, **1998**.

[177] J.H. Lee, D.E. Kim and S.J. Lee, *Statistical analysis of cutting force ratios for flank-wear monitoring*, Journal of Materials Processing Technology, vol. 74, pp. 104-114, **1998**.

[178] D.E. Dimla, P.M. Lister and N.J. Leighton, *Investigation of a single-layer perceptron neural network to tool wear inception in a metal turning process*, Proceedings of the 1997 IEE Colloquium on Modelling and Signal Processing for Fault Diagnosis, pp. 3/1 - 3/4, **1996**.

[179] K. Venkatesh, M. Zhou and R.J. Caudill, *Design of artificial neural networks for tool wear monitoring*, Journal of Intelligent Manufacturing, vol. 8, pp. 215-226, **1997**.

[180] D.E. Dimla, P.M. Lister and N.J. Leighton, *Tool condition monitoring in metal cutting through application of MLP neural networks*, Proceedings of the 1997 IEE Colloquium on Fault Diagnosis in Process Systems, pp. 9/1-9/3, **1997**.

[181] S. Das, R. Roy and A.B. Chattopadhyay, *Evaluation of wear of turning carbide inserts using neural networks*, International Journal of Machine Tools and Manufacture, vol. 36, pp. 789-797, **1996**.

[182] G. Luetzig, M. Sanchez-Castillo and R. Langari, *On tool wear estimation through neural networks*, Proceedings of the 1997 IEEE International Conference on Neural Networks - part 4 of 4, Houston, Texas, U.S.A., pp. 2359-2363, **1997**.

[183] S. Das, A.B. Chattopadhyay and A.S.R. Murthy, *Force parameters for on-line tool wear estimation: A neural network approach*, Neural Networks, vol. 9, pp. 1639-1645, **1996**.

[184] T.I. Liu, W.Y. Chen and K.S. Anatharaman, *Intelligent detection of drill wear*, Mechanical Systems and Signal Processing, vol. 12, pp. 863-873, **1998**.

[185] C.M. Talbott, *Prognosis of residual machine life*, Proceedings of the Maintenance and Reliability Conference (MARCON 98), University of Tennessee, Knoxville, pp. 11-14, **1998**.

[186] A. C. Okafor and O. Adetona, *Predicting quality characteristics of end-milled parts based on multi-sensor integration using neural-networks: individual effects of learning parameters and rules*, Journal of Intelligent Manufacturing, vol. 6, pp. 389-400, **1995**.

[187] D.E. Dimla, *Application of perceptron neural networks to tool-state classification in a metal-turning operation*, Engineering Applications of Artificial Intelligence, vol. 12, pp. 471-477, **1999**.

[188] C. Scheffer and P.S. Heyns, *Monitoring of turning tool wear using vibration measurements and neural network classification*, Proceedings of the 25th International Seminar on Modal Analysis, Leuven, Belgium, pp. 899-906, **13-15 September 2000**.

[189] Q. Liu and Y. Altintas, *On-line monitoring of flank wear in turning with multilayered feed-forward neural network*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1945-1959, **1999**.

[190] R.J. Kuo, *Multi-sensor integration for on-line tool wear estimation through artificial neural networks and fuzzy neural network*, Engineering Applications of Artificial Intelligence, vol. 13, pp. 249-261, **2000**.

[191] P.S. Pai, T.N. Nagabhushana and P.K.R. Rao, *Tool wear estimation using resource allocation network*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 673-685, **2001**.

[192] A. Ghasempoor, T.N. Moore and J. Jeswiet, *Tool wear prediction in turning*, Proceedings of the 13th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2000), Houston, Texas, U.S.A, pp. 399-406, **3-8 Dec, 2000**.

[193] A. Zawada-Tomkiewicz, *Classifying the wear of turning tools with neural networks*, Journal of Materials Processing Technology, vol. 109, pp. 300-304, **2001**.

[194] P.S. Pai, T.N. Nagabhushana and P.K.R. Rao, *Tool wear estimation using resource allocation network*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 673-685, **2001**.

[195] C. Chungchoo and D. Saini, *On-line tool wear estimation in CNC turning operations using fuzzy neural network model*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 29-40, **2002**.

[196] H. Demuth and M. Beale, *Neural Network Toolbox for use with MATLAB*, http://www.mathworks.com, **1998**.

[197] D.E. Dimla, P.M. Lister and N.J. Leighton, *Neural network solutions to the tool condition monitoring problem in metal cutting--a critical review of methods*, International Journal of Machine Tools and Manufacture, vol. 37, pp. 1219-1241, **1997**.

[198] E. Govekar and I. Grabec, *Self-organizing neural network application to drill wear classification*, Transactions of the ASME: Journal of Engineering for Industry, vol. 116, pp. 233-238, **1994**.

[199] C.L. Jiaa and D.A. Dornfeld, *A self-organizing approach to the prediction and detection of tool wear*, ISA Transactions, vol. 37, pp. 239-255, **1998**.

[200] C. Scheffer and P.S. Heyns, *Development of an adaptable tool condition monitoring system*, Proceedings of the 13th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2000), Houston, Texas, U.S.A, pp. 361-370, **3-8 Dec, 2000**.

[201] K. Lou and C. Lin, *An Intelligent Sensor Fusion System for Tool Monitoring on a Machining Centre*, International Journal of Advanced Manufacturing Technology, vol. 13, pp. 556-565, **1997**.

[202] J.C. Chen and JT. Black, *A Fuzzy-nets-in-process (FNIP) system for tool breakage monitoring in end-milling operations*, International Journal of Machine Tools and Manufacture, vol. 37, pp. 783-800, **1997**.

[203] B. Sick, *Fusion of Hard and Soft Computing Techniques in Indirect, Online Tool Wear Monitoring*, Personal correspondence, University of Passau, **2001**.

[204] L. Xiaoli, Y. Yingxue and Y. Zhejun, *On-line tool condition neural network with improved fuzzy neural network*, High Technology Letters, vol. 3, pp. 30-38, **1997**.

[205] P. Fu, A.D. Hope and M.A. Javed, *Fuzzy classification of milling tool wear*, Insight, vol. 39, pp. 553-557, **1997**.

[206] S. Li and M.A. Elbestawi, *Fuzzy clustering for automated tool condition monitoring in machining*, Mechanical Systems and Signal Processing, vol. 10, pp. 533-550, **1996**.

[207] R.J. Kuo and P.H. Cohen, *Multi-sensor integration for on-line tool wear estimation through radial basis function networks and fuzzy neural network*, Neural Networks, vol. 12, pp. 355-370, **1999**.

[208] R.J. Kuo and P.H. Cohen, *Intelligent tool wear estimation system through artificial neural networks and fuzzy modelling*, Artificial Intelligence in Engineering, vol. 12, pp. 229-242, **1998**.

[209] R. Du, Y. Liu, Y. Xu, X. Li, Y.S. Wong and G.S. Hong, *Tool condition monitoring using transition fuzzy probability*, Metal Cutting and High Speed Machining, Kluwe Academic / Plenum Publishers, New York, pp. 375-392, **2002**.

[210] M. Beynon, B. Curry and P. Morgan, *The Dempster-Shafer theory of evidence: an alternative approach to multicriteria decision modelling*, Omega, vol. 28, pp. 37-50, **2000**.

[211] H.M. Ertunc, K.A. Loparo and H. Ocak, *Tool wear condition monitoring in drilling operations using hidden Markov models (HMMs)*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 1363-1384, **2001**.

[212] H.M. Ertunc and K.A. Loparo, *A decision fusion algorithm for tool wear condition monitoring in drilling*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 1347-1362, **2001**.

[213] Y.K. Chou and C.J. Evans, *White layers and thermal modeling of hard turned surfaces*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1863-1881, **1999**.

[214] H.A. Kishawy and M.A. Elbestawi, *Effects of process parameters on material side flow during hard turning*, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1017-1030, **1999**.

[215] T.G. Dawson and T.R. Kurfess, *Wear trends of PCBN cutting tools in hard turning*, Metal Cutting and High Speed Machining, Kluwe Academic / Plenum Publishers, New York, pp. 221-231, **2002**.

[216] M. Zimmermann, M. Lahres, D.V. Viens and B.L. Laube, *Investigations of the wear of cubic boron nitride cutting tools using Auger electron spectroscopy and X-ray analysis by EPMA*, Wear, vol. 209, pp. 241-246, **1997**.

[217] Y.K. Chou and C.J. Evans, *Cubic boron nitride tool wear in interrupted hard cutting*, Wear, vol. 225-229, pp. 234-245, **1999**.

[218] T. Ozel and A. Nadgir, *Prediction of flank wear by using back propagation neural network modeling when cutting hardened H-13 steel with chamfered and honed CBN tools*, International Journal of Machine Tools and Manufacture, vol. 42, pp. 287-297, **2002**.

[219] R. Kothamasu and S. H. Huang, *Intelligent tool wear estimation for hard turning: Neural-Fuzzy modeling and model evaluation*, Proceedings of the 3rd International Conference on Intelligent Computation in Manufacturing Engineering, Ischia, Italy, pp. 343-346, **3-5 July 2002**.

[220] H.K. Tonshoff, M. Jung, S. Mannel, W. Rietz, *Using acoustic emission signals for monitoring of production process*, Ultrasonics, vol. 37, pp. 681-686, **2000**.

[221] T. Kohonen, *The self-organizing map*, Neurocomputing, vol. 21, pp. 1-6, **1998**.

[222] J. Vesanto, *SOM-based data visualization methods*, Intelligent Data Analysis, vol. 3, pp. 111-126, **1999**.

[223] D.W. Smithey, S.G. Kapoor and R.E. DeVor, *A worn tool force model for three-dimensional cutting operations*, International Journal of Machine Tools and Manufacture, vol. 40, pp. 1929-1950, **2000**.

[224] J. Kennedy and R. C. Eberhart, *Particle swarm optimisation*, Proceedings of the 1995 IEEE International Conference on Neural Networks, vol. 4, Perth, Australia, pp. 1942-1948, **1995**.

[225] J. Rech, M. Lech and J. Richon, *Surface integrity in finish hard turning of gears*, Metal Cutting and High Speed Machining, vol. Kluwe Academic / Plenum Publishers, New York, pp. 211-220, **2002**.

[226] B. Sick, *Signalinterpretation mit Neuronalen Netzen unter Nutzung von modellbasiertem Nebenwissen am Beispiel der Verschleißüberwachung von Werkzeugen in CNC-Drehmaschinen*, VDI Verlag, Düsseldorf, Fortschritt-Berichte VDI, series 10 "Informatik / Kommunikation", no. 629, pp. , **2000**.

[227] R. Allen and H. Shi, *Tool wear monitoring using a combination of cutting force and vibration signals*, International Journal of COMADEM, vol. 4, pp. 26-32, **2001**.

[228] G.J. Tarmal and P. Opavsky, *Signal processing in measurement of milling forces*, Proceedings of the 13th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2000), Houston, Texas, U.S.A., pp. 389-397, **2000**.

[229] H. Schulz, A. Versch, U. Fiedler and W. Huerkamp, *Process monitoring with mechatronic toolholders*, Proceedings of the 3rd International Conference on Metal Cutting and High Speed Machining, Metz, France, pp. 225-235, **27-29 June 2001**.

[230] C. Li and A.G. Ulsoy, *High-precision measurement of tool-tip displacement using strain gauges in precision flexible line boring*, Mechanical Systems and Signal Processing, vol. 13, pp. 531-546, **1999**.

[232] F. Failli, M. Beghini, G. Dini, M. Santochi and G. Tantussi, *A sensor integrated tool for two-components cutting force monitoring*, Proceedings of the 1st International Seminar on Progress in Innovative Manufacturing Engineering (PRIME 2001), Sestri levante (Genoa), Italy, 20-22 June, **2001**.

[233] Mechanical Vibrations, *S.S. Rao*, Addison Wesley Publishing Company, Massachusetts, **1995**.

[234] Microchip Corporation, *Internet web site*, http://www.microchip.com, **2000**.

[235] System Identification - theory for the user, *Lennart Ljung*, PTR Prentice Hall Inc., Englewood Cliffs, New Jersey, **1987**.

[236] System monitoring and the use of models, *J.H. Williams*, Handbook of Condition Monitoring, Chapman & Hall, London, **1998**.

[237] A. Swami, J.M. Mendel and C.L. Nikias, Higher-Order Spectral Analysis Toolbox for use with MATLAB user's guide, http://www.mathworks.com, **2000**.

[238] C.R. Houck, J.A. Joines and M.G. Kay, *A Genetic Algorithm Function for Optimisation: A Matlab Implementation: NCSU-IE Technical Report 95-09*, http://ie.ncsu.edu/mirage/, **2001**.

[239] F. Klocke. M. Weck, H. Kratz and S. Platen, *SIMON: Increasing the Productivity by Adjusted Monitoring Solutions*, Proceedings of the International Intelligent Manufacturing Systems (IMS) Forum 2001, Monte Verita, Switzerland, pp. 342-350, **8-10 October 2001**.

[240] B.Y. Lee, Y.S. Tarng and H.R. Lii, *An investigation of modeling of the machining database in turning operations*, Journal of Materials Processing Technology, vol. 105, pp. 1-6, **2000**.

[241] S.G. Kapoor, R.E. DeVor and R. Zhu, *Development of Mechanistic Models for the Prediction of Machining Performance: Model-Building Methodology*, Proceedings of the International Workshop on Modeling of Machining Operations, Atlanta, Georgia, U.S.A., pp. 109-120, **May 19, 1998**.

[242] S.G. Kapoor, R.E. DeVor, R. Zhu, R. Gajjela, G. Parakkal and D. Smithey, *Development of mechanistic models for the prediction of machining performance: model building technology*, Machining Science and Technology, vol. 2, pp. 213-238, **1998**.

[243] D.W. Smithey, S.G. Kapoor and R.E. DeVor, *A New Mechanistic Model for Predicting Worn Tool Cutting Forces*, Personal correspondence, UIUC, **2001**.

[244] K. Weinert and A. Zabel, *Simulation-Based prediction of tool wear in milling*, Proceedings of the 3rd International Conference on Intelligent Computation in Manufacturing Engineering, Ischia, Italy, pp. 279-282, **3-5 July 2002**.

[245] Montronix Company, *Internet web site*, http://www.montronix.com, **2000**.

[246] Kistler Company, *Internet web site*, http://www.kistler.ch, **2000**.

[247] K. Jemielniak, *Commercial Tool Condition Monitoring Systems*, International Journal of Advanced Manufacturing Technology, vol. 15, pp. 711-721, **1999**.

[248] A. Kirchheim, P. Wolfer, S. Platen and H. Kratz, *SIMON: Sensor fused Intelligent Monitoring System for Machining*, Proceedings of the International Intelligent Manufacturing Systems (IMS) Forum 2001, Monte Verita, Switzerland, pp. 292-296, **8-10 October 2001**.

[249] Brankamp Company, *Internet web site*, http://www.brankamp.com, **2000**.

[250] Artis Company, *Internet web site*, http://www.artis-systems.com, **2000**.

[251] PCB Company, *Internet web site*, http://www.pcb.com, **2000**.

[252] Entran Company, *Internet web site*, http://www.entran.com, **2000**.

[253] Prometec Company, *Internet web site*, http://www.prometec.com, **2000**.

[254] F. Klocke and M. Rhese, *Intelligent Tools through Integrated Micro Systems*, Annals of the German Society for Production Engineering, vol. 2, **1997**.

[255] U.S. Patent Office, *Internet web site*, http://www.uspto.gov, **2000**.

[256] Ovation Engineering Company, *Internet web site*, http://www.ovationeng.com, **2000**.

[257] M. A. Kaiser, *Advancements in the Split Hopkinson Bar Test*, Master of Science Report, Virginia Polytechnic Institute and State University, Virginia, U.S.A., **1998**.

[258] VISHAY Intertechnology, *Interactive guide to Strain Measurement Technology*, http://www.vishay.com/brands/measurements_group/guide/guide.htm, **2002**.

[259] J.A. Snyman and A.M. Hay, *The Spherical Quadratic Steepest Descent (SQSD) Method for Unconstrained Minimization with No Explicit Line Searches*, Computer and Mathematics with Applications, vol. 42, pp. 169-178, **2001**.

[260] J.A. Snyman, *ETOPC: A FORTRAN program for solving general constrained minimization problems by the conjugate gradient method without explicit line searches*, Research Report, Department of Mechanical Engineering, University of Pretoria, **1998**.

[261] J.A. Snyman, *Unconstrained minimization by combining the dynamic and conjugate gradient methods*, Quaestiones Mathematicae, vol. 8, pp. 33-42, **1985**.

[262] J.A. Snyman, *A new and dynamic method for unconstrained minimization*, Applied Mathematical Modelling, vol. 6, pp. 449-462, **1982**.

[263] J.A. Snyman, *An improved version of the original leap-frog dynamic method for unconstrained minimization LFOP1(b)*, Applied Mathematical Modelling, vol. 7, pp. 216-218, **1982**.

[264] Y. Shi and R.C. Eberhart, *A modified particle swarm optimizer*, Proceedings of the IEEE International Conference on Evolutionary Computation, IEEE Press, Piscataway, New Jersey, U.S.A., pp. 69-73, **1998**.

[265] Eagle Technologies, *Internet web site*, http://www.eagle.co.za, **2000**.

[266] Wavelet Toolbox for use with MATLAB user's guide, *M. Misiti, Y. Misiti, G. Oppenheim and J.P. Poggi*, http://www.mathworks.com, vol., **2000**.

[267] S. Jayaram, S.G. Kapoor and R.E. DeVor, *Estimation of the specific cutting pressures for mechanistic cutting force models*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 265-281, **2001**.

[268] R.G. Reddy, R.E. DeVor and S.G. Kapoor, *A mechanistic force model for combined axial-radial contour turning*, International Journal of Machine Tools and Manufacture, vol. 41, pp. 1551-1572, **2001**.

[269] The Self-Organizing Map, *Helsinki University of Technology, Neural Networks Research Centre*, http://www.cis.hut.fi/nnrc, **2000**.

[270] S. Cho, *Ensemble of structure-adaptive self-organizing maps for high performance classification*, Information Sciences, vol. 123, pp. 103-114, **2000**.

[271] J. Kangas and T. Kohonen, *Developments and applications of the self-organizing map and related algorithms*, Mathematics and Computers in Simulation, vol. 41, pp. 3-12, **1996**.

[272] A.S. Tolba and A.N. Abu-Rezeq, *A self-organizing feature map for automated visual inspection of textile products*, Computers in Industry, vol. 32, pp. 319-333, **1997**.

[273] K. Kiviluoto, *Predicting bankruptcies with the self-organizing map*, Neurocomputing, vol. 21, pp. 191-201, **1998**.

[274] U.R. Kulkarni and M.Y. Kiang, *Dynamic grouping of parts in flexible manufacturing systems - A self-organising neural networks approach*, European Journal of Operational Research, vol. 84, pp. 192-212, **1995**.

[275] Talysurf 10 Operators Handbook, , Rank Taylor Hobson Limited, Leicester, England, pp. 241-7, **1995**.

A. An overview of commercial sensors for TCM

A.1 Introduction

Although many researchers have considered the complex problem of TCM in many different approaches, a need still exists for adequate sensors for TCM. Many authors state that force and vibration sensors yield the best results in a TCMS, although a number of different approaches proposed by others also yielded satisfactory results. This includes methods like ultrasonic sensing, laser focus methods, temperature measurements and optical methods and Acoustic Emission (AE) monitoring. However, most of these sensors seem to be operational only in laboratory setups. In the case of AE, many authors have had success in establishing a reliable AE methodology for TCM, overcoming the sensor problems. The use of AE for TCM has also found its way into industry. However, there is still a lack of understanding / interpreting the AE signals generated during machining. As was shown in Chapter 4, the trends in the AE signal due to tool wear is not very consistent, and it will not always be possible to utilise AE for any TCM problem.

As far as the other methods are concerned, they are not without their respective shortcomings. For instance, the presence of cutting fluid and chips breaking away from the workpiece have a very large impact on the performance of vision-based systems. Although tool wear and cutting temperature are closely related, it will not be wise to design a TCMS based only on this feature. Also, the presence of coolant will make this approach impossible. Figure A.1 shows a sensor selection table for TCM suggested by Montronix [245]. The table suggests force measurement for tool wear monitoring. However, the performance of any sensor is always process dependent, and therefore this table must not be used as the ultimate sensor reference guide.

The sensor problem in the case of vibration and cutting force still require further developments. Dimla [41] state that the monitoring of both vibration and cutting force is of indispensable value to successful TCM. In the case of the cutting force, the static and dynamic component of the force is of importance. The dynamic behaviour of the cutting process is embodied in the vibration and the dynamic cutting force. In industry only the static cutting force is used for TCM, but this is mainly due to the fact that the physical measurement, analysis and interpretation of the dynamic component is more complex.

Due to nature of the machining environment, very robust sensors are required to monitor tool vibrations and cutting forces. Usually, a need exists to monitor as close to the machining process as possible to minimise the effect of external influences on the sensor performance. However, cutting fluid and chips breaking away from the workpiece can damage a sensor beyond repair.

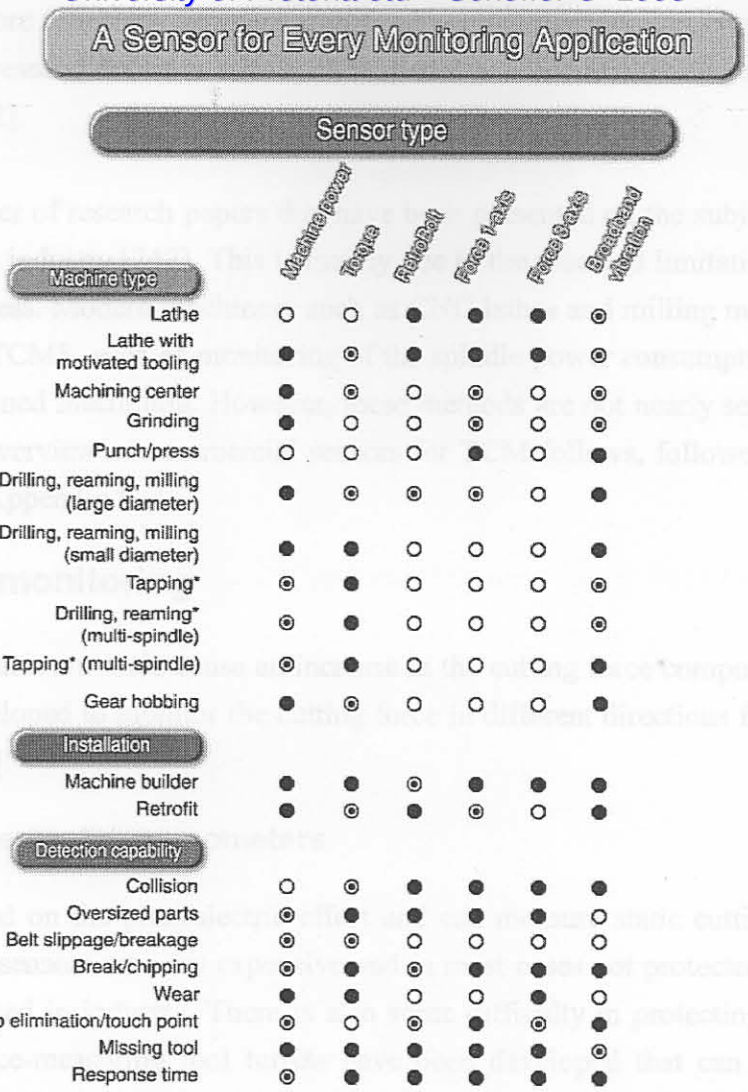


Figure A.1: Sensors for TCM, as suggested by Montronix

Key:

- ● Common, recommended application
- ⊙ Acceptable depending on application, but may not be preferred solution.
- ○ Unusual application - generally not recommended.

Also, in order to develop a TCMS for use in industry, sensors must have a realistic lifetime of at least a few years (also refer to sensor requirements listed in Chapter 3). Many modern sensors comply with some of these requirements, but a vibration or force sensor that complies with all of these requirements does not exist yet. This gives rise to another issue: Should a new sensor be developed specifically for TCMS, or should the designer of the TCMS attempt to get best results with available sensors?

The difficulty with the first question is the fact that there are so many different processes. Thus, a process specific sensor will probably not justify the costs involved. However, in the case of cutting force monitoring, Kistler [246] developed a range of sensors to monitor cutting forces for particular processes like turning, milling, drilling, grinding and even micro-machining. In the case of accelerometers, a need still exists for more robust and smaller sensors that can be integrated into the tool holder. A few sensors developed by Kistler and PCB are available that can be used for TCM applica-

tions. The design of more reliable sensors to continuously monitor machining processes is a topic for future research. Some researchers have already identified this niche market and are currently working in this area [32,103, 232].

Despite the large number of research papers that have been presented on the subject of TCM, very little found their way into industry [247]. This is mainly due to the practical limitations to obtain reliable sensor data on the process. Modern machinery such as CNC lathes and milling machines are equipped with relatively simple TCMS, such as monitoring of the spindle power consumption for detecting tool breakage during unmanned machining. However, these methods are not nearly sensitive enough to estimate tool wear. An overview of commercial sensors for TCM follows, followed by an overview of commercial TCMS in Appendix B.

A.2 Force-based monitoring

It is well recognized that worn tools cause an increase in the cutting force components. Many types of sensors have been developed to monitor the cutting force in different directions for a number of processes. These include [5]:

A.2.1 Direct measurement dynamometers

These sensors are based on the piezoelectric effect and can measure static cutting forces very accurately. However, these sensors are very expensive and in most cases not protected from overload, and therefore not widely used in industry. There is also some difficulty in protecting the sensors against cutting lubricants. Force-measuring tool turrets have been developed that can measure three force components, but these still have a very high cost. Kistler is the leader in supplying force dynamometers for machining processes.

Most academics active in the area of TCM have used the Kistler dynamometers for cutting force measurements. These are usually two or three axis quartz dynamometers. Quartz is a natural crystalline material with extreme ruggedness and ultrastability. The piezoelectric force measuring system differs fundamentally from most other methods. The forces acting on the quartz elements are directly converted into proportional electrical signals, and the resulting displacement amounts to a few thousands of a millimetre. Dynamometers are very accurate and can measure quasi-static forces. Their frequency range is unfortunately fairly low, allowing dynamic force measurements to about 1kHz. Many different types of dynamometers exist, developed for specific machining operations.

The Kistler Three Component Dynamometer 9121, shown in Figure A.2, is a quartz three-component dynamometer for measuring the three orthogonal components of a force. The force components are measured practically without any displacement. The sensors are fitted so that they are ground isolated. This largely excludes ground loop problems [246]. The dynamometer is corrosion-resistant and protected against penetration by spray water and cutting fluid. The control unit of the dynamometer is easy to operate and contains a power pack and a keyboard with status displays together with a connector for signal input. The output voltages are proportional to the forces occurring at the tool tip. Figure

A.3 shows a recently developed three-component dynamometer for hard turning [248].

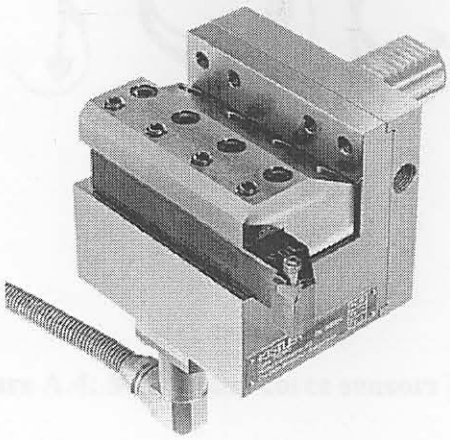


Figure A.2: Kistler Three Component Dynamometer 9121 [246]

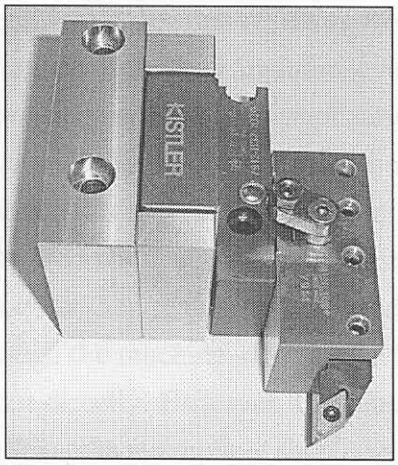


Figure A.3: Kistler 7914 dynamometer for hard turning [248]

The IntelliTool developed by Sandvik is a force sensor integrated tool holder for machining centres, developed around 1995. A spokesperson from Sandvik confirmed the existence of the IntelliTool, but stated that it is not in production anymore. Sandvik is currently not exploring further developments of the IntelliTool concept.

A.2.2 Sensors in plates and rings

Force-measuring plates can be fitted relatively easy on turning machines between the turret housing and the cross slide. These thin plates are fitted with piezoelectric force measuring sensors. This approach has some advantages, but is subject to many disturbing factors such as thermal expansion. Montronix provide a fairly wide range of force sensors for TCM, primarily used for fixed tool applications such as lathes, presses, and broaches. Typical sensors from Montronix are shown in Figure A.4.

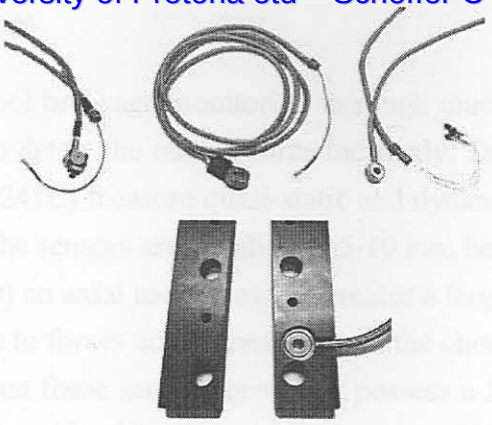


Figure A.4: Montronix force sensors [245]

The Montronix RetroBolt is a retrofit force sensor typically used on turret and tool block lathes. This single axis sensor is installed as a washer under a bolt head. The selected bolt must be in the load path. The sensor measures the small changes in bolt tension resulting from cutting forces exerted on the tool. Typical monitoring performance using the RetroBolt includes overload detection, machine protection and detection of catastrophic tool breakage in roughing applications. Examples of ring-shaped force transducers from Montronix that can be fitted on a CNC machine are shown Figure A. [245]. Kistler also offers a number of ring-shaped sensors that can be fitted in a similar way onto CNC machines. Some sensors offer the possibility for 3-component force monitoring. An advantage of using these sensors is that their frequency range is higher than dynamometers, and the cost is lower. The disadvantage is that force must be transmitted through the machine structure to the sensor, in contrast to dynamometers, which measure the cutting forces directly. However, the sensors are very robust and can be used for a wide range of machining operations. An example is shown in Figure A.6.

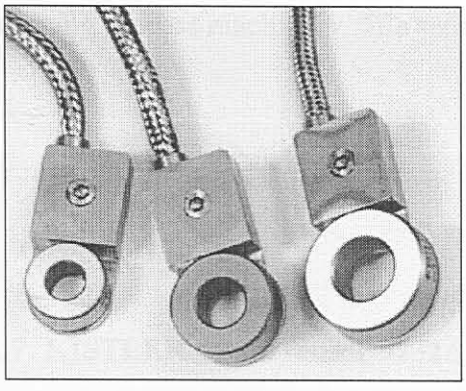


Figure A.5: Examples of ring-shaped force transducers for TCM from Montronix [245]

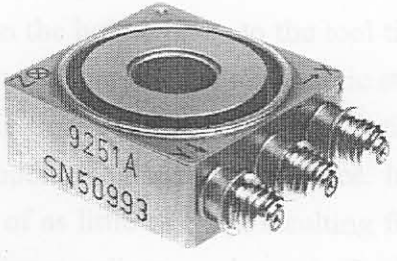


Figure A.6: Force sensor from Kistler [246]

A.2.3 Pins, extension sensors

These sensors are suitable for tool breakage monitoring in rough machining. They are fitted on force-carrying machine components to detect the cutting force indirectly. Transverse and longitudinal measuring pins (*e.g.* Kistler 9243A/9241C) measure quasi-static and dynamic strains in the structure of machine components or fixtures. The sensors are installed in 5-10 mm holes. Owing to the larger measuring range (length of pressure bar) an axial measuring pin creates a larger measuring signal than a radial pin. Both sensors are insensitive to forces acting transverse to the chosen axis. In view of the unavoidable interference and the fact that these sensors generally possess a low level of sensitivity, they are normally only suitable for breakage identification during rough machining and for press force monitoring. The identification of a suitable fitting position can only be determined experimentally, which is another disadvantage.

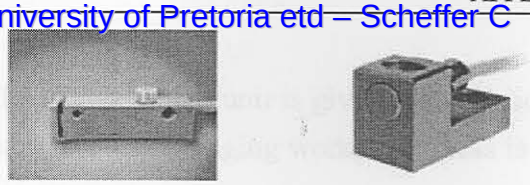
A.2.4 Measurement of displacement and strain

Non-contact sensors to measure the displacement or bending of tools can be mounted directly on the tool holder. However, these sensors are subjected to the high risk of damage and disturbances due to chips, dirt and cooling lubricant. It is relatively easy to retrofit various strain and displacement sensors. Their working principle is indirect force measurement. A quartz strain transducer can be mounted on a part of the machine where the mechanical stress is large and disturbances are low.

All force transmitting parts in machines and fixtures are deformed elastically by the forces acting on them. As a result of this deformation, a detectable displacement occurs on the surface of the force transmitting parts. Because of high machine rigidity, the elastic deformations are extremely small. A strain sensor development for measuring dynamic and quasi-static forces on stationary and moving machinery is the Kistler 9232A strain sensor. Its high sensitivity and acceleration-compensated design allows a sensor application on fast running process machinery. The sensor is shown in Figure A.7.

Figure A.7: KISTLER strain sensor 9232A [246]

The PCB dynamic strain sensor has also proved its worth in TCM applications [165]. These are very small sensors that can be adhered on the holder close to the tool tip. Another advantage of this sensor is the fact that it is extremely sensitive and can measure dynamic strain accurately in the range of 5Hz - 10 kHz. One disadvantage though is the fact that they cannot estimate static strains, which would be useful for estimating the static component of the cutting force. Induction-based distance sensors are capable of resolving displacements of as little as 5 nm resulting from forces acting on machine parts. Examples of such sensors from leading suppliers are shown in Figure A.8.



NORDMANN BDA-Kralle PROMETEC DSA 50

Figure A.8: Induction-based displacement sensors

Resistance strain gauges are widely used for stress analysis, but are not often found in TCM applications. A standard design of such a sensor consists of a 0.025 mm layer of polyimide film that covers the gauge except that portion of the tab necessary for lead attachment. The polyimide overlay provides protection of the sensing element during installation handling, and promotes better long-term stability with the foil grid protection from airborne contaminants or fingerprints. Figure A.9 shows how HBM resistance strain gages are applied to the shaft of milling tools [248,239].

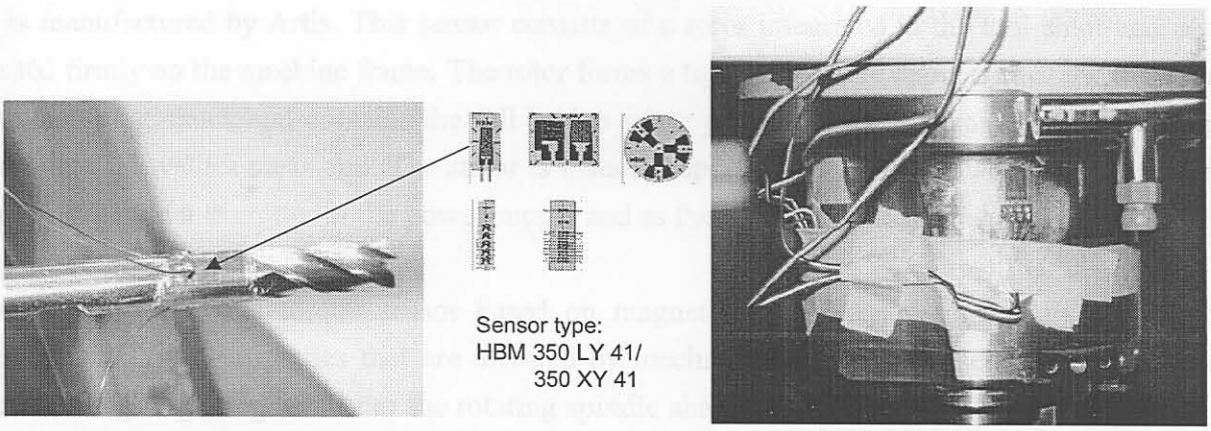


Figure A.9: HBM resistance strain gage application [248]

A.2.5 Force-measuring bearings

Bearings and bushes can be specially fitted with strain gauges in certain positions to measure cutting forces. Force-measuring bearings require a low-pass filter due to disturbances from the ball contact frequency, and therefore high frequency signal processing is not possible. The force-measuring bushes are only accepted in special cases because they reduce the rigidity of the machine.

A.2.6 Force and torque at spindles

These systems can be very complex because they have to monitor the torque of the spindle with high resolution and within the entire range of the motor. Furthermore, the signal must be transmitted on a non-contact basis. The installation of such a system is not possible on most machines due to the limited space available for sensor mounting. The Kistler two-component high-speed dynamometer type 9125A is a part of a rotating cutting force measuring system and allows measurements of both the axial feed force and torque on the rotating tool. The instrument consists of a two-component piezoelectric sensor, a built-in charge amplifier and a built-in two-channel telemetry system for wireless data transmission. The main advantages of such a machine tool integrated, rotating cutting force measuring system are:

- The force measurement is directly on the rotating tool holder, i.e. the signal is recorded close to

the tool tip.

- A constant mass of the measurement unit is given, i.e. a steady resonance frequency of the dynamometer is realised, since the changing workpiece mass is not connected to the dynamometer.
- Low and constant inertia forces are acting during the cutting process.
- Any orientation of the dynamometer is possible.

Much more accurate tool and process monitoring can be achieved by measuring mechanical torque directly instead of measuring the power consumed by the spindle motor. It is especially useful for tapping and a multi-spindle application where power monitoring is insufficient. The sensor can monitor tool wear, tool breakage, did-not-cut condition, thread depth, oversized or undersized pre-drilled holes, and damaged or missing threads on taps.

A torque tool sensor that uses strain gauges for torque measurement and non-contact signal transmission is manufactured by Artis. This sensor consists of a rotor integrated in the tool shaft and a stator mounted firmly on the machine frame. The rotor forms a tight ring in the upper part of the tool shaft or of the clamping chuck and contains the full bridge strain gages and the electronics for acquiring and transmitting the values measured. The stator is installed approximately 5 μm away from the rotor and serves both as the transmitter of the power supply and as the receiver of the signal.

Montronix introduced a torque sensor based on magneto-elastic sensing methods (Accu-Torque), which has magnetic properties that are affected by mechanical torque. The sensor includes a small torque-sensing ring integrated onto the rotating spindle shaft, and non-rotating pick-up. The ring converts mechanical shaft torque into a linearly proportional magnetic field. The pickup converts this field into a linearly proportional electrical signal and acts as a non-contact means of gathering shaft torque information.

Excellent rotating quartz four-component (Fx, Fy, Fz and Mz) dynamometers for measuring cutting forces and torque on a rotating tool spindle are also available from Kistler. However, because of their high cost they can be considered as laboratory tools only. Examples are shown in Figure A.10.

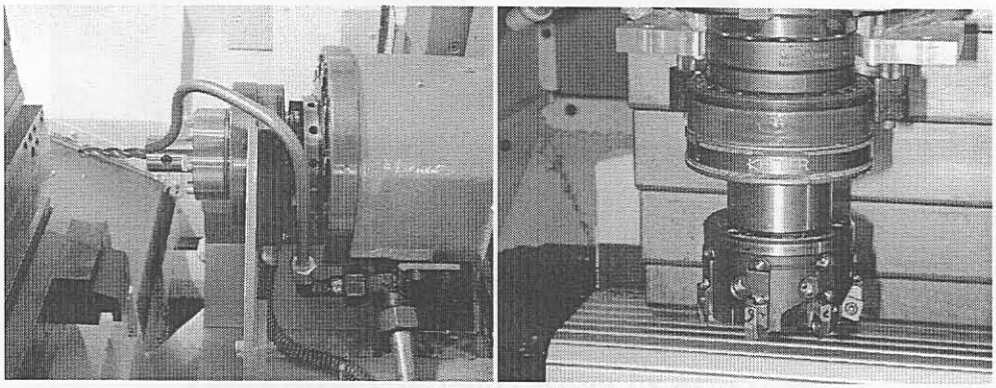


Figure A.10: KISTLER RCD 9123 sensor used in drilling and milling applications [246]

A.3 Measurement of motor current

The measurement of motor current is an easy alternative to the above-mentioned systems and can be installed without much difficulty. A wide range of sensors is available for this purpose. However, due to fluctuations in the signal due to friction, this approach is not accurate. Furthermore, tool breakage is not detected directly, but only after damage has occurred. Spindle power is also proportional to the cutting force in the primary motion, which is not always sensitive enough for TCM. The cutting process consumes only a small portion of the measured power of the spindle, which also makes monitoring difficult. However, monitoring systems based on the principle of spindle current can be successful when used with an appropriate process. Such sensors are often found in the automotive industry in drilling and tapping applications. A typical sensor for estimating the spindle power is shown in Figure A.11 [249].

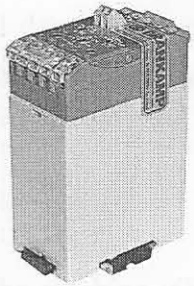


Figure A.11: Hall-effect sensor for measuring spindle power [249]

A power sensor measures the spindle or axis drive power for AC, DC or variable frequency motors (frequency range 0-200 kHz). In all commercial applications currently available, power is measured directly by means of a voltage and current measurement. The power sensor is installed directly in the electrical cabinet, and measures with 1-3 Hall effect sensors. A typical application is shown in Figure A.12.

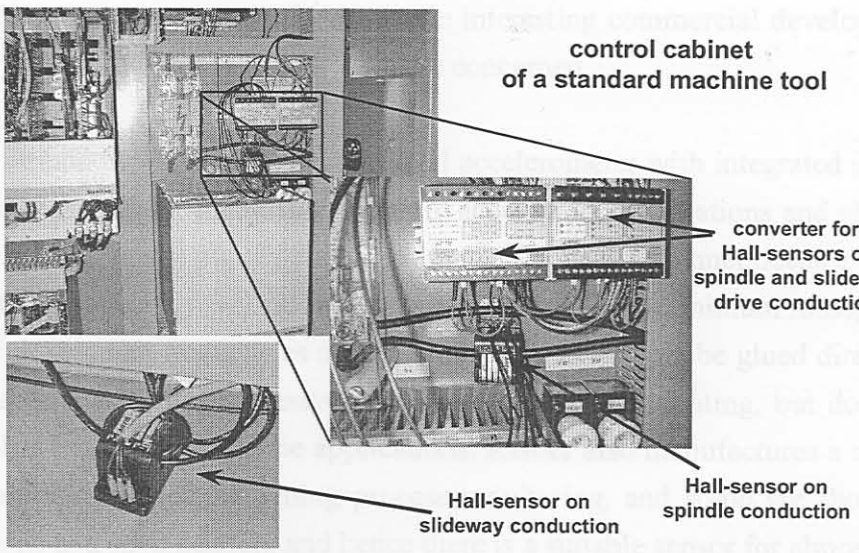


Figure A.12: True power sensor by Artis [250]

Use of three balanced Hall effect sensors eliminates large phase shift errors. When two sensors in two

phases are used, the third current quantity results from the two that are measured, making it possible to take the usual grid fluctuation between three phases into account. Effective power measurement has the advantage over simple current measurements in that the idle current, which provides no information about the motor load on the tool, is not measured. Power is linear, so that a change in motor load is a change in power. Current is not a sensitive indicator of power at low loads in three-phase motors. The sensor can provide an indication of a missing tool or tool wear in certain applications. Using an additional logarithmic signal amplifier, small tools can also be monitored. If the power-measuring curve is wavy, or displays ripples, peaks or even brief collapses (<1s), the power measuring value can be smoothed. The quality of effective power monitoring depends largely on the relationship between the cutting power and the nominal drive power of the motor. This means that small tools (such as drills with a diameter of less than 1mm) can be monitored only on tight fitting spindle drives. Owing to the inert masses, the output signal has a low-pass filter characteristic. Therefore, tool breakage is not detected directly, but only after consequential damage has occurred. In many cases, force, vibration or AE signals have to be used as complementary information to extend the capability of the monitoring system.

A.4 Acceleration measurement

Piezoelectric accelerometers can measure the machine vibration caused by oscillations of the cutting forces. It is known that vibration levels change with tool wear. Some accelerometers are inexpensive and can measure vibration levels within a very wide frequency range. For these reasons, accelerometers are often used for TCM [5,41]. Miniature accelerometers available on the market are not always robust enough to withstand cutting fluid and chips for very long. The ones that may be able to withstand the environment are either too big or cannot measure in the frequency range of interest. The sensor cables must also be watertight, and preferably inside a steel shielding. Hence, a dedicated accelerometer must be used for machining process monitoring. The development of an accelerometer integrated into a tool holder is definitely a worthwhile research topic. Such a system was recently proposed by Lago [32]. Lately, there have been some interesting commercial developments as far as accelerometers for machining process monitoring are concerned.

The Kistler Piezotran® 8694 is a miniature triaxial accelerometer with integrated impedance converter and low impedance output for measuring dynamic accelerations, vibrations and shocks. The sensor is shown Figure A.13. For power supply and signal processing each channel needs a Piezotron® coupler. The accelerometer is especially well suited for applications where minimum mass, small mounting dimensions and high resonant frequencies are essential. The sensor can be glued directly to the object or to mounting adapters. A contact adhesive is recommended for mounting, but double sided adhesive tape or wax is also acceptable for some applications. Kistler also manufactures a range of other accelerometers that are suitable for machining process monitoring, and some are shown in Figure A.14. These vary and size and specification, and hence there is a suitable sensor for almost any application.

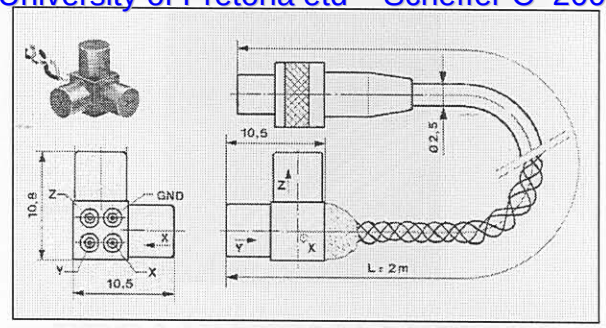


Figure A.13 KISTLER Piezotran® 8694 accelerometer [246]

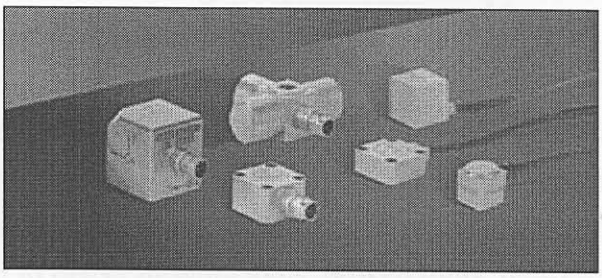


Figure A.14: Kistler three-component accelerometers for TCM [246]

The IMI 356A11 miniature triaxial ICP® accelerometer, depicted in Figure A.15, is a miniature, four-gram, hermetically sealed titanium accelerometer. The hermetic titanium case provides a rugged, reliable sensor for minimal mass loading during triaxial vibration measurements. The 10mV/g output from ceramic shear mode sensing elements provides a wide measurement range from 2 to 10 000 Hz (±5%) in the "Z" axis and 2 to 7 000 Hz (±5%) in the "X" and "Y" axes. The 10ft integral shielded cable terminates in a four-pin connector. This robust and small sensor would work well for TCM applications.

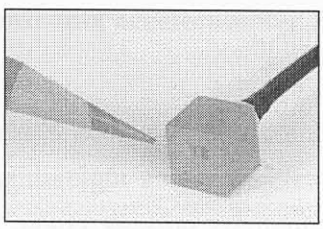


Figure A.15: IMI 356A11 Miniature Triaxial ICP® Accelerometer [251]

An accelerometer that can be used in a corrosive environment compatible with 304 Stainless Steel (SS), and survive long term submersion in water (including the cable) is ENTRAN's EGS Series. Its case is fabricated from 304SS with welded seals and its cable is sheathed with a 3.2mm diameter soft annealed SS tube, which can be easily bent or shaped for routing around corners. The sensor is shown in Figure A.16.

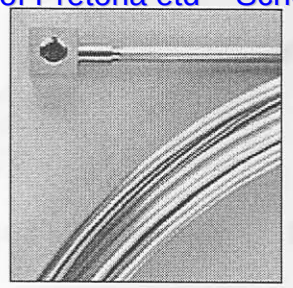


Figure A.16: Entran EGS series [252]

Montronix supplies the BV100 sensor designed specifically for application on machining centres that use a wide variety of tooling. The sensor is shown in Figure A.17. The sensor boasts an extremely broad frequency range for improved sensitivity to the condition of small and large diameter tools. It is easily to install: The 5 meter cable is rated to IP67 and NEMA 6 industrial standards and is shielded against electro-magnetic interferences, hence 100% inner and outer shielding. An inner polyurethane jacket protects against coolant damage and the stainless steel verbraid protects against hot chips and sharp corners. It can be used to detect missing tools, broken tools, out-of-tolerance parts, machine collision and severe process faults. It is also possible to monitor excessive vibration on bearings or spindles. The vibration sensor is easy to install on new or existing machines.

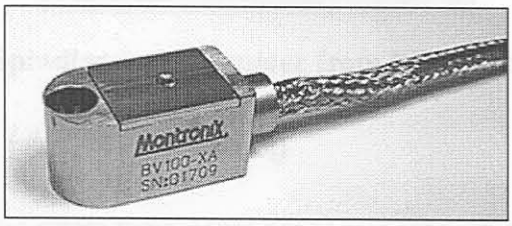


Figure A.17: Montronix accelerometer BV 100-XA [245]

The Kistler 8730A500 miniature K-Shear® accelerometer combines a reliable, easy to handle connector with an integral stud to create a lightweight accelerometer that fits into narrow areas [246]. The hermetically sealed titanium case and quartz sensing elements ensure proper mechanical protection. The sensor is shown in Figure A.18

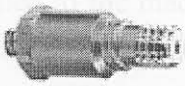


Figure A.18: Kistler 8730A500

Prometec supplies a number of sensors for TCM, optimised for use with their other TCM hardware. These sensors include: force, torque, displacement, strain, position, hydraulic pressure, power, AE, and vibration. All of these sensors are robust, and designed specifically for the purpose of TCM [253]. The force and vibration sensors from Prometec are shown in Figure A.19.

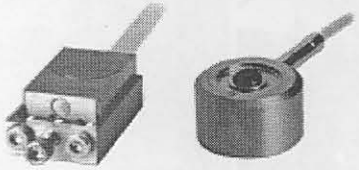


Figure A.19: Force and vibration sensors from Prometec [253]

The ICP compatible "Spindler" accelerometer from PCB, shown in Figure A.20, is especially designed for machining process monitoring. The accelerometer is provided with a stainless steel cable shielding that can withstand cutting fluids and chips. Furthermore, the sensor can swivel around its base, making it easy to install in limited space applications.

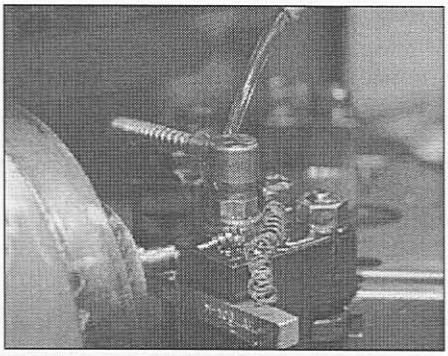


Figure A.20: Spindler accelerometer from PCB piezotronics [251]

A.5 Acoustic Emission (AE) monitoring

Cutting processes produce elastic stress waves that propagate through the machine structure. Different sources in the cutting process generate these stress waves known as Acoustic Emission (AE). AE is defined to occur in the range of 50kHz - 250kHz. An AE sensor measures the high-frequency energy signals produced by cutting processes. When a tool breaks, the sensor also measures the AE energy resulting from the fracture. An AE sensor is best suited to applications where the level of background AE is low compared to that of tool breakage. The sensor is easy to install on both new and existing machines. In combination with true power, it increases the reliability of breakage monitoring significantly. It is used especially with solid carbide tools, or very small tools on large machines and multi-spindles. Most AE sensors have to be attached to the machine tool surface (sometimes with different mounting variants, *e.g.* side, top, or bottom connection, and spring disk fixing). An example of mounting for a drilling application is shown in Figure A.21.

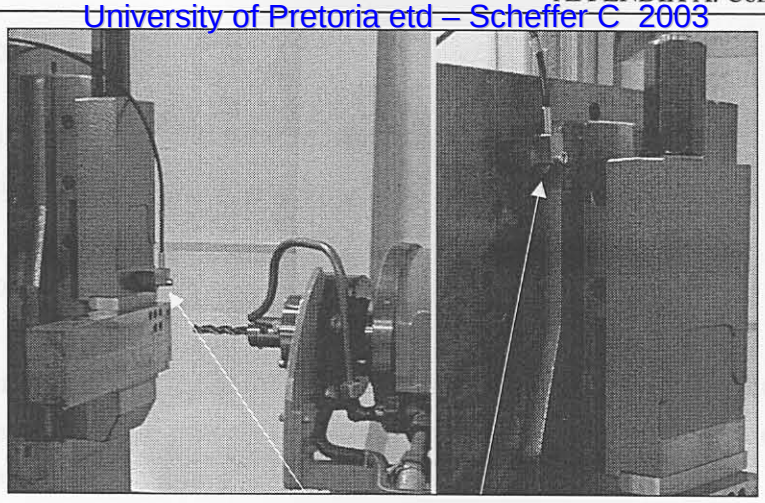


Figure A.21: Kistler AE 8152A1 sensor mounted

The primary and secondary cutting shear zones are important sources of AE. In the presence of flank wear, the interface between tools and work materials becomes an additional zone of AE generation due to intensive friction between the tool and the workpiece surface which move past one another at relative high velocities. Changes of cutting conditions also affect the behaviour of AE signals.

There are a number of ways through which AE waves are transmitted. The Prometec AEL 200 is a rotating, wireless AE sensor. It is suitable for applications where signals from moving parts have to be passed to a fixed receiver for analysis and monitoring. When applied in grinding, the sensor enables precise detection of any sparking, which enables optimisation of the depth setting. Another approach is the Prometec fluid sound sensor WAE100, an AE sensor that receives the acoustic waves via a jet of cooling lubricant, which can be connected directly to the tool (or workpiece). The sensor and an example of its mounting is shown in Figure A.22. Its advantage over the conventional AE sensors is that it measures close to the tool tip. The fluid AE sensor allows detection of high frequency stress waves from moving or rotating workpieces or components, or from materials with very rough surfaces. In some cases the acoustic waves in the coolant stream are damped so much by the air bubbles that the resulting measurement is too low. Particles of metal and dirt, on the other hand, do not cause problems.

Although announced a few years ago, a dual-mode sensor for the simultaneous measurement of AE and one to three orthogonal force components never reached the production line. The advantage of this dual-mode sensor would be to compliment the force measurement with AE.

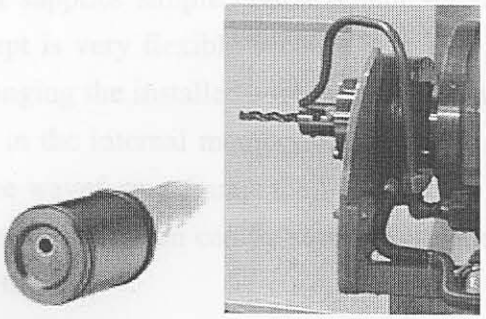


Figure A.22: Prometec WAE 100 [253]

A.6 Other methods

Other methods for indirect / direct tool wear monitoring are:

- non-contact capacitive sensors
- vision systems
- laser scatter methods
- stereo imaging

At the time of the review, none of these methods were available commercially, besides vision systems, which are available for machining process monitoring in general, but not dedicated to TCM. However, examples of these systems can be found in the literature (refer Chapter 3).

A.7 Latest developments

A.7.1 SeTac

The Sequoia Triaxial Acceleration Computer (SeTAC) is a device that includes three accelerometers on each Cartesian axis and a microprocessor for the signal analysis and data communication. SeTAC belongs to the category of smart sensors, which offer a direct connection between signal recording and processing [239,248]. Regarding the signal understanding process, the SeTAC sensor aims at integrating the filtering functionality into the sensor. The SeTac functionality is schematically depicted in Figure A.23.

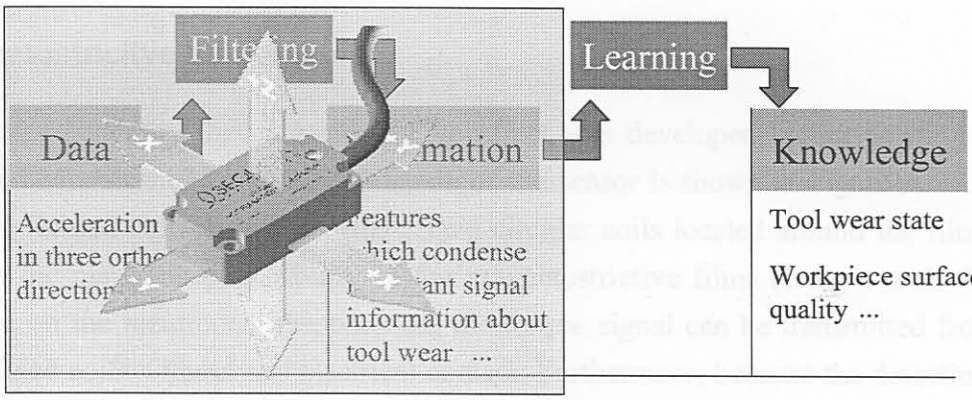


Figure A.23: SeTAC tri-axial accelerometer

The accelerometer signals are processed in the SeTAC by a built-in micro-controller that is programmable and fully customisable. It supplies simple digital output suitable for common devices, PLCs, actuators *etc*. The SeTAC concept is very flexible because it is able to monitor and evaluate a wide range of phenomena by only changing the installed software. A common Personal Computer (PC) can be used to upload the programs in the internal memory. The PC can even operate as a user interface where it is possible to display the waveform, change the filtering and set the thresholds. When the required behaviour is achieved the configuration can be saved in the memory and the SeTAC is able to operate without connection with the PC.

A.7.2 Intelligent tool coating

Various new developments are based on the notion to integrate sensors into the tool coating, *e.g.* German BMBF project IDEE or the EMO 2001 presentation of Kyocera Fineceramics Gmbh. Such a direct measurement would offer the advantage to exclude a large degree of external disturbances that are normally reflected in any indirect measured signal pattern. The tool coating has to be restructured by adding several layers of different materials without changing the quality of the coating. By integrating the sensor in the tool coating the conversion of signal data to process knowledge is extremely simplified. An example is shown in Figure A.24.

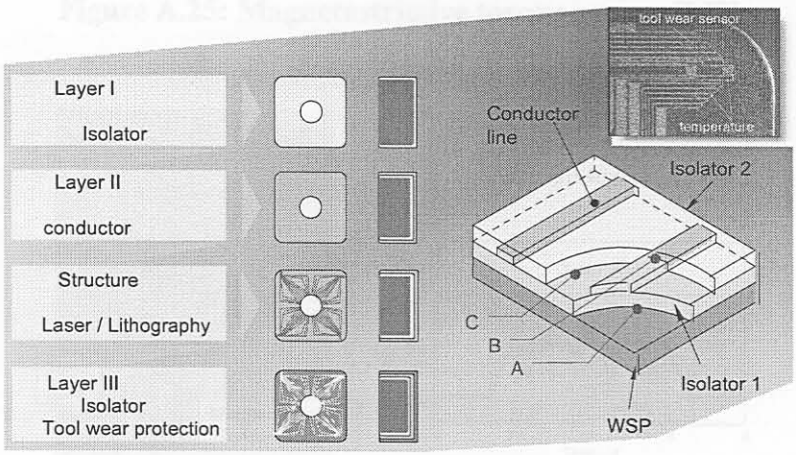


Figure A.24: Intelligent tool coating – direct tool wear and temperature measurement [254]

A.7.3 Magnetostricitve torque sensor

A torque sensor based on the magnetostrictive effect was developed by cooperation between Keio University and Mitsubishi Materials. A diagram of the sensor is shown in Figure A.25. Magnetostrictive films are formed onto the sensor shaft. Two circular coils located around the films measure the torque, since the magnetic permeability of the magnetostrictive films changes with the variation of torque. Based on the mentioned phenomenon, the torque signal can be transmitted from the rotating shaft to the fixed coils without any electrical contact. Furthermore, because the detecting elements are close to the cutting point of the tool, the magnetostrictive torque sensor provides a high quality cutting torque signal being independent of the workpiece location.

In order to simplify application of the sensor system, the torque sensor is applied to the tool holder. A prototype of the system together with example signals from a drilling application is shown in Figure A.26. The development of a magnetostrictive sensor that additionally sense a three component cutting force being applied on the cutting edge of a rotating tool is also under way.

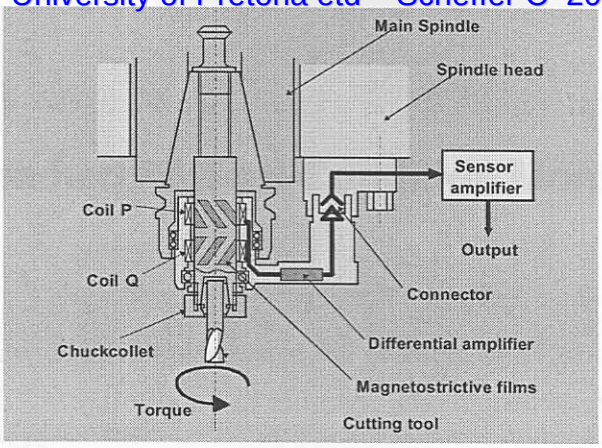


Figure A.25: Magnetostrictive torque sensor [248]

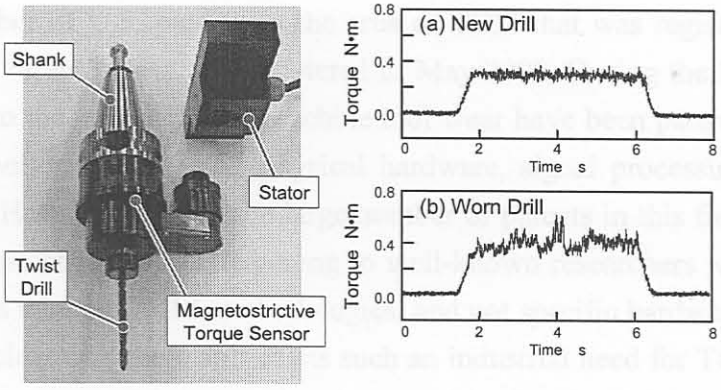


Figure A.26: Magnetostrictive torque sensor prototype [248]

B. Overview of commercial Tool Condition Monitoring Systems

B.1 Introduction

Numerous researchers worldwide have investigated the problem of TCM. However, due to many practical limitations, not many TCMS are found in industry. The objective of this appendix is to give an overview of commercial TCMS.

B.2 U.S.A. Patents

Table B.1 lists a number of U.S. patents in the area of TCM that was registered within the last ten years [255]. The most recent patent was registered in May 2000. During the last 15 years, more than 50 inventions related to the monitoring of machine tool wear have been patented in the U.S.A. These inventions include monitoring strategies, physical hardware, signal processing techniques and even mathematical models. However, despite the large number of patents in this field, only a few commercial versions exist. Most of these patents belong to well-known researchers within the field of TCM, and most of the patents relate to TCM methodologies, and not specific hardware. When looking at this list of patents, it is unclear why there still exists such an industrial need for TCMSs. The only reasonable explanation is that the patents were never developed into commercial products, either due to the costs involved or limited accuracy or applicability in industry. It is also often unclear if these methods were tested under conditions resembling a realistic industrial environment. If the systems were not tested under typical shop floor conditions, it remains a question whether they will be effective in industry.

Table B.1: U.S.A. patents related to TCM in the last 10 years

no.	*Description*
6059494	Tool bit monitoring system for machine tools **(2000)**
6055484	Tool monitor and assembly qualifier
5904457	Detecting tool wear by thermal monitoring of workpiece
5773949	Cutter fracture detecting system
5587931	Tool condition monitoring system
5542304	Magnetostrictive torque sensor magnetostrictive torque measuring apparatus and condition-monitoring apparatus for a cutting tool using the same
5414632	System and method for predicting failure in machine tool
5361308	3-D measurement of cutting tool wear
5266929	Metal cutting tool wear indicator method and system
5251144	System and method utilising a real time expert system for tool life prediction and tool wear diagnosis
5212645	Flexible real-time multi-tasking architecture for tool condition monitoring
5076102	Tool monitor **(1991)**

B.3 Commercial TCMS

B.3.1 Montronix

Montronix supply a fairly wide range of TCM products that can be used with most CNC machines. However, the signal processing methods proposed by Montronix do not employ the state-of-the-art methods available today. A comparison between the on-line and a reference signal is generally used as a tool state classification strategy. Furthermore, the positioning of the sensors on machines not always justified properly. Although the systems supplied by Montronix seem to be the best available on the market, it lacks the use of modern signal processing techniques that will enhance the capabilities of the system.

The systems supplied by Montronix are either stand-alone systems, or integrate with the CNC machine. A diagrammatic representation of a Montronix system is shown in Figure B.1. In Figure B.2, some of the hardware for process monitoring is shown. The integrated process monitor works with open architecture or PC-based machines where the Montronix software directly loads into the operator interface. A monitoring unit without graphical display is available for other machines (without open architecture), or an LCD display unit can be used. These systems are mounted on top of the machining centres. A remote display unit is also available [245].

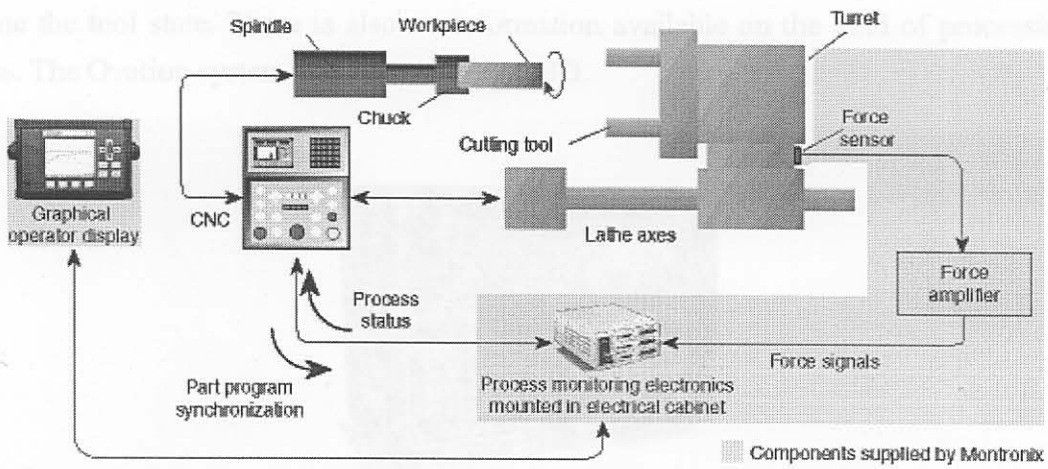


Figure B.1: Montronix TCM scheme for lathes [245]

B.3.2 Ovation

Ovation engineering developed a CNC tool monitor that can monitor the state of a range of tools that must be pre-programmed. The features of the Ovation CNC Tool Monitor includes [256]:

- Easy to use and program.
- Very fast response time (less than 10msec).
- Upload and download of all limits via RS232.
- Saves the last 300 alarms with date and time stamp.
- Extreme, wear and undercut limits for every tool and section.
- Up to 99 tools and 5 sections per tool.
- Parallel or serial machine interface.

- Monitoring of coolant flow and spindle speed.
- GE Specification Power monitoring.

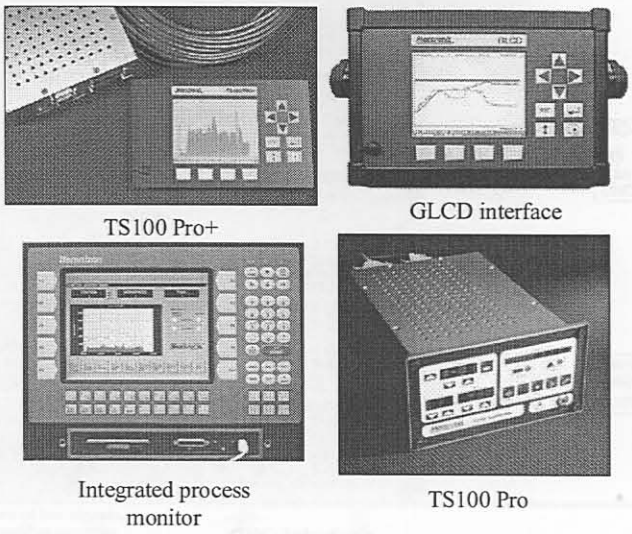


Figure B.2: Some of the TCM products from Montronix [245]

Additional transducers can be added to allow a different range to be used for smaller tools or for live tooling on lathes, and adaptive control can be added. The system only uses spindle power monitoring to determine the tool state. There is also no information available on the kind of processing that the system uses. The Ovation system is shown in Figure B.3.

Figure B.3: Ovation Engineering CNC Tool Monitor [256]

B.3.3 Artis Systems Inc.

Artis Systems specialises in tool and process monitoring. Their ranges of equipment for monitoring and diagnosis technology are used in the automotive industry worldwide. Artis supplies sensors, monitoring modules and data handling hardware. In Figure B.4, the range of hardware supplied by Artis is shown. The data visualisation can take place either through the CNC (open architecture) or a PC [250].

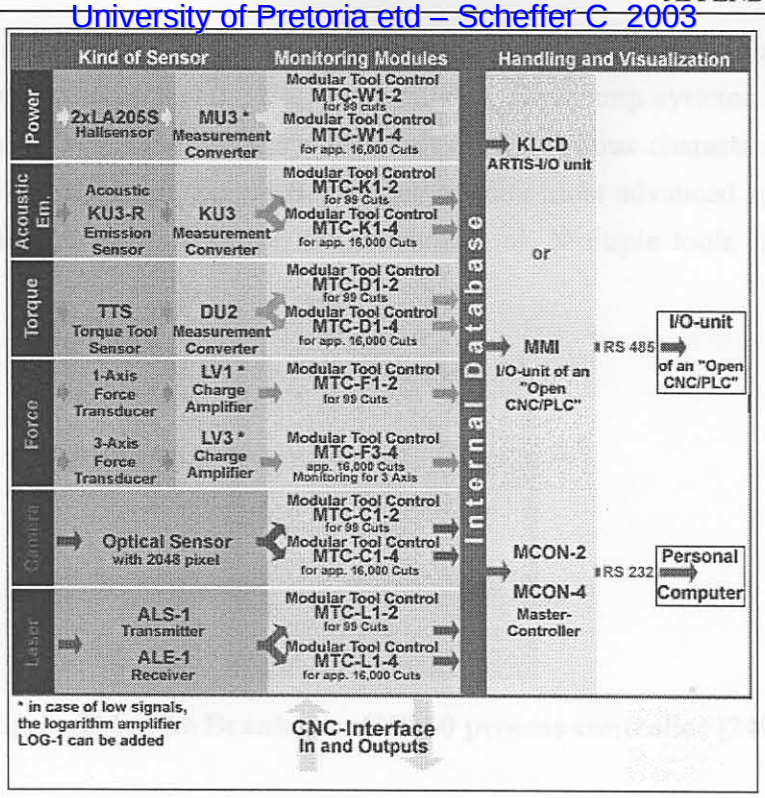


Figure B.4: Artis hardware for TCM [250]

Figure B.5 shows a diagrammatical layout of the Artis approach to TCM. The supported sensors are power, torque, force, AE, cameras and lasers. The tool control modules can be programmed for certain cuts.

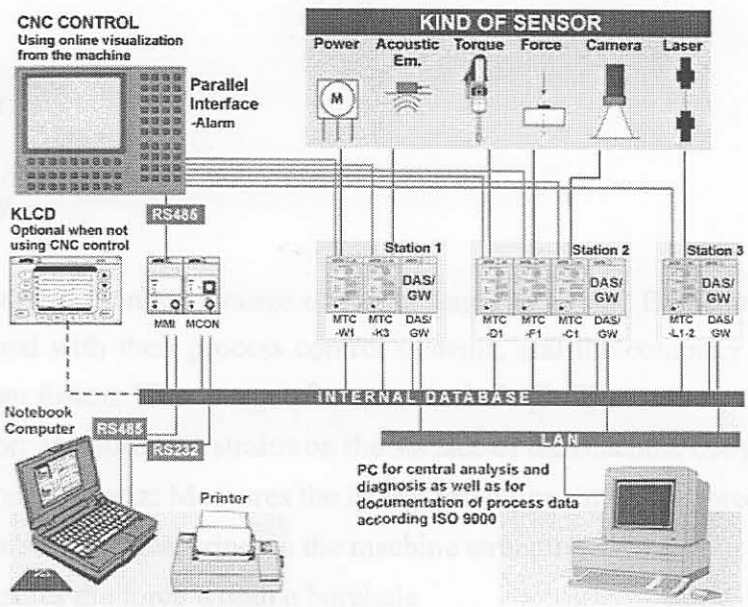


Figure B.5: Schematic representation of Artis TCM methodology [250]

B.3.4 Brankamp

Brankamp is a well-established German company that focuses on process control and monitoring. They offer a wide range of hardware for product quality, computer aided quality control software, and monitoring units mostly with large LCDs to allow process observation. The units are designed for all

cutting processes: presses, stamping machines, cold forming and general assembly and manufacturing equipment. A range of hardware for TCM is also available. Brankamp systems are either open architecture type of controllers, or stand-alone systems that boasts various characteristics. The Brankamp C8060 process controller, shown in Figure B.6, is one of their most advanced systems. The LCD display can show a variety of information on the machine tool. Multiple tools can also be monitored [249].

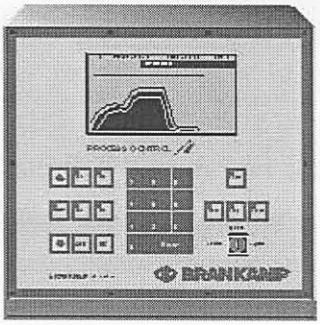


Figure B.6: The Brankamp C 8060 process controller [249]

This monitoring system has four sensors or monitoring channels. The system presents a lot of different process visualisations, for instance the envelope curve on multistage production machines. The whole process is wrapped up and monitored against the envelope limits. The following screens can be seen on the LCD display:

- Actual process values
- Process memory
- Tool life counter
- Productive counter
- Stop/Go-diagram
- Stop-code-memory

Brankamp offers sensors to monitor a range of machining operations. Brankamp sensors are also developed to work optimal with their process control systems, and the company attempts to emphasise the advantages of sensor fusion. Their range of sensors include [249]:

- Piezoelectric sensor: Measures the strains on the surface of the machine body
- One/Three- component-quartz: Measures the lateral force flow in one or three directions
- Magnetic inductive sensor: Measuring on the machine structure
- Vario sensor: Measures the force within a borehole
- Acoustic emission sensor: Measures high frequency noises
- Brankamp power sensor: Measures the power of motors on machine tools
- Micro sensor: Measures the strains on the surface of the machine body

B.3.5 Prometec

Prometec is a German company specialising in the fields of tool and machine monitoring, as well as process analysis for turning, drilling, milling, and grinding. Their systems are widely used in the auto-

motive industry. Their tool monitoring systems are also complemented by new systems for the detection of oscillations and imbalance. Prometec supplies a number of sensors for TCM, optimised for use with their other TCM hardware. These sensors include: force, torque, displacement, strain, position, hydraulic pressure, effective power, AE, structure-borne sound and vibration. All of these sensors are robust, and designed specifically for the purpose of TCM [253].

The TCMS supplied by Prometec consist of operator panel modules, machine interface modules and monitoring modules. The processing of the sensor signals are conducted by the monitoring modules, which work independently of one another and are assigned to a machining station. Various monitoring modules are available for differing extents of monitoring or numbers of cycles (tools or cuts). All monitoring modules are supplied in a compact housing for mounting in switch cabinets. Single dual channel monitoring modules are available. This allows, for example, a cost-effective solution for a twin-turret drilling machine by using a dual-channel monitoring module.

The monitoring module's notification of breakage, wear *etc.* are transmitted via a machine interface module. The display of the changes in the signals occurs over time, and an operator panel module is necessary for data visualisation. This allows for the adjustment of the monitoring parameters to the process at hand. This is done first of all during the teach-in where the progression of the signals is learned and the limits then defined. The characteristics of these systems are shown in Figure B.7. Figure B.8 shows the monitoring module, the machine interface module, the operator panel module and the PROVIS visualisation tool supplied by Prometec [253].

Operator panel module	OPM12	OPM 20 with PROVIS	Open CNC with PROVIS
Display	LCD, 4 x 20 characters	VGA, color	PROVIS software in the operator unit of the open NC. Connection to monitor module via serial interface. For DOS, Windows 95 and Windows NT 4.0.
Program-dependent data storage	—	x	
RS 232 terminal for notebook, modem etc.	x	x	
Printer terminal	—	x	
Terminal for PC keyboard	—	x	
External connector module for the terminals listed above	Option	Option	
Protective housing	Option	Option	

Machine interface module	MINI	MIDI	BAPSI	PROFI-BAPSI INTER-BAPSI
1-channel				As with BAPSI, but serial operation in accordance with PROFIBUS-DP/ INTERBUS-S protocol
Cycles	8	—	250	
Output messages	3	—	7	
2-channel				
Cycles per channel	2	30	250	
Output messages per channel	1 Plus supplementary, joint output for collision message	3	7 Two BAPSI interface modules are to be used for 2-channel operation	
Connecion to machine interface module	Ribbon cable	2-wire cable	2-wire cable	2-wire cable

Figure B.7: Prometec TCMS [253]

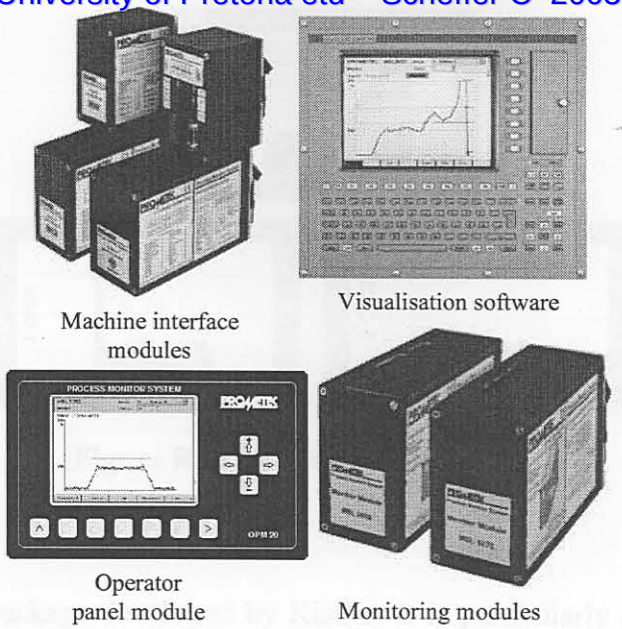


Figure B.8: TCM modules by Prometec [253]

B.4 TCM software

B.4.1 Introduction

A number of companies supply Windows supported software for TCM. Most of these software packages only enable relatively simple data processing. The software programs are usually graphical interfaces displaying the sensor signals and a statistical summary of events. Normally, these packages are used to view trends in sensor signals over a period of time, and to perform comparative TCM studies. Table B.2 summarise the visualisation methods supplied by various companies (from Jemielnaik [247]). Some are discussed in more detail.

Table B.2: Visualisation methods applied in TCM systems [247]

Visualisation method	Supplier					
	Artis	Brankamp	Kistler	Montronix	Nordmann	Prometec
Simple digital / bar display	●	●		●	●	●
Graphic display	●	●	●		●	●
Via open CNC control system	●			●	●	●
Via PC / Notebook computer	●	●	●	●		●

B.4.2 Montronix

The M-View package from Montronix was developed to enhance the performance of the Montronix range of TCM products. M-View is a typical visualisation tool for trends and comparative TCM studies. Typical windows from the M-View package are shown in Figure B.9.

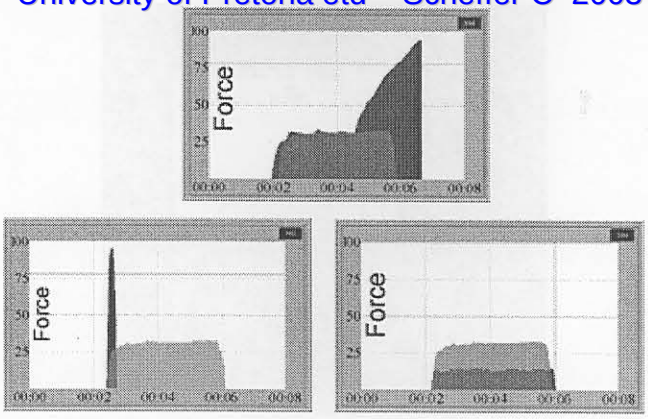


Figure B.9: M-View windows [245]

B.4.3 Kistler

DynoWare is a software package developed by Kistler. It is particularly suitable for force measurements with Kistler dynamometers and force sensors. DynoWare may be used to analyse, optimise and document cutting force measurements in a simple and efficient way. Programmed in the Windows environment, DynoWare is a powerful data acquisition system that can be used universally. Eight free selectable measuring signals may be displayed and analysed in several windows. DynoWare contains an easy to use integrated database containing default set up values for the most common factors. (sensitivity, scale, *etc.*). DynoWare also features [246]:

- 8 measuring channels with 12-bit resolution each.
- Measuring range per channel: ±10 V.
- Max. sampling rate of A/D card, divided by number of channels selected.
- 100 kHz/ number of channels.
- Remote control of Kistler electronics via RS-232C or IEEE-488 interface.
- Several measuring cycles can be recorded during a trial.
- Import / Export of measured data in ASCII format (Labview, Dia Dago, Excel).
- Trigger with keyboard, analogue, digital.

B.4.4 Prometec

The PROVIS software by Prometec offers facilities for visualisation of the process signal to enable the diagnosis and optimisation of machining operations. PROVIS software are available in Windows or DOS versions for all of the leading control manufacturers including Siemens, Indramat, GE-Fanuc, Schleicher *etc.* Figure B.10 shows an example of a PROVIS window [253].

B.5 Monitoring strategy

Monitoring strategies applied by different suppliers can be grouped into six approaches, summarised in Table B.3 (after Jemielniak [247]). Most of these are based on static limits. Static parameters are process parameters that remain fixed during the processing of the current workpiece. The basic monitoring strategies will be discussed in more detail.

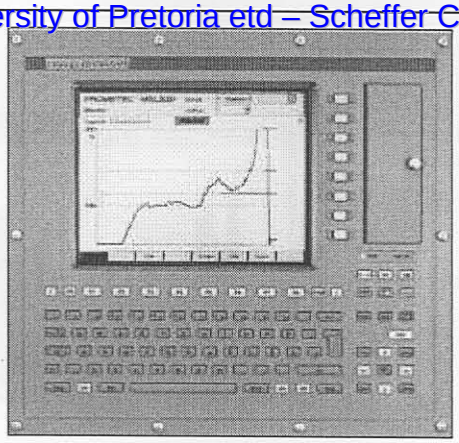


Figure B.10: PROVIS software by Prometec, in operator display panel [253]

Table B.3: Strategies applied in TCMS [247]

Strategy	Supplier					
	Artis	Brankamp	Kistler	Montronix	Nordmann	Prometec
Simple fixed limits				●	●	●
Time defined limits			●		●	●
Part signature	●	●	●	●	●	
Pattern recognition				●		
Wear estimator				●		
Dynamic limits						●

B.5.1 Simple fixed limits

This simple strategy is based on fixed limits that apply to the signals from the sensors (after simple pre-processing). During machining of the first workpiece, the monitor reference signal is constructed, by normalising it between 0 and 100 % for each cut and each tool. The positions of limits relative to the set standards are defined (manually or automatically). For example, three limits can be set, one for a worn tool, one for a broken tool, and one for collision. Instead of fixed limits, floating limits can also be used which track cycle-to-cycle trends by using information from the current cycle to adjust the limits of the next cycle automatically.

B.5.2 Time defined limits

Fixed limits for a machining process can also be time displaced, which means that limits for certain sub-areas of the signal are set. This can enhance the performance of the fixed limit strategy to detect sudden tool wear, a premature cut, incorrect adjustments or sudden tool fracture.

2.5.3 Part signature

The cutting of a certain part is divided into segments using stepped limits based on time and / or position. This can create a "part signature" that can more closely track a complex cutting cycle than a single set of fixed limits. The part signature is like a combination of floating and time defined limits. Whenever the current signal has higher or lower values of the reference signature, an alarm goes off. Figure B.11 shows an example of the part signature concept from Montronix [245].

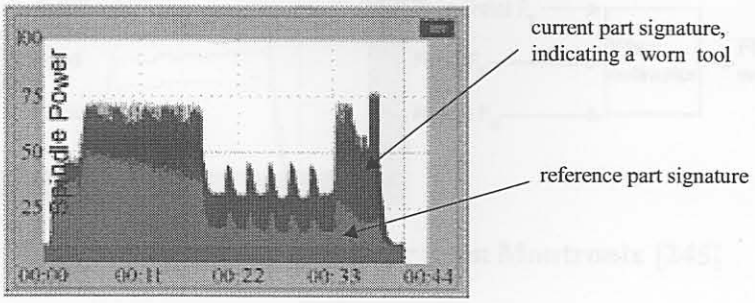


Figure B.11: Part signature concept from Montronix [245]

B.5.4 Pattern recognition

Montronix developed a tool breakage monitoring strategy based on pattern recognition for turning. The system stores a number of reference signals indicative of tool breakage, such as breakage of a carbide or ceramic tool. The system continuously monitors the current signal for one of the breakage patterns. When the pattern occurs, the machine stops. This method has the advantage that the pattern is independent of the process signal value, and therefore it can be applied to a number of different cuts. It is therefore also independent of the tool condition and other machining parameters. Figure B.12 displays an example of the pattern recognition process by Montronix [245].

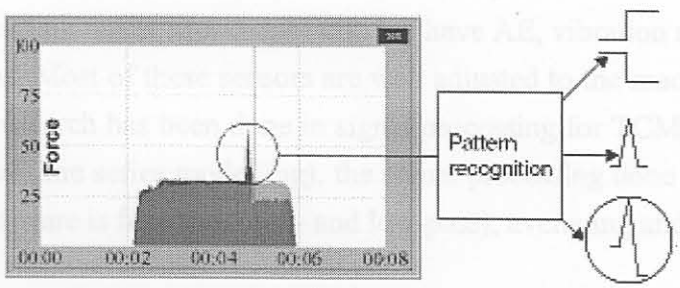


Figure B.12: Pattern recognition for tool breakage detection in turning form Montronix [245]

B.5.5 Wear estimator

The wear estimator is a proprietary technique developed by Montronix used for turning tool flank wear estimation. The method uses the relationship between all three cutting force components, and requires a three-component force sensor. The technique can distinguish between normal tool wear and variations in process parameters. These variations, such as workpiece hardness or runout, can adversely affect other techniques such as part signature analysis. Figure B.13 shows the wear estimator concept from Montronix [245].

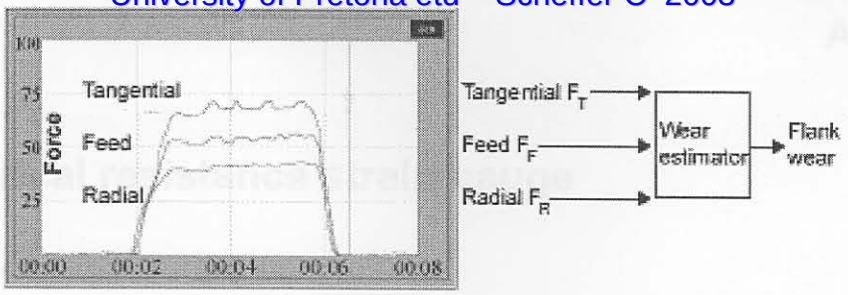


Figure B.13: Wear estimator from Montronix [245]

B.5.6 Dynamic limits

The PROMOS system for Prometec incorporates the use of dynamic limits for tool breakage detection. Two dynamic limits, above and below the monitoring signal, follow the signal continuously for every load level. In the case of an extremely fast crossing of one of the two dynamic limits, the tool state is determined by visual comparison of the monitoring signal. Another version of the dynamic limit by Prometec, is when a special feature is generated from the force signal. The monitoring strategy then combines the feature generation with dynamic limits, to enhance the performance of the TCMS to automatically adjust over a range of sensor signals in different machining situations.

B.6 Conclusion

The most commonly used sensors for TCM in commercial systems are sensors measuring force or quantities related to the cutting force. Most suppliers also have AE, vibration and optical sensors available for some applications. Most of these sensors are well adjusted to the machining environment. Despite the fact that much research has been done in signal processing for TCM, (such as spectral analysis, statistical analysis and time series modelling), the signal processing done by commercial suppliers of TCM hardware and software is filtering (high- and low-pass), averaging and rms calculations.

Since constant limit methods only work when all restrictions remain constant, the use of self-adjusting limits is more appropriate. Only Prometec with its "dynamic limit" method and Montronix with the pattern recognition and wear estimator use more sophisticated signal processing, although these are not nearly the state-of-the-art methods in modern signal processing and data classification. Monitoring strategies developed by researchers nowadays almost always incorporate more sophisticated signal processing where features are extracted from sensor signals. In commercial systems, the one-sensor, one process approach dominates. The processing and fault detection techniques on the sensor signals in commercial TCMS primarily remain independent. Only the Montronix wear estimator uses more than one signal to classify the tool condition.

C. The electrical resistance strain gauge

C.1 Introduction

The most common type of strain transducer is the electrical resistance strain gauge. Strain gauges are extremely versatile due to their small size, low cost and the fact that they are available in a wide range of configurations for different applications. A diagram of the single element strain gauge is shown in Figure C.1.

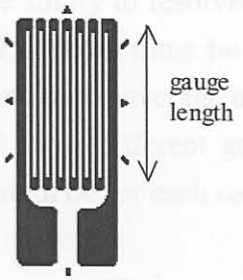


Figure C.1: Typical single element strain gauge

Strain gauges are usually connected in a Wheatstone bridge configuration. A basic Wheatstone bridge configuration is shown in Figure C.2.

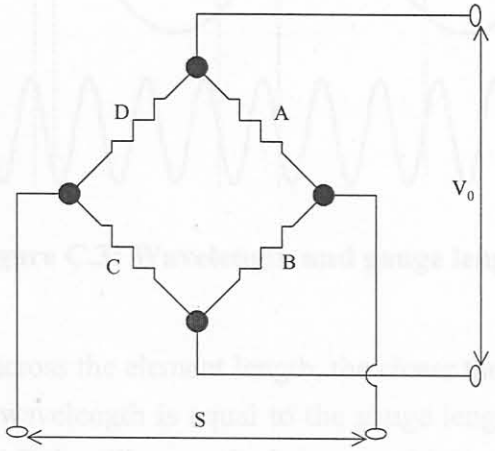


Figure C.2: Strain gauge bridge configuration

In Figure C.2, four strain gauges are represented by A, B, C and D, and their resistances would typically be 120Ω each. A voltage V_0 is supplied across the bridge, and this is typically 10V. The signal S is then measured across the other ends of the bridge. A change in resistance of the gauge is caused by the base strain (elongation or compression) of the surface it is adhered to. If all four gauges (A, B, C and D) are adhered to the surface it is called a full-bridge configuration. If two gauges are attached to the surface, the other two are "dummy" gauges, and this is called a half-bridge configuration. If only one is adhered to the surface it is called a quarter bridge. Many other configurations are possible and

are used for particular applications. The types of strain gauge configurations and their applications are beyond the scope of this text. However, there are at least three significant considerations when using strain gauges for measuring dynamic signals. The following limitations should be considered before choosing a strain gauge:

- Frequency resolution of the gauge
- Spatial resolution of the gauge
- The so-called rise time of the gauge

C.2 Frequency resolution

The ability to resolve a wave spatially becomes limited for pulses near the order of the gauge length. As the gauge length increases, so does the ability to resolve the spatial location of the pulse decrease. Therefore the total length of the pulse of concern must be much greater than the gauge length. The output of a strain gauge tends to give an integrated average of the strains imposed over its length. Consider the three different wavelengths and three different gauge lengths, shown in Figure C.3. What would the average strain over the gauge length be for each case?

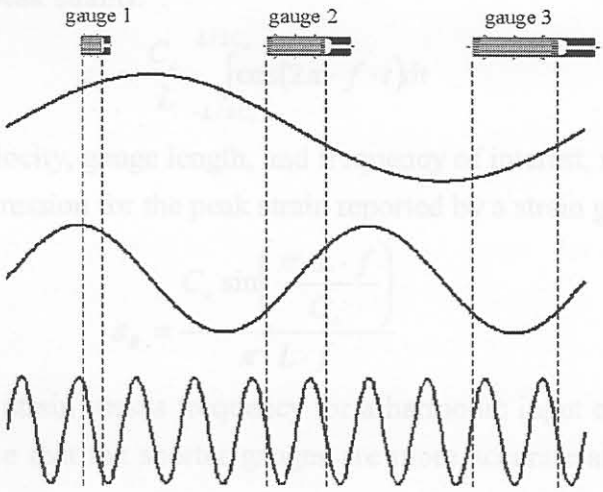


Figure C.3: Wavelength and gauge length

The smaller the strain gradient across the element length, the closer the output will be to the true strain. For the special case where the wavelength is equal to the gauge length, the average output would be zero, but the actual strain is not! This will occur for integer multiples of the wavelength. As the wavelength decreases, so too does the ability to resolve the peak strain due to the averaging effect over the gauge length. The peak strain is consequentially always estimated low. The concept is illustrated in Figure C.4 - the steeper the gradient (i.e. short wavelength), the less the peak strain can be resolved [257].

C.3 Spatial Resolution

For harmonic waves propagating through a strain gauge, the peak output from the gauge will always occur when the peak strain is centred in the gauge. To determine the closed form solution for the peak

strain reported by a strain gauge, the harmonic wave must be integrated over the gauge length.

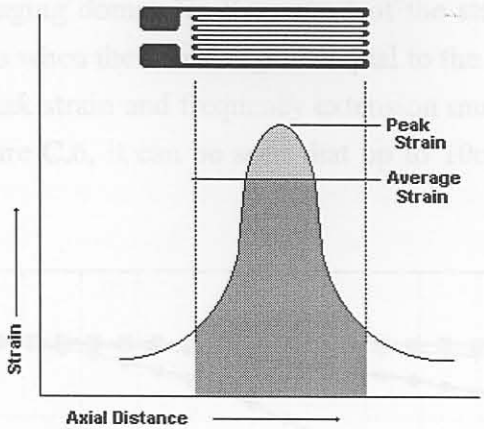


Figure C.4: Averaging effect on peak strain [258]

By recognizing that the peak of a cosine wave coincides with the centre of the gauge, the following expression yields the peak strain reported by the strain gauge (remember the peak output of the gauge is always less than the true peak strain):

$$\varepsilon_R = \frac{C_o}{L}\int_{-L/2C_o}^{L/2C_o}\cos(2\pi \cdot f \cdot t)\mathrm{d}t \tag{C.1}$$

C_o, L, and f are the wave velocity, gauge length, and frequency of interest, respectively. The evaluation of this integral yields an expression for the peak strain reported by a strain gauge:

$$\varepsilon_R = \frac{C_o \sin\left(\frac{\pi \cdot L \cdot f}{C_o}\right)}{\pi \cdot L \cdot f} \tag{C.2}$$

A plot of the reported peak strain versus frequency for a harmonic input of amplitude one for steel is shown in Figure C.5. Notice that the shorter gauges are more accurate at measuring high frequency strains.

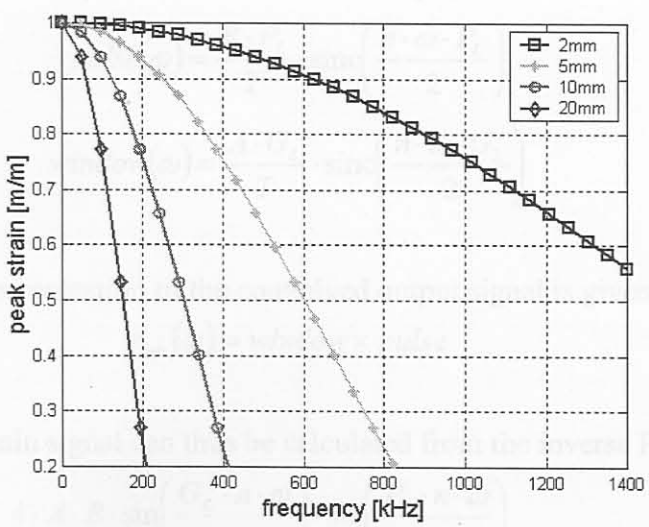


Figure C.5: Reported peak strain from various gauge length strain gauges

Also note that accurate static (DC) strain values are reported. As frequency increases, however, the reported strain becomes drastically attenuated. As was expected, shorter gauges offer better frequency response. The effect of the averaging dominates the output of the strain gauge for high frequencies. The output from the gauge is zero when the wavelength is equal to the gauge length. To choose a strain gauge, the desired accuracy of peak strain and frequency extension must be considered. Looking at the 0-50kHz region of the plot, Figure C.6, it can be seen that up to 10mm gauges can be used on steel with reasonable accuracy.

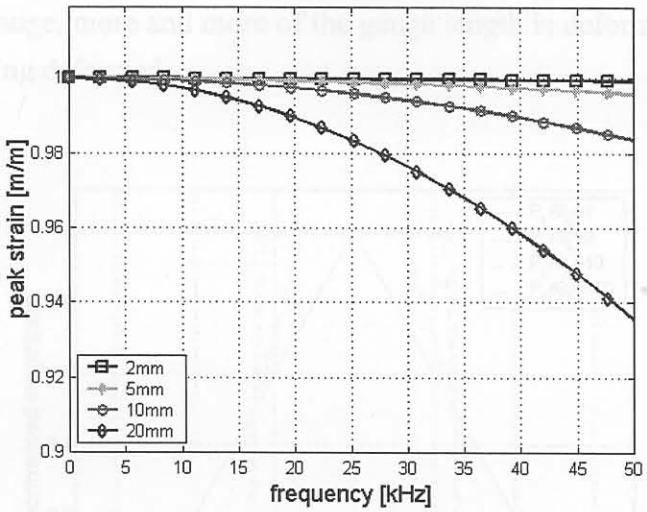


Figure C.6: Reported peak strain from various gauge length strain gauges

Another condition (related to the physics of the strain gauge) exists that distorts the output of strain gauges. This is a windowing effect that distorts the actual strain pulse. The effective gauge length represents a rectangular window that in effect changes the shape of the true strain pulse. To calculate the effect of windowing requires convolution. Recall that convolution in time is equal to multiplication in frequency. Hence, the inverse Fourier transform of the product of two functions of frequency is equal to the convolution of the two same functions in the time domain. Consider a square pulse of magnitude B and length P_L / 2, measured by a rectangular window of magnitude A and length G_L / 2, and both pulses having period T. The Fourier series of the pulses and window are:

$$pulse(\omega) = \frac{B \cdot P_L}{T} \cdot \text{sinc}\left(\frac{n \cdot \omega \cdot P_L}{2}\right) \tag{C.3}$$

$$window(\omega) = \frac{A \cdot G_L}{T} \cdot \text{sinc}\left(\frac{n \cdot \omega \cdot G_L}{2}\right) \tag{C.4}$$

The frequency domain representation of the convolved output signal is given by:

$$\varepsilon_{out}(\omega) = window \times pulse \tag{C.5}$$

The convoluted time domain signal can thus be calculated from the inverse Fourier transformation:

$$\varepsilon_{out}(t) = \sum_{-\infty}^{\infty} \frac{4 \cdot A \cdot B \cdot \sin\left(\frac{G_L \cdot n \cdot \omega}{2}\right) \cdot \sin\left(\frac{P_L \cdot n \cdot \omega}{2}\right)}{T^2 \cdot n^2 \cdot \omega^2} \exp(-i \cdot n \cdot \omega \cdot t) \tag{C.6}$$

C.4 Rise time of strain gauge

By plotting this rather complex equation, the windowing effects of the strain gauge can be seen, as shown in Figure C.7. This assists to predict gauge output for many configurations of gauge lengths (G_L) and pulse lengths (P_L). Figure C.7 is a plot of the convolved output of various pulse lengths to gauge length ratios. Notice that all of the convolved pulses exhibit increased rise time, not characteristic of the square input pulse. This rise time is due to the period of time when the pulse just begins and just exits the strain gauge. Just before entering the gauge, the strain gauge output is zero. As the pulse progresses through the gauge, more and more of the gauge length is deformed, until the point when all of the gauge length is being deformed.

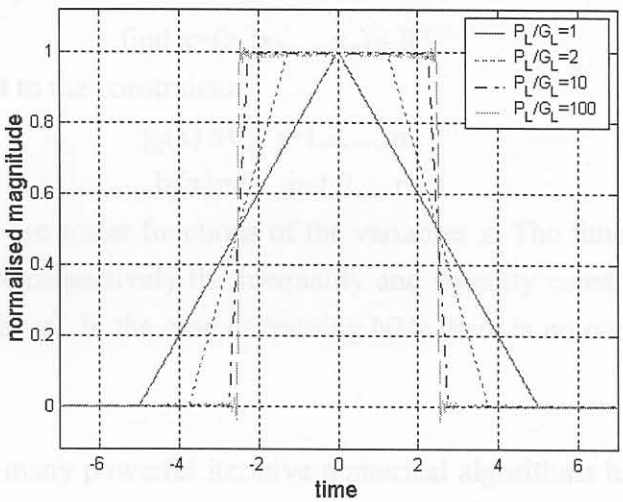


Figure C.7: Convolved output of strain gauges

In the limiting case where the gauge length equals the pulse length, the maximum value is only reached only once because there exists only one time when the gauge window and pulse window perfectly overlap. It can be seen that the windowing effect can be minimized by selecting a sufficiently large P_L / G_L ratio. However, with the measured strain signals in hand, corrections for the windowing effect of the gauge can also be employed by dividing the Fourier transform of the strain signal by the Fourier transform of the gauge measurement window [257].

C.5 Conclusion

With reference to the strain gauges selected for measurements as described in Chapter 4, these considerations were taken into account when selecting the gauges and also when analysing the signals obtained from them. In essence, very small gauge lengths were used compared to the fairly low frequency range that was investigated. Thus, the effects of rise-time and frequency distortion were minimal. Furthermore, it should be mentioned that the selected strain gauges were self-temperature-compensated, which means that the traditional use of dummy gauges was not necessary.

D. Mathematical Optimisation

D.1 General formulation

Training neural networks is in actual fact nothing but solving an unconstrained optimisation problem. Many algorithms exist whereby Neural Networks (NNs) can be trained. During the course of this research, different algorithms were investigated for their feasibility to train the dynamic NNs. For on-line implementation, the algorithm must be robust, quick and stable. The general formulation of the mathematical optimisation problem with constraints is:

$$\text{find } \mathbf{x}=(x_1,x_2,\ldots,x_n)\in R^n \tag{D.1}$$

that minimizes $f(\mathbf{x})$ subject to the constraints:

$$g_j(\mathbf{x})\leq 0,\quad j=1,2,\ldots,m \tag{D.2}$$

$$h_j(\mathbf{x})= 0,\quad j=1,2,\ldots r \tag{D.3}$$

where $f(\mathbf{x})$, $g_j(\mathbf{x})$ and $h_j(\mathbf{x})$ are scalar functions of the variables $\mathbf{x}$. The function f is called the objective function and g_j and h_j are respectively the inequality and equality constraint functions. A local optimum solution is denoted by $\mathbf{x}^*$. In the case of training NNs there is no constraints and thus no g and h functions.

Over the past thirty years many powerful iterative numerical algorithms have been developed to solve the above-mentioned general problem. It is however true to say that no single algorithm dominates in being superior to all others when applied to different subclasses of the general problem. Depending on the degree of non-linearity, the presence of noise or discontinuities in the functions, the number of variables involved and the time required to evaluate the functions, some methods will be preferred above others depending on the efficiency, accuracy or reliability required. Difficulties were experienced with conventional methods, and for this reason new methods were investigated for training the dynamic NNs.

D.2 Algorithms

D.2.1 SQSD

The Spherical Quadratic Steepest Descent (SQSD) Method [259] was employed to train the dynamic network. The method is a gradient-only algorithm for unconstrained minimisation, which applies the steepest descent method to successive spherical quadratic approximations of the objective function. No explicit line searches are preformed which makes the method computationally very efficient. It has also been shown that method performs extremely well with ill-conditioned problems, which is sometimes the case with neural networks. The gradient functions of the neural networks were calculated with finite differences. The selection of the step size when using this technique requires some care, especially if the functions are not analytical (*e.g.* NNs). The algorithm provided excellent results in less time than conventional training procedures.

The algorithm can be described as follows by specifying the convergence criteria $\varepsilon_g, \varepsilon_x$, steplimit $d > 0$ and selecting a starting point $\mathbf{x}^0$. Set $c_0 := \left\|\nabla f\left(\mathbf{x}^0\right)\right\| / d$. Let $k := 1$ and proceed to (1).

(1) If $\left\|\nabla f\left(\mathbf{x}^{k-1}\right)\right\| < \varepsilon_g$, then $\mathbf{x}^* \cong \mathbf{x}^c = \mathbf{x}^{k-1}$ and stop; otherwise set:

$$\mathbf{x}^k := \mathbf{x}^{k-1} - \frac{\nabla f(\mathbf{x}^{k-1})}{c_{k-1}} \tag{D.4}$$

2) If $\left\|\mathbf{x}^k - \mathbf{x}^{k-1}\right\| > d$, then set

$$\mathbf{x}^k := \mathbf{x}^{k-1} - d\frac{\nabla f(\mathbf{x}^{k-1})}{\left\|\nabla f(\mathbf{x}^{k-1})\right\|} \tag{D.5}$$

if $\left\|\mathbf{x}^k - \mathbf{x}^{k-1}\right\| < \varepsilon_x$, then $\mathbf{x}^* \cong \mathbf{x}^c = \mathbf{x}^{k-1}$ and stop.

3) Set

$$c_k := \frac{2\left[f\left(\mathbf{x}^{k-1}\right) - f\left(\mathbf{x}^k\right) - \nabla^T f\left(\mathbf{x}^k\right)\left(\mathbf{x}^{k-1} - \mathbf{x}^k\right)\right]}{\left\|\mathbf{x}^{k-1} - \mathbf{x}^k\right\|^2} \tag{D.6}$$

if $c_k < 0$, then set $c_k := 10^{-60}$.

4) Let $k := k + 1$ and then go to step 1 for next iteration.

The algorithm was programmed in Matlab for training the neural networks.

D.2.2 ETOPC

The ETOPC algorithm can solve unconstrained and constrained problems by means of the conjugate gradient method. Constrained problems are solved via a penalty function formulation. The specific implementation used is that of Snyman [260] which is unique in the sense that it employs no explicit line searches. With the ETOPC conjugate gradient program the user may select either a Fletcher-Reeves or a Polak-Ribiére implementation. The ETOP (unconstrained) algorithm is formulated as follows [261]:

Assume x_0, Δt given and set $v_0 \leftarrow -\nabla F_0 \Delta t = s_0 \Delta t$. Do for $k = 0,1,2,....$ until convergence:

1) Set

$$x_{k+1} \leftarrow x_k + v_k \Delta t = x_k + s_k \Delta t^2 \tag{D.7}$$

2) Determine:

$$\Delta F_k = -v_k^T a_k' \Delta t \tag{D.8}$$

where:

$$a_k' = (a_k + a_{k+1})/2 \tag{D.9}$$

$$a_k = -\nabla F(x_k) \tag{D.10}$$

$$\theta = -\rho / (\Delta F_k - \rho) \tag{D.11}$$

$$x_{k+1}^* = x_k + \theta \Delta x_k / 2 \tag{D.12}$$

3) Set:

$$x_{k+1} \leftarrow x^*_{k+1} \tag{D.13}$$

4) Determine:

$$v_{k+1} \leftarrow \left(-\nabla F_{k+1} + \beta_{k+1} v_k / \Delta t\right)\Delta t = s_{k+1}\Delta t \tag{D.14}$$

with the Polak-Ribiere formula:

$$\beta_i = \frac{\left(\nabla F_i - \nabla F_{i-1}\right)^T \nabla F_i}{\left\|\nabla F_{i-1}\right\|^2} \text{ for } i = 1,2,... \tag{D.15}$$

The algorithm was implemented in Matlab for training the dynamic networks.

D.2.3 LFOP

This algorithm is the LFOPC algorithm of Snyman [261-263]. It is a gradient method that generates a dynamic trajectory path from any given starting point towards a local optimum. This method differs conceptually from other gradient methods, such as SQP, in that no explicit line searches are performed.

The original leap-frog method [262-263] for unconstrained minimization problems seeks the minimum of a function of n variables by considering the associated dynamic problem of a particle of unit mass in a n-dimensional conservative force field, in which the potential energy of a particle at point $\mathbf{x}(t)$ at time t is taken to be the function $f(\mathbf{x})$ to be minimized. The reader is referred to [261-263] for a mathematical formulation of LFOP.

The optimisation is terminated when either of the convergence criteria becomes active:

$$\|\Delta \mathbf{x}\| < \varepsilon_x \text{ or } \|\nabla f(\mathbf{x})\| < \varepsilon_g. \tag{D.16}$$

It is recommended that the step size, δ, be of the same order of magnitude as the "diameter of the region of interest", hence:

$$\delta = \frac{1}{10}\sqrt{\sum_i^n (\text{range}_i)^2} \tag{D.17}$$

where range_i is the size of the region of interest in the i-th variable direction. In practice the choice of μ and μ_{max} should be coupled to the accuracy required.

D.2.4 PSOA

The PSOA simulates the physical movement of social creatures, for instance the movement of a flock of birds. Consider a flock of p particles of birds. For particle i, Kennedy and Eberhart [224] proposed that the position $\mathbf{x}^i$ be updated as:

$$\mathbf{x}^i_{k+1} = \mathbf{x}^i_k + \mathbf{v}^i_{k+1} \tag{D.18}$$

where $\mathbf{x}$ represents the position of each bird, $\mathbf{v}$ the velocity, and k a unit pseudo-time increment, and would in this case be an inner step. Shi and Eberhart [264] proposed that the velocity $\mathbf{v}^i$ be updated with:

$$\mathbf{v}_{k+1}^{i} = w\mathbf{v}_{v}^{i} + c_1 r_1 \left(\mathbf{p}_k^i - \mathbf{x}_k^i\right) + c_2 r_2 \left(\mathbf{p}_k^g - \mathbf{x}_k^i\right) \tag{D.19}$$

where $\mathbf{p}_k^i$ represent the best ever position of particle i at time k, and $\mathbf{p}_k^g$ represents the global best position in the swarm at time k. The numbers r_1 and r_2 are random uniform numbers between 0 and 1. The inertia term w, is chosen as 1 and then decreased linearly during optimisation. The cognitive and social scaling factors c_1 and c_2 are chosen as $c_1 = c_2 = 2$ in order to allow a mean of unity.

In the case of the training of the dynamic networks, the network biases and weights were taken as the position vector **x**. The algorithm is stopped when it reaches a suitable convergence tolerance on the objective function value. If the algorithm does not reach the tolerance value, the vector **x** is re-initialised. This is necessary for the case when an old tool is replaced with a new tool, because the swarm moved with the wear of the older tool, it will be too far away from the solution to converge within limited steps for the new tool. The number of particles that were used was 25, and the maximum number of steps 40. The random initialisation for bias and weight values was made between -1 and +1. The algorithm usually achieved convergence within less than about 15 inner steps in a matter of seconds on a PIII computer.

E. Custom hardware specifications

E.1 Introduction

As mentioned in the main text of the thesis, a special measurement system was developed in-house for cost-effective strain gauge measurements on the tool holder. The details of the system are described here. The system was developed to measure strain with three strain gauge rosettes, each connected in a half-bridge configuration. Note that 120Ω self-temperature compensated strain gauges should be used. The basic components of the system are:

- HBM clip strain gauge amplifiers type AE101
- HBM clip voltmeter type DA101
- HBM clip power supply type NT101
- variable "dummy" gauges
- low-pass filter with power supply
- user interface for Eagle A/D cards (PC-30U)

All of these components will be described in more detail, besides the particular details of the HBM hardware, which could be sourced from HBM directly. The user interface for the Eagle card is very simple, and only brief details are given. A special connection cable was made for the strain gauge measurement system, and only this cable should be used to connect the custom hardware to Eagle cards. The complete system was built into a robust steel enclosure, depicted in Figure E.1. A picture of the enclosure with the front panel open is shown in Figure E.2.

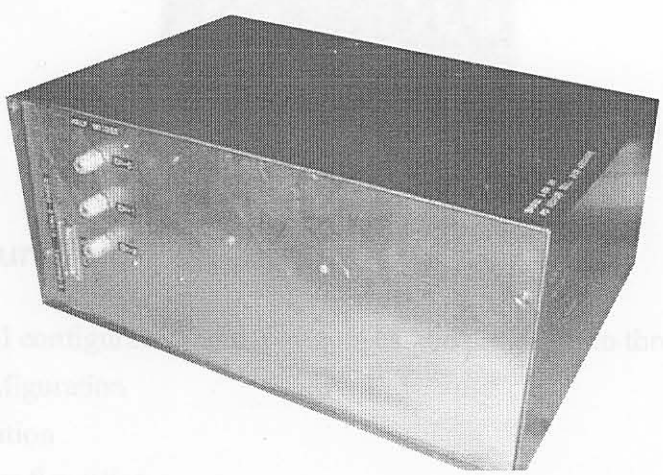


Figure E.1: Strain gauge measurement system

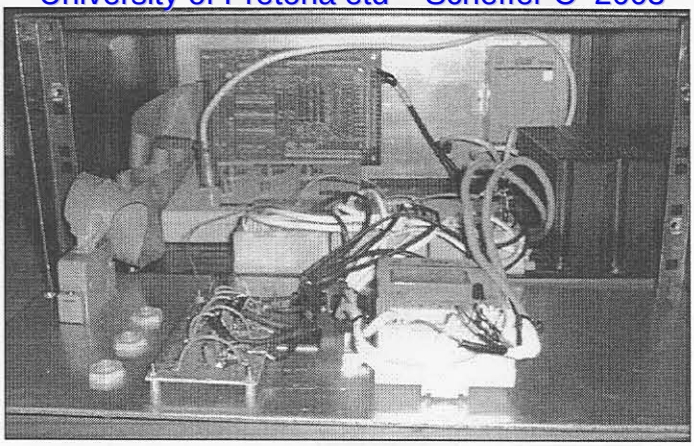


Figure E.2: Enclosure with front panel open

E.2 PC-30U

The PC-30U card is a simple analogue card that protects the PC-30 A/D card against high voltage peaks. A brief description follows:

- External adapter board to be used with any PC-30xx DAQ Card
- High degree of protection from high voltage spikes, accidental shorts and overloads.
- Internal PSU is assumed (thus PC-30xx card must be configured with ±12V and ±5V line enabled).
- Fuses F1, F2, F3 rated @ 125mA.

Only use the supplied cable to connect the strain gauge box to the PC-30 card. Contact Eagle Technologies [265] for additional information. The PC-30U is shown in Figure E.3.

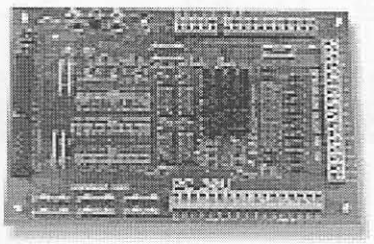


Figure E.3: PC-30U

E.3 Overall configuration

For simplicity, the overall configuration of the system can be divided into three parts:

- power supply configuration
- channel configuration
- cable screening configuration

E.3.1 Power supply configuration

The setup of the power supply is graphically depicted in Figure E.4. One normal South African mains plug is provided for the complete system, and it is shared with the HBM power supply and the power supply for the low-pass filter.

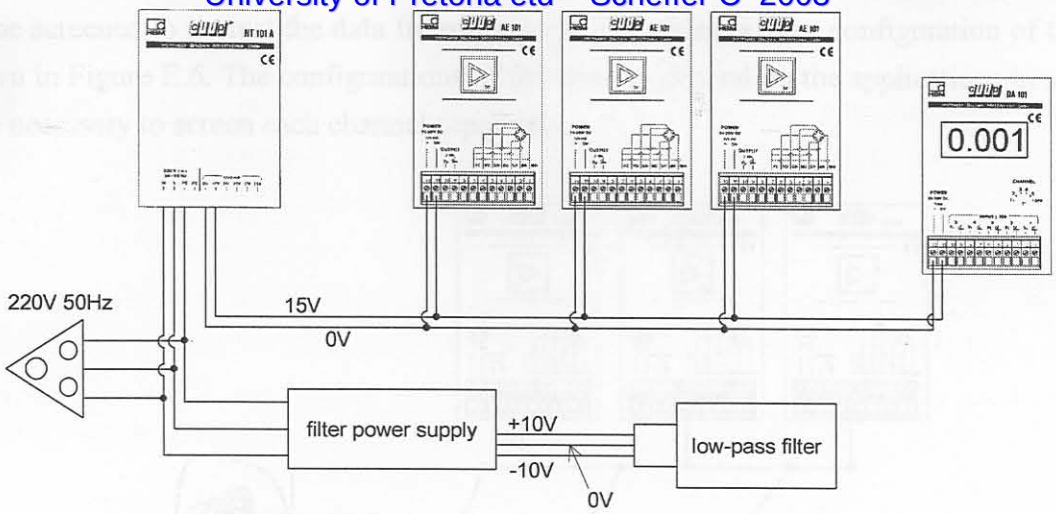


Figure E.4: Power supply configuration

E.3.2 Channel configuration

The three channels are formed from the output of the AE101 amplifiers. From this point, a single analogue ground and three channels are formed. Hence, the system is single-ended and the A/D card should be configured likewise. From the amplifiers the outputs pass through the low-pass filter, and then to the PC-30U. The output cable from the PC-30U is fastened on the front panel of the box and can be connected directly to an Eagle A/D card with the provided cable. The setup of the channels is shown schematically in Figure E.5.

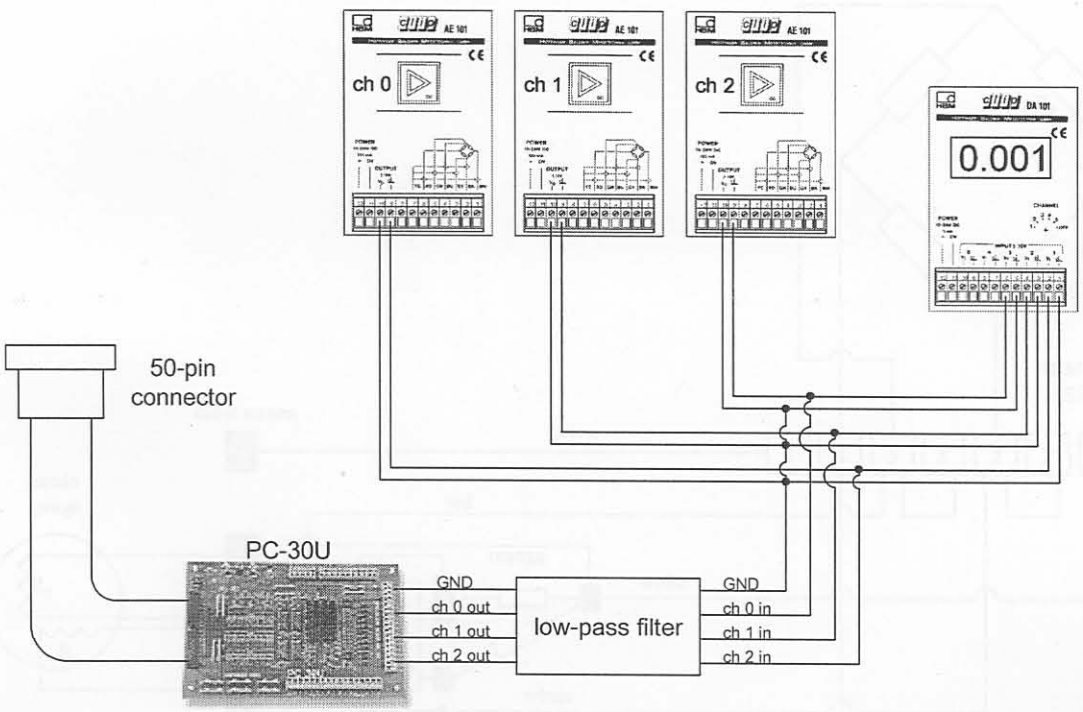


Figure E.5: Channel configuration

E.3.3 Cable screening configuration

All cables carrying very low voltage signals should be screened. The screenings can be connected to the appropriate pins on the AE101 amplifiers. Note that the actual strain gauges and the dummy resis-

tors must be screened to protect the data from electrical disturbances. The configuration of the screening is shown in Figure E.6. The configurations of the screens depend on the application: In some cases it might be necessary to screen each channel separately.

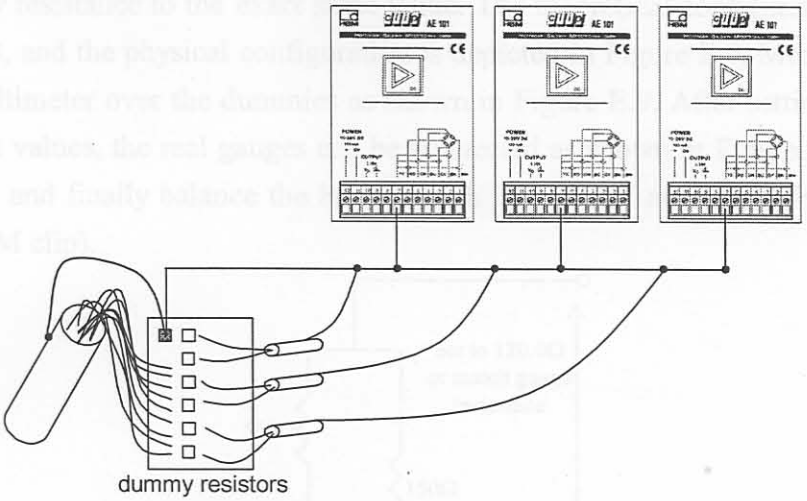


Figure E.6: Cable screening configuration

E.3.4 Per-channel configuration for half-bridges

The per-channel configuration for each half-bridge is graphically depicted in Figure E.7. Each half-bridge within the system is connected in the exact same way. The physical colours of the cables are also indicated on the figure.

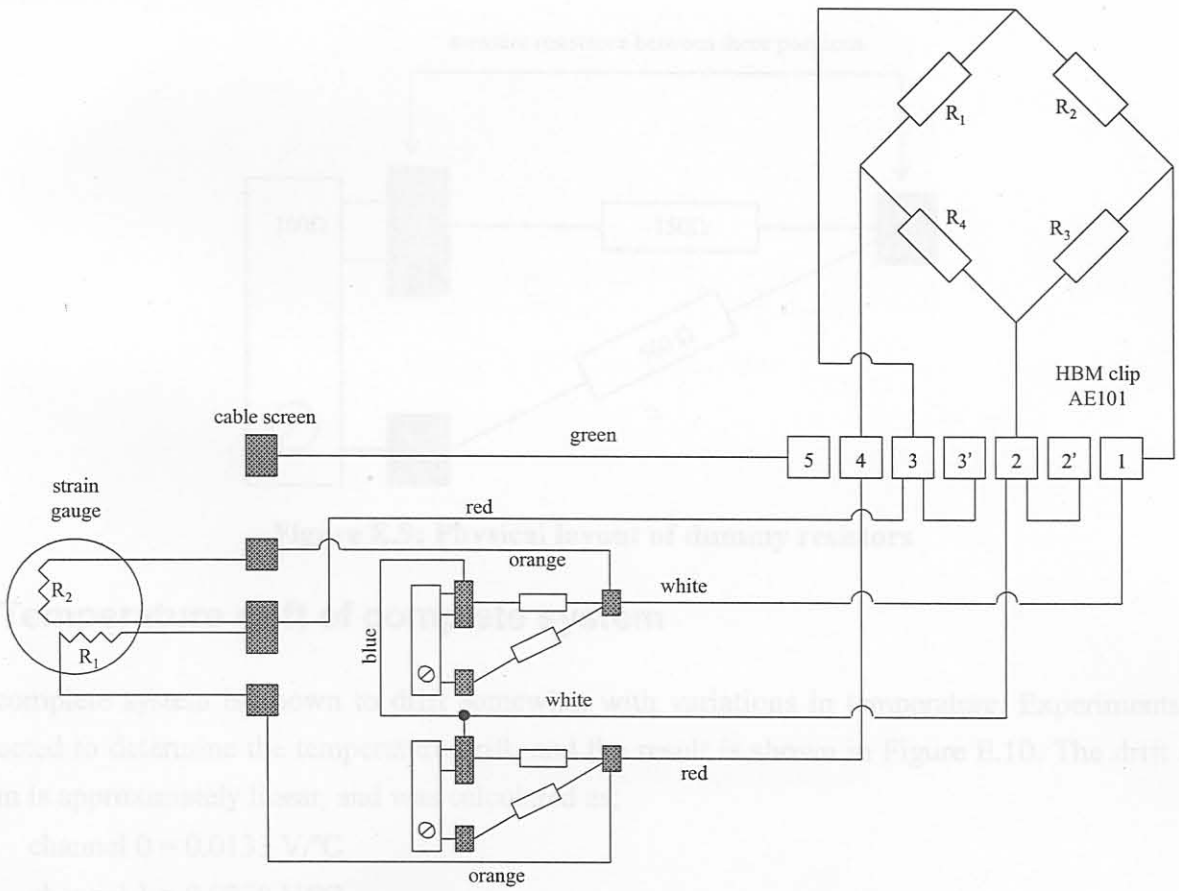


Figure E.7: Detailed per-channel bridge configuration

E.3.5 Dummy gauge configuration

The dummy gauges must be set to the ***exact*** same resistance as the strain gauges. The resistance of the strain gauges should be measured with their cables included. Use the multi-turn potentiometer to change each dummy resistance to the exact same value. The theoretical configuration of the system is shown in Figure E.8, and the physical configuration is depicted in Figure E.9. Measure the dummy resistances with a multimeter over the dummies as shown in Figure E.9. After setting the dummy resistances to the correct values, the real gauges can be connected as shown in Figure E.7. Switch the system main power on and finally balance the bridges with the AE101 amplifiers for each channel (see user manual for HBM clip).

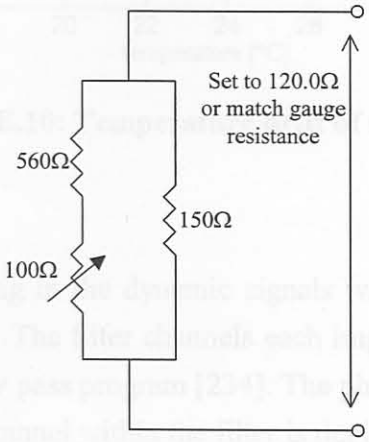


Figure E.8: Theoretical layout of dummy resistance

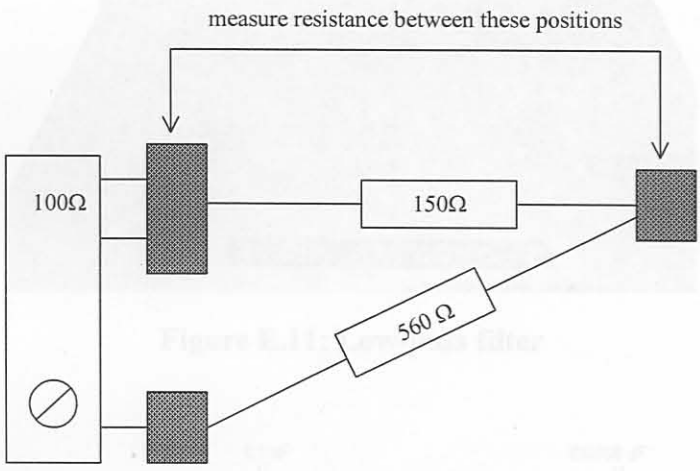


Figure E.9: Physical layout of dummy resistors

E.4 Temperature drift of complete system

The complete system is known to drift somewhat with variations in temperature. Experiments were conducted to determine the temperature drift, and the result is shown in Figure E.10. The drift of the system is approximately linear, and was calculated as:

- channel 0 = 0.0133 V/°C
- channel 1 = 0.0278 V/°C
- channel 2 = 0.0022 V/°C

The effects of temperature are compensated for (refer Chapter 5).

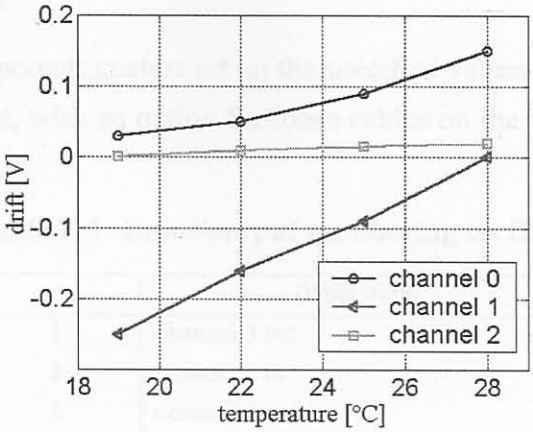


Figure E.10: Temperature drift of system

E.5 Low-pass filter

The low-pass filter to prevent aliasing in the dynamic signals was built in-house, and is part of the complete strain measurement system. The filter channels each implements a 4th order Chebyshev type filter designed with the FilterLab Low pass program [234]. The physical layout of the filter is shown in Figure E.11. The circuit of a single channel within the filter is depicted in Figure E.12:

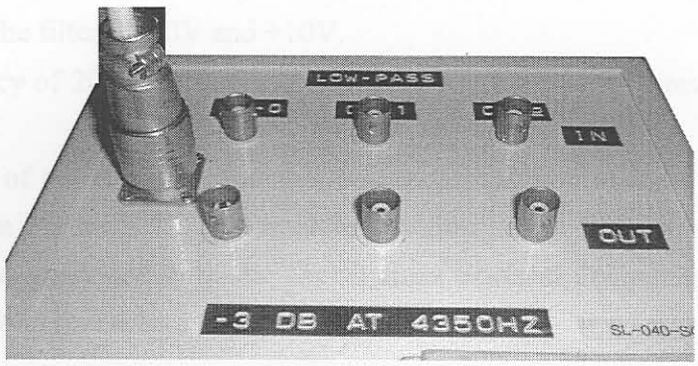


Figure E.11: Low-pass filter

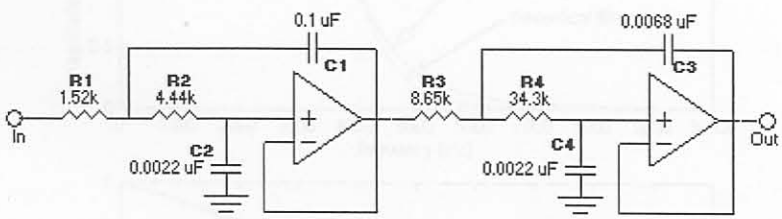


Figure E.12: Filter circuit (one channel)

Herewith some basic details of the filter:

- Three identical channels were built to pass DC - 4000Hz.
- The -3dB point is reached at approx. 4350Hz.

- All components are standard:
  LM324 op-amps
  Standard capacitors
  Resistors are standard potentiometers set on the specified values.
- BNC connectors in and out, with an option for loose cables on the right of the panel. The numbering is as shown in Table E.1:

Table E.1: Input/output numbering on filter

	designation
1	channel 3 out
2	channel 3 in
3	channel 2 out
4	channel 2 in
5	channel 1 out
6	channel 1 in
7	not used
8	analogue ground (0 V)

- A custom-built power supply is supplied with the unit, and is connected with the special 7-pin socket connector to the unit.
- The power supply supplies a 0 V, +12V and -12V DC to the unit from 220V 50 Hz AC power.
- Batteries can also power the unit.
- Maximum range of the filter is -10V and +10V.
- A sampling frequency of 20kHz is suggested when using the filter to prevent aliasing in the DC - 4000 Hz band.

The real characteristics of the channels, together with a comparison to the theoretic characteristic of the circuit, are shown in Figure E.13. The transfer function of the filter was constructed by using a wideband random signal generated with the Siglab from Spectral Dynamics. The transfer was also measured with the Siglab in real time.

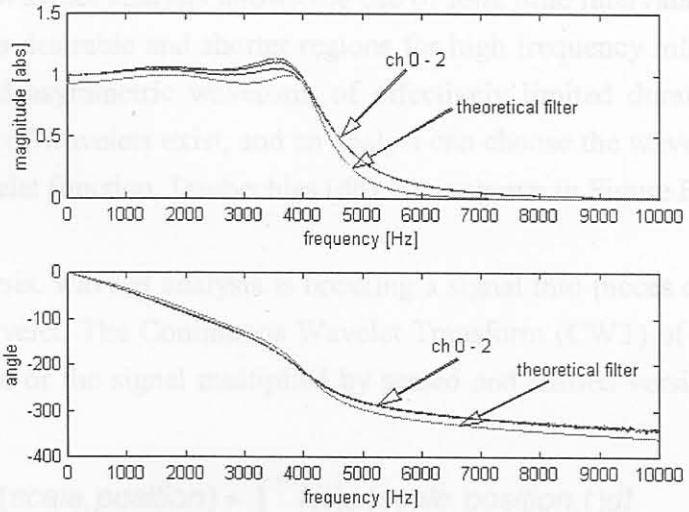


Figure E.13: Filter characteristics

F. Wavelet analysis

F.1 Introduction

The wavelet transform is a relatively new method of signal processing that has been applied to many engineering studies with great success. Recent studies also proved that wavelet analysis could be utilised for monitoring of machining processes [161-163]. The success of the wavelet transform is generally attributed to the natural shape of the wavelet, which is said to be more descriptive of natural phenomena than the sine and cosine functions used in Fourier analysis. Signals with sharp and sudden changes might be better analysed with an irregular wavelet than with a smooth sinusoid. Wavelet analysis is capable of revealing aspects of data that other signal analysis techniques miss, like trends, breakdown points, discontinuities in higher derivatives and self-similarity [266]. In this research, wavelet packet analysis was used to generate features that may show consistent trends towards tool wear. In this appendix, the basic principles of wavelet analysis are discussed.

F.2 Wavelet analysis background

Fourier analysis breaks a signal down into constituent sinusoids of different frequencies, which transforms our view of the signal from the time to the frequency domain. The drawback of Fourier analysis is that the time information is lost, which may be important if the signal contains non-stationary characteristics. This drawback may be overcome by the Short-Time Fourier Transform (STFT), but is of limited precision and not very flexible. These methods are based on windowing the signal and analysing each short time window separately. The signal can then be mapped onto a two-dimensional display of time and frequency [266].

To overcome the limitations of the STFT, wavelet analysis is based on a windowing technique with variable-sized regions. Wavelet analysis allows the use of long time intervals where more precise low frequency information is desirable and shorter regions for high frequency information. A wavelet is a sometimes irregular and asymmetric waveform of effectively limited duration that has an average value of zero. A variety of wavelets exist, and an analyst can choose the wavelet that suits his application best. A typical wavelet function, Daubechies (db) 10, is shown in Figure F.1.

Similar to Fourier analysis, wavelet analysis is breaking a signal into pieces of shifted and scaled versions of the original wavelet. The Continuous Wavelet Transform (CWT) of a function $f(t)$ is defined as the sum over all time of the signal multiplied by scaled and shifted versions of the wavelet function ψ [266]:

$$C(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(scale, position, t)dt \qquad \text{(F.1)}$$

The result of the CWT is many wavelet coefficients C, which are functions of scale and position.

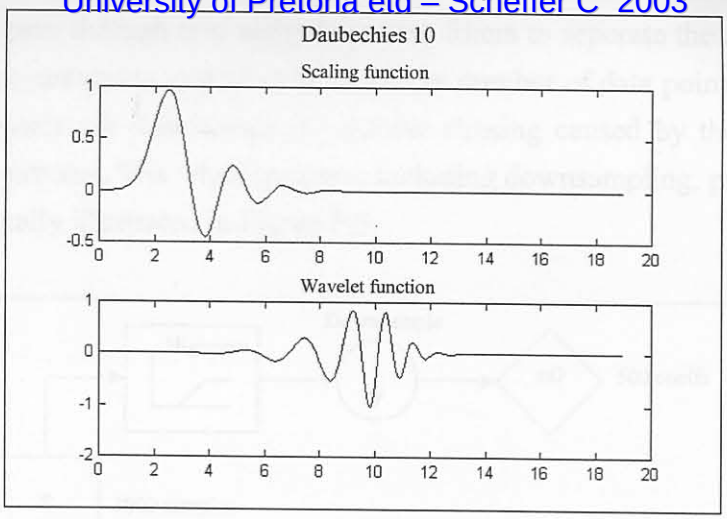


Figure F.1: Typical wavelet function, the db10

F.3 Scaling and shifting

Scaling a wavelet means stretching or compressing it, which is denoted by the scale factor a. The smaller the scale factor, the more "compressed" the wavelet, therefore the scale factor is related to the frequency of the signal. In wavelet analysis the low and high frequency contents of the signal are referred to as Approximations (A) and Details (D), respectively:

- Low scale / Compressed wavelet / High frequency / Details.
- High scale / Stretched wavelet / Low frequency / Approximations.

Shifting a wavelet simply means delaying or hastening its onset. For example, delaying wavelet function $\psi(t)$ by k will be represented by $\psi(t-k)$ [266].

F.4 The Discrete Wavelet Transform (DWT)

If the scales and positions are chosen based on powers of two, the analyses are much more efficient and just as accurate. This is called the Discrete Wavelet Transform (DWT). An efficient way to implement the DWT is by using filters. The filtering process is actually very complex, and will only be discussed in principle. At its most basic level, the filtering process can be illustrated as shown in Figure F.2.

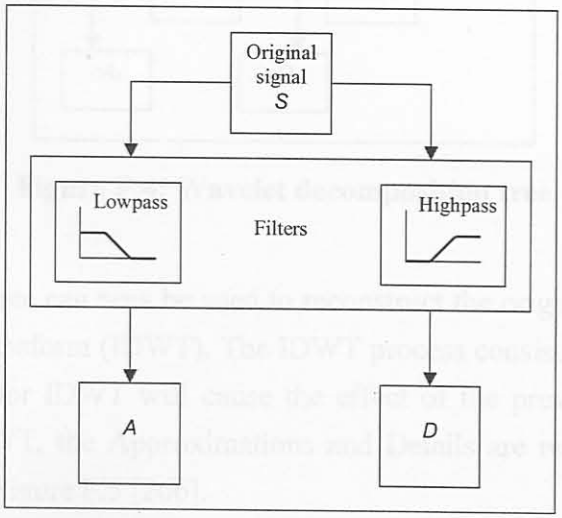


Figure F.2: DWT at its most basic level, using filters

The original signal, *S*, pass through two complementary filters to separate the approximations and the details. This causes computation to end up with twice the number of data points of the original signal. To correct this, the signals are downsampled, and the aliasing caused by the downsampling is accounted for later in the process. The whole process, including downsampling, produces the DWT coefficients as diagrammatically illustrated in Figure F.3.

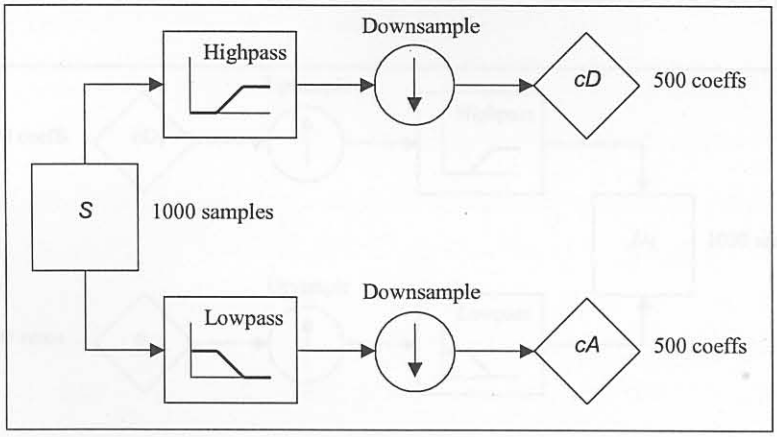


Figure F.3: DWT with downsampling

F.5 Multiple Level Decomposition and Reconstruction

The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower-resolution components. This is called the wavelet decomposition tree, which in theory can be calculated to an infinite level, but in practice can only continue until the details consist of a single data point. In most cases, the optimal level of decomposition can be calculated based on an energy approach, such as entropy. The wavelet decomposition tree, as shown in Figure F.4, illustrates the multiple-level decomposition diagrammatically [266].

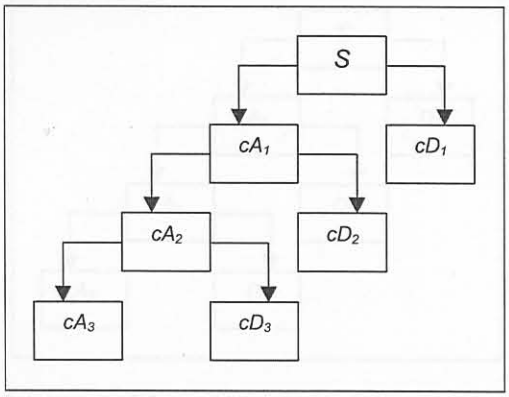


Figure F.4: Wavelet decomposition tree

The wavelet decomposition tree can now be used to reconstruct the original signal, which is called the Inverse Discrete Wavelet Transform (IDWT). The IDWT process consists of upsampling and filtering. A specific choice of filters for IDWT will cause the effect of the previously mentioned aliasing to 'cancel out'. During the IDWT, the Approximations and Details are reconstructed with the wavelet coefficients, as illustrated in Figure F.5 [266].

The original signal can now be reconstructed with a number of combinations from the Approximations and Details on the decomposition tree, for example:

$$\begin{aligned} S &= A_1 + D_1 \\ &= A_2 + D_2 + D_1 \\ &= A_3 + D_3 + D_2 + D_1 \end{aligned} \tag{F.2}$$

The wavelet tree with reconstructed components is illustrated in Figure F.6 [266].

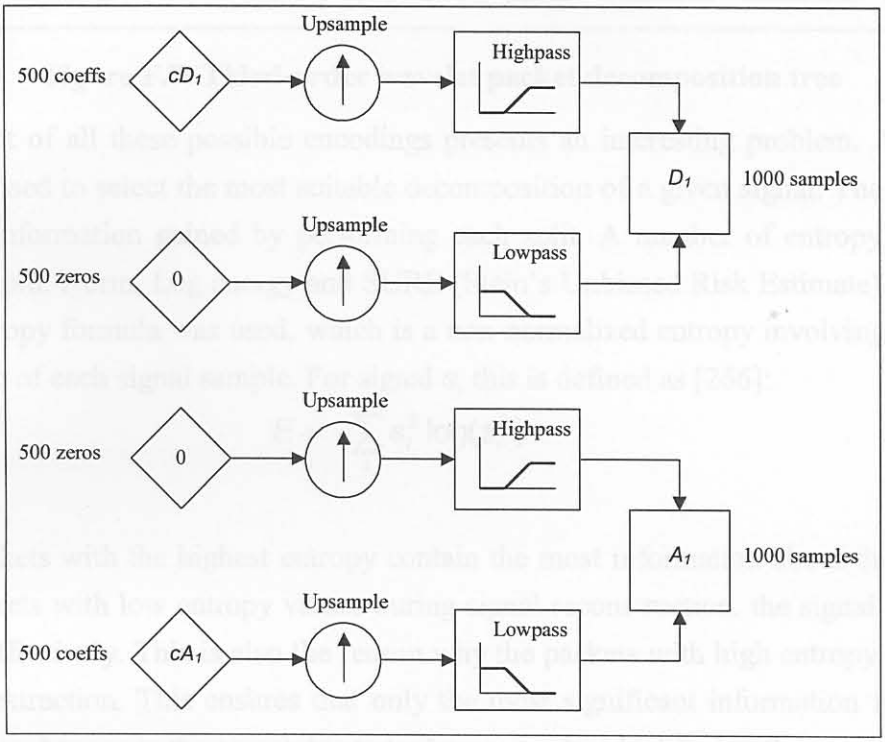


Figure F.5: Reconstruction of Approximations and Details

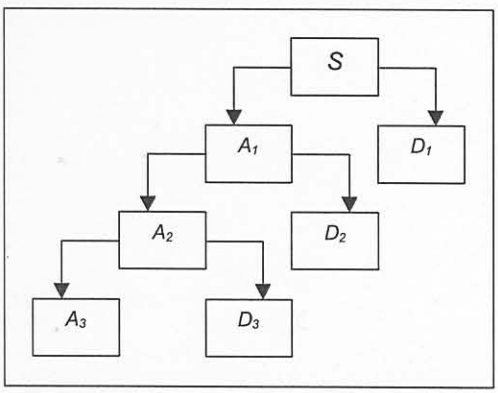


Figure F.6: Wavelet tree with reconstructed Approximations and Details

F.6 Wavelet Packet Analysis

The wavelet packet method offers a wider range of possibilities for signal analysis. With normal wavelet analysis, only the approximations are split in every step. In wavelet packet analysis, the details as well as the approximations are split in every step, illustrated in Figure F.7 [266]. The signal can be reconstructed using any number of packets on the wavelet packet decomposition tree, for example:

$$S = A_1 + AAD_3 + DAD_3 + DD_2 \tag{F.3}$$

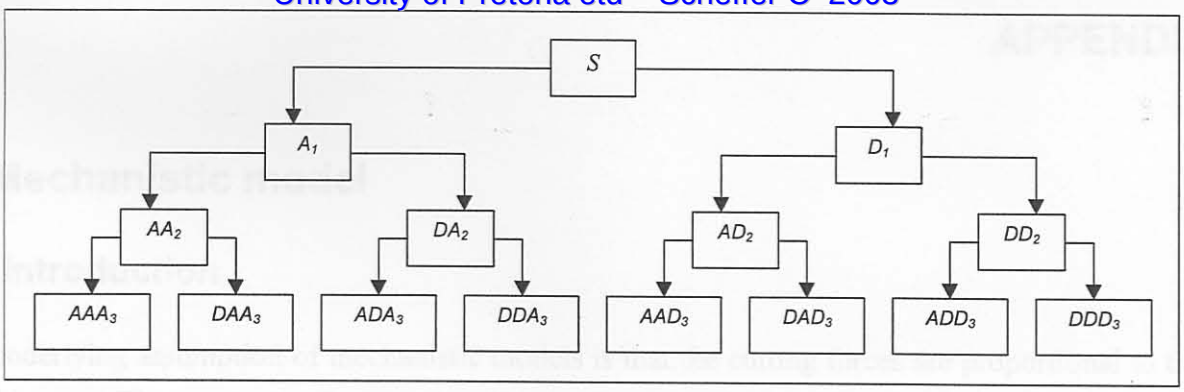


Figure F.7: Third-order wavelet packet decomposition tree

Choosing one out of all these possible encodings presents an interesting problem. An entropy-based criterion can be used to select the most suitable decomposition of a given signal. The entropy is an indication of the information gained by performing each split. A number of entropy types exist, like Shannon, Threshold, Norm, Log energy and SURE (Stein's Unbiased Risk Estimate). In this research, the Shannon entropy formula was used, which is a non-normalized entropy involving the logarithm of the squared value of each signal sample. For signal s, this is defined as [266]:

$$E = -\sum_i s_i^2 \log(s_i^2) \tag{F.4}$$

The wavelet packets with the highest entropy contain the most information about the process. By neglecting the packets with low entropy values during signal reconstruction, the signal can be de-noised or compressed effectively. This is also the reason why the packets with high entropy values were chosen for feature extraction. This ensures that only the most significant information from the signal is extracted, and any changes in the signal due to tool wear can be identified with ease.

G. Mechanistic model

G.1 Introduction

The underlying assumption of mechanistic models is that the cutting forces are proportional to the uncut chip area. The constants of proportionality are determined by the cutting geometry, material properties and machining conditions. The procedure for calculating the worn tool cutting forces is based on calculating the sharp tool forces and adding the forces caused by the worn tool. The complete model is built around a combination of models for estimating several unknowns, and also relies on a calibration procedure via selected experiments. The references for this entire appendix are the research papers by DeVor, Kapoor and their colleagues at UIUC [45,223,242,243,267,268].

G.2 Procedure for modelling the sharp tool forces in turning

G.2.1 Basic mechanistic relationship

The magnitudes of the normal and friction forces are modelled as being proportional to the process geometry characteristics that represent the amount of material being removed. These relations are:

$$F_n = K_n \cdot A_c \tag{G.1}$$

$$F_f = K_f \cdot A_c \tag{G.2}$$

where:

F_n is the normal cutting force

F_f is the frictional cutting force

K_n is the normal cutting pressure

K_f is the frictional cutting pressure

A_c is the area of chip load

G.2.2 Force transformation relations

For turning, an equivalent oblique cutting geometry can be obtained by taking the nose radius of the tool into account. The associated geometrical parameters can be obtained by integrating along the curved cutting edge. The local tool cutting forces can be determined by:

$$\begin{bmatrix} F_c \\ F_t \\ F_a \end{bmatrix} = [A] \cdot F_n + [B] \cdot F_f \tag{G.3}$$

where:

$$[A]=\begin{bmatrix}\cos(\alpha_{en})\cos(i_e)\\ \cos(\alpha_{en})\sin(i_e)\sin(\gamma_{eL})-\sin(\alpha_{en})\cos(\gamma_{eL})\\ -\cos(\alpha_{en})\sin(i_e)\cos(\gamma_{eL})-\sin(\alpha_{en})\sin(\gamma_{eL})\end{bmatrix} \quad (G.4)$$

$$[B]=\begin{bmatrix}\sin(\eta_c)\sin(i_e)+\cos(\eta_c)\sin(\alpha_{en})\sin(i_e)\\ \cos(\eta_c)\cos(\alpha_{en})\cos(\gamma_{eL})-[\sin(\eta_c)\cos(i_e)-\cos(\eta_c)\sin(\alpha_{en})\sin(i_e)]\sin(\gamma_{eL})\\ [\sin(\eta_c)\cos(i_e)-\cos(\eta_c)\sin(\alpha_{en})\sin(i_e)]\cos(\gamma_{eL})+\cos(\eta_c)\cos(\alpha_{en})\sin(\gamma_{eL})\end{bmatrix} \quad (G.5)$$

and:

γ_{eL} is the effective lead angle

i_e is the effective inclination angle

α_{en} is the effective normal rake angle

η_c is the chip-flow angle

For turning operations, these force components coincide with the external coordinate system:

$$\begin{bmatrix}F_x\\ F_y\\ F_z\end{bmatrix}=\begin{bmatrix}F_c\\ F_t\\ F_a\end{bmatrix} \quad (G.6)$$

G.2.3 Chip load model

To calculate the area of chip load, a numerical integration along the edge of the cutting tool intersecting with the workpiece is performed. The uncut chip thickness is also calculated during this procedure.

G.2.4 Chip flow model

The direction of chip flow is computed by dividing the round cutting edge into a large number of smaller straight edges and assuming oblique cutting on each (Stabler rule). The unified chip flow direction is then computed as the weighted average of all the individual chip flow directions, where the weighting for each is determined by the magnitude of the friction force at that edge. This process is done iteratively by initially guessing the friction force at each edge. The unified chip flow direction, measured with respect to a line on the rake face normal to the equivalent cutting edge is given by:

$$\eta_c=\frac{\sum\left|F_f(w)\cdot(\eta_c(w)+\gamma'(w))\right|}{\sum F_f(w)}-\gamma'_{eL} \quad (G.7)$$

G.2.5 Model calibration

Model calibration is achieved by conducting experiments over a range of machining conditions. Empirical equations are used to capture the effects of machining parameters on the normal and friction pressures:

$$\ln(K_n)=a_0+a_1\ln(t_c)+a_2\ln(V)+a_3\alpha_n+a_4\ln(V)\ln(t_c) \quad (G.8)$$

$$\ln(K_f)=b_0+b_1\ln(t_c)+b_2\ln(V)+b_3\alpha_n+b_4\ln(V)\ln(t_c) \quad (G.9)$$

where:

t_c is the uncut chip thickness

V is the cutting speed

α_n is the normal rake angle

G.3 Basic procedure for modelling the worn tool forces

G.3.1 Introduction

The basic assumption of this model is that for three-dimensional cutting operations, the plastic flow region grows linearly as the total wearland width of the tool increases, and this growth can be modelled independent of the cutting conditions. The width of the plastic flow region is of importance because the tool flank stresses (and consequently the cutting forces) depend on it. An integral part of determining the tool flank stresses is the development of a contact model for determining the normal stress at the tool tip. The three-dimensional cutting forces due to tool flank wear can then be determined by discretisising the rounded flank face into two-dimensional elements and adding the forces from all the elements.

G.3.2 Linear growth of plastic flow region

It has been shown that the plastic flow region on worn tools increase linearly with respect to the length of cut for a range of materials. Based on this, a piecewise linear function is proposed to model the growth of tool flank wear:

$$VB_p = c_0 + c_1 VB \text{ if } VB < VB_{cr} \text{ then } VB_p = 0 \tag{G.10}$$

where:

$$VB_{cr} = \frac{-c_0}{c_1} \tag{G.11}$$

VB is the total wearland width

VB_p is the width of the plastic flow region

c_0, c_1 are linear coefficients that can be obtained by fitting the function through experimental data. It has been shown that this model is independent of cutting conditions which means that only one wear experiment is required to obtain a model for the growth of the plastic flow region.

G.3.3 Two-dimensional stress on the tool flank

The cutting force and thrust force due to tool flank wear can be determined by integrating the normal and shearing tool flank stresses over the flank wear width:

$$F_{tw} = w \int_0^{VB} \sigma_w(x)\mathrm{d}x \tag{G.12}$$

$$F_{cw} = w \int_0^{VB} \tau_w(x)\mathrm{d}x \tag{G.13}$$

where:

F_{tw} is the thrust force due to tool flank wear

F_{cw} is the cutting force due to tool flank wear

σ_w is the normal tool flank stress

τ_w is the shearing tool flank stress

w is the width of cut

x is the distance from tool tip

In the plastic flow region at the tool tip, the stresses are equal to the tool tip stresses, σ_o and τ_o. By applying the linear growth of the plastic flow region, the following functions can be derived:

$$\text{if } x < VB_p \begin{cases} \sigma_w = \sigma_o \\ \tau_w = \tau_o \end{cases} \text{ Plastic flow} \tag{G.14}$$

$$\text{if } x > VB_p \begin{cases} \sigma_w = \sigma_o \cdot \left(\dfrac{VB - x}{VB - VB_p} \right)^2 \\ \tau_w = \tau_o \cdot \left(\dfrac{VB - x}{VB - VB_p} \right)^2 \end{cases} \text{ Elastic contact} \tag{G.15}$$

where:

σ_o is the tool tip normal stress

τ_o is the tool tip shear stress

G.3.4 Contact model to predict the normal tool tip stress

Since plastic deformation of the workpiece occurs under the flank of the tool, the maximum effective stress in the workpiece must be equal to the yield strength of the workpiece material. It is therefore assumed that the stresses in the tool flank are limited by the yield stress of the workpiece material. The contact model presented here calculates the maximum effective stress in the workpiece to be equal to the yield strength of the workpiece material. A plane strain assumption is followed because it has been shown that the tangential stress does not play an important role. The contact model is only applied once to a workpiece material, and the value of σ_o is assumed to remain constant for all values of VB.

The stresses in the workpiece at an arbitrary point A due to the load can be found by integrating a set of concentrated loads over the loaded region. The following equations can be used to obtain the effective stress in the workpiece:

$$\sigma_x = -\frac{2z}{\pi} \int_{-a}^{a} \frac{\sigma_w(s)(x-s)^2 \, ds}{\left[(x-s)^2 + z^2\right]^2} \tag{G.16}$$

$$\sigma_z = -\frac{2z^3}{\pi} \int_{-a}^{a} \frac{\sigma_w(s) ds}{\left[(x-s)^2 + z^2\right]^2} \tag{G.17}$$

$$\tau_{xz} = -\frac{2z^2}{\pi} \int_{-a}^{a} \frac{\sigma_w(s)(x-s) ds}{\left[(x-s)^2 + z^2\right]^2} \tag{G.18}$$

where:

$$\sigma_w(s) = \sigma_o \left(\frac{s-a}{2a} \right)^2 \tag{G.19}$$

and s, x and z are defined in the Figure G.1.

To obtain the effective stress σ_{eff} the Von Mises criterion can be used:

$$\sigma_{eff} = \sqrt{(\sigma_1 - \sigma_2)^2 + \sigma_1^2 + \sigma_2^2} \tag{G.20}$$

where:

$$\sigma_{1,2} = \frac{\sigma_x + \sigma_z}{2} \pm \sqrt{\left(\frac{\sigma_x - \sigma_z}{2} \right)^2 + \tau_{xz}^2} \tag{G.21}$$

To find the maximum effective stress in the workpiece, the workpiece is divided into small square elements. The effective stress is calculated at all four corners of the squares and the maximum effective stress is the maximum of all these calculated values. To find the normal stress that causes the maximum effective stress to be equal to the workpiece yield strength, an iterative approach is used. A value for normal stress is chosen and the maximum effective stress is calculated. When the maximum effective stress is equal to the yield strength, the correct value for the normal stress is found.

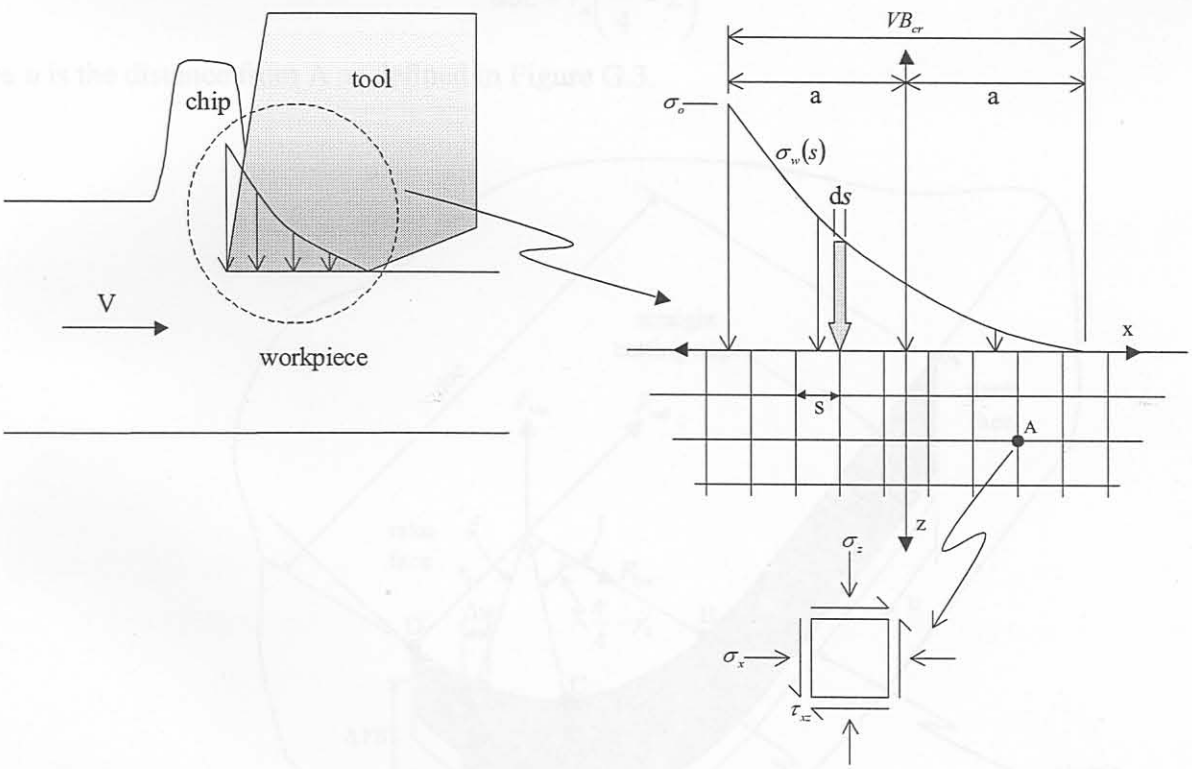


Figure G.1: Stress distribution on workpiece

G.3.5 Prediction of shearing tool tip stress

The shearing tool tip stress is found by assuming that the friction on the flank of the tool is adhesive in nature. Thus, the calculated value for the normal stress can be used to determine the shearing tool tip stress, using:

$$\tau_o = \frac{\sigma_o}{1+2\pi} \tag{G.22}$$

G.3.6 Calculating the three-dimensional forces

Once the tool tip stresses are known, the tool flank stresses can be found as described earlier. The three-dimensional forces are found by dividing the tool flank into small two-dimensional elements. For the straight edge section, the width of each two-dimensional element is found by choosing the number of elements, N :

$$\Delta w = \frac{doc - r_n}{N\cos(\gamma_1)} \tag{G.23}$$

where:

Δw is the width of the element

doc is the depth of cut

r_n is the tool nose radius

γ_1 is the lead angle of the tool

The individual element's VB values are given by:

$$\Delta VB = \frac{(u + 0.5\Delta w)(VB - 0.5 \cdot VB)}{doc + r_n\left(\frac{\pi}{4} - 2\right)} \tag{G.24}$$

where u is the distance from A as defined in Figure G.3.

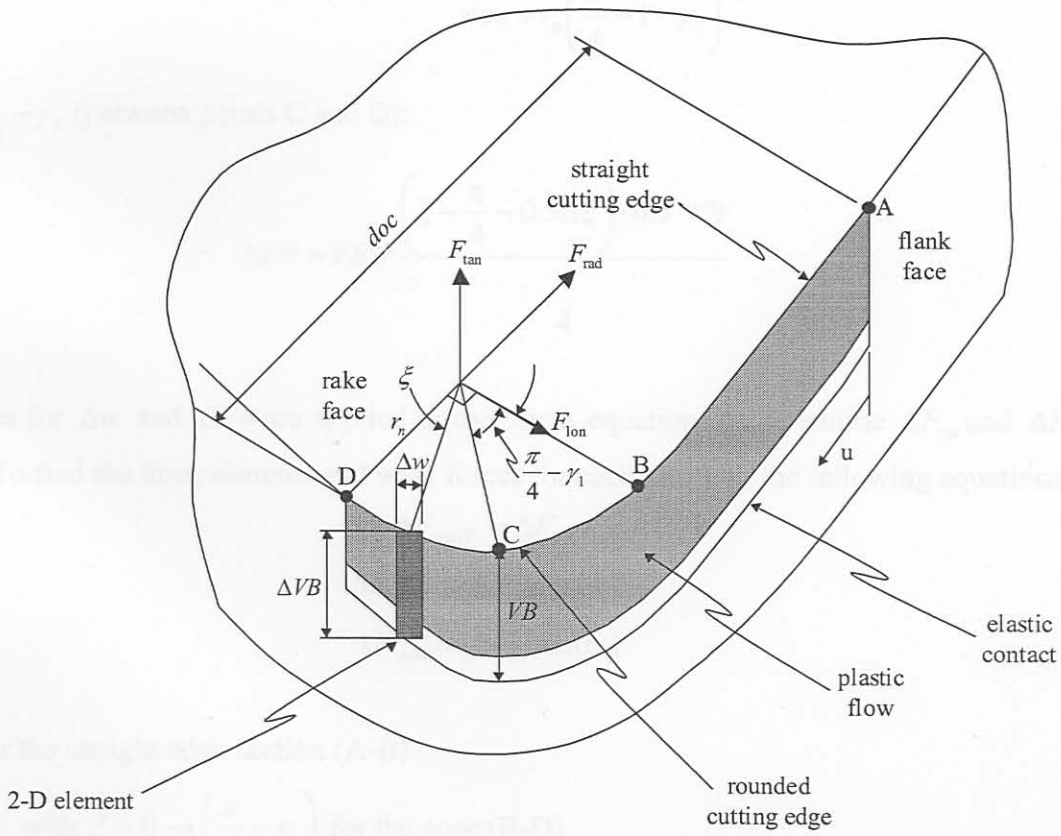


Figure G.2: Analysis of tool (1)

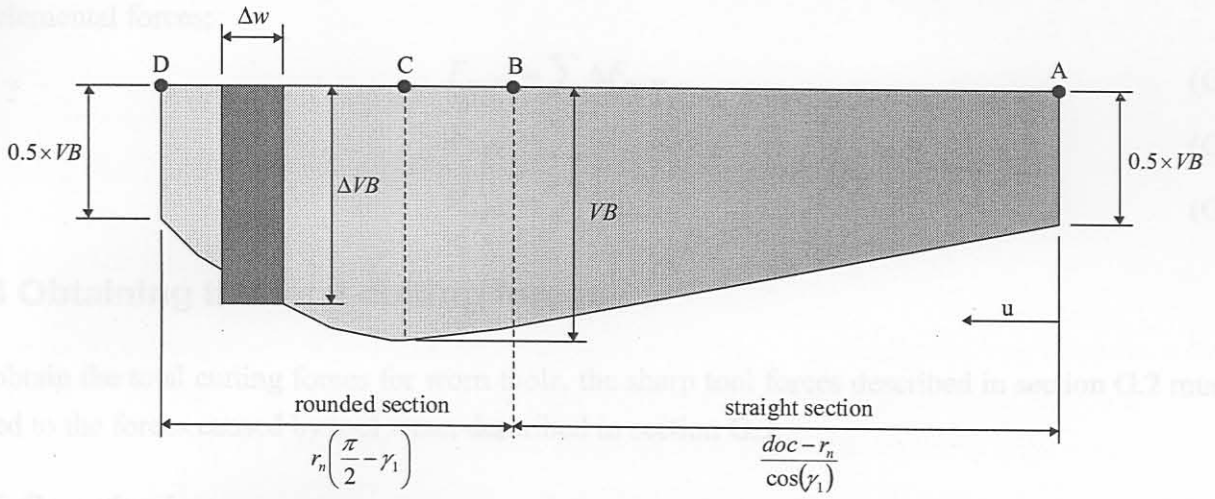


Figure G.3: Analysis of tool (2)

For the rounded nose section, Δw can be found by dividing the total included angle on the nose of the tool into $\Delta\xi$ sections:

$$\Delta\xi = \frac{\frac{\pi}{2}-\gamma_1}{N}, \Delta w = 2r_n \sin\left(\frac{\Delta\xi}{2}\right) \tag{G.25}$$

For $\xi \le \frac{\pi}{4}-\gamma_1$ (between points B and C):

$$\Delta VB = 0.5\cdot VB + \frac{doc + r_n\left(\xi - 0.5\Delta\xi\right)\cdot 0.5\cdot VB}{doc + r_n\left(\frac{\pi}{4}-1-\gamma_1\right)} \tag{G.26}$$

For $\xi > \frac{\pi}{4}-\gamma_1$ (between points C and D):

$$\Delta VB = VB - \frac{\left(\xi - \frac{\pi}{4} - 0.5\Delta\xi\right)\cdot 0.5\cdot VB}{\frac{\pi}{4}} \tag{G.27}$$

The values for Δw and ΔVB are applied in the stress equations to determine ΔF_{cw} and ΔF_{tw} for each element. To find the three dimensional wear forces for each element, the following equations apply:

$$\Delta F_{\tan W} = \Delta F_{cw} \tag{G.28}$$

$$\Delta F_{lonW} = \Delta F_{tw}\cos(\psi) \tag{G.29}$$

$$\Delta F_{radW} = \Delta F_{tw}\sin(\psi) \tag{G.30}$$

where:

$\psi = \gamma_1$ for the straight edge section (A-B)

$\psi = \xi + \gamma_1$ with $\xi = 0 \rightarrow \left(\frac{\pi}{2}-\gamma_1\right)$ for the nose (B-D)

Once these forces are known, the total cutting forces due to tool wear can be calculated by adding all

the elemental forces:

$$F_{\tan W} = \sum \Delta F_{\tan W} \tag{G.31}$$

$$F_{lonW} = \sum \Delta F_{lonW} \tag{G.32}$$

$$F_{radW} = \sum \Delta F_{radW} \tag{G.33}$$

G.4 Obtaining the total cutting forces

To obtain the total cutting forces for worn tools, the sharp tool forces described in section G.2 must be added to the forces caused by tool wear, described in section G.3.

G.5 Conclusion

Despite the fact that the model described in this appendix seems to be the most novel approach up to date for calculating worn tool forces using basic theoretical principles and combining experimental data, the implementation of the model proved unsuccessful due to reasons described in Chapter 6.

H. The Self-Organising Map (SOM)

H.1 Introduction

The Self-Organising Map (SOM), developed by Kohonen [212], is a fairly new and effective software tool for data analysis. The SOM has been implemented successfully in numerous applications in fields such as process analysis, machine perception, control and communication. The SOM implements the orderly mapping of high-dimensional data onto a two-dimensional grid of neurons. Thereby the SOM is able to identify hidden relationships between high-dimensional data with simple geometric relationships that can be displayed on a single figure [269]. The SOM can be described as a NN with self-organising capabilities. Most NNs require information and interaction from the user for classification. The training of the SOM is based on unsupervised learning, which means that the data is automatically arranged without output patterns. Although the SOM was originally intended as a data visualisation tool, it can be used for data classification as well. Cho [270] describes how the SOM can be used as a high-performance classifier for noisy data sets. The SOM automatically arranges the data on a two dimensional grid of neurons where similar observations are placed close to one another and dissimilar ones further away. If the classes of the training observations are known, certain regions on the grid could be allocated for these classes.

H.2 Computation of the SOM

H.2.1 Structure of the SOM

A SOM is formed of neurons on a 2-dimensional grid. Higher dimensional grids can also be used, but their visualisation is very problematic. Each neuron i of the SOM is represented by an n-dimensional weight or reference vector $\boldsymbol{m}_i = [m_{i1}\ m_{i2}\ \ldots\ m_i]$, where n is equal to the dimension of the input vectors. The map shape is usually rectangular, but other shapes have also been used successfully. The number of neurons is fixed before training. The number of neurons affects the accuracy and the generalisation capability of the SOM. As the size of the map increases, training becomes very time consuming.

H.2.2 Neighbourhood relation

The neurons on the map are connected to adjacent neurons by a neighbourhood relationship. Immediate neighbours belong to the 1-neighbourhood N_{i1} of the neuron i. In the 2-dimensional case the neurons of the map can be arranged either on a rectangular or a hexagonal lattice. Neighbourhoods of different sizes in rectangular and hexagonal lattices are illustrated in Figure H.1 [269].

H.2.3 Initialisation

Before the training phase initial values are given to the weight vectors. The SOM is robust with respect to initialisation, but a proper initialisation allows the algorithm to converge faster to a reliable solution.

One of the three following initialisation procedures can be used [269]:

- Random initialisation, whereby the initial values of the weight vectors are selected randomly.
- Sample initialisation, whereby the initial values of the weight vectors are selected based on samples from the training data.
- Linear initialisation, whereby the initial values of the weight vectors are generated linearly from the lowest to the highest value of the training data.

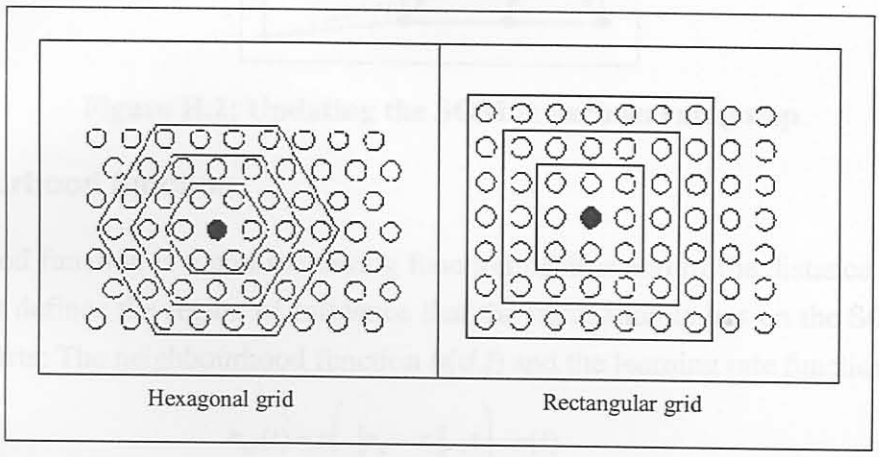


Figure H.1: Neighbourhood structures

H.2.4 Training the SOM

The computation of the SOM is a nonparametric, recursive regression process. In each training step, one sample vector $\boldsymbol{x}$ from the input data set is chosen randomly and a similarity measure is calculated between it and all the weight vectors of the map. The Best-Matching Unit (BMU), denoted as c, is the unit whose weight vector has the greatest similarity with the input sample $\boldsymbol{x}_i$. The similarity is usually defined by means of a distance measure, typically the Euclidian distance. Formally the BMU is defined as the neuron for which [221]:

$$\forall i,\ \|\boldsymbol{x}(t)-\boldsymbol{m}_c(t)\| \le \|\boldsymbol{x}(t)-\boldsymbol{m}_i(t)\| \tag{H.1}$$

which means that $\boldsymbol{m}_c(t)$ is the model that matches best with $\boldsymbol{x}(t)$. This is the BMU.

After finding the BMU, the weight vectors of the SOM are updated. The weight vectors of the BMU and its neighbours are moved closer to the input vector in the input space. This adaptation procedure stretches the BMU and its neighbours towards the sample vector. This is illustrated in Figure H.2, where the input vector given to the network is marked by an x [269]. The SOM update rule for the weight vector of the unit i is:

$$\boldsymbol{m}_i(t+1) = \boldsymbol{m}_i(t) + h_{c(x),i}(\boldsymbol{x}(t)-\boldsymbol{m}_i(t)) \tag{H.2}$$

where t is the index of the learning step, and learning is performed recursively for each presentation of a sample of $\boldsymbol{x}$, denoted $\boldsymbol{x}(t)$. The scalar multiplier $h_{c(x),i}$ is called the neighbourhood function, which causes similar observations to be placed in the same region on the map.

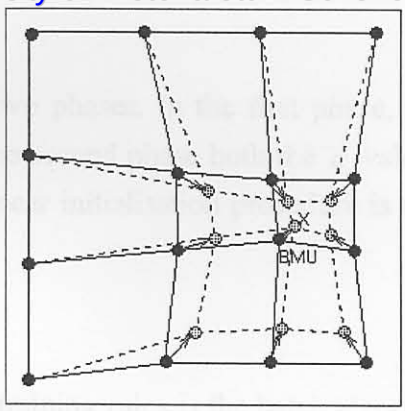


Figure H.2: Updating the SOM in each learning step

H.2.5 Neighbourhood function

The neighbourhood function is a non-increasing function of time and of the distance of unit *i* from the winning unit *c*. It defines the region of influence that the input sample has on the SOM. The function consists of two parts: The neighbourhood function $h(d,t)$ and the learning rate function $\alpha(t)$:

$$h_{ci}(t) = h\left(\|\mathbf{r}_c - \mathbf{r}_i\| , t \right) \alpha(t) \tag{H.3}$$

where $\mathbf{r}_i$ is the location of unit *i* on the map grid.

The simplest neighbourhood function is the Bubble: It is constant over the whole neighbourhood of the winner unit and zero elsewhere. Another is the Gaussian neighbourhood function. It gives slightly better results, but is computationally slower. Figure H.3 displays the two common neighbourhood functions [269].

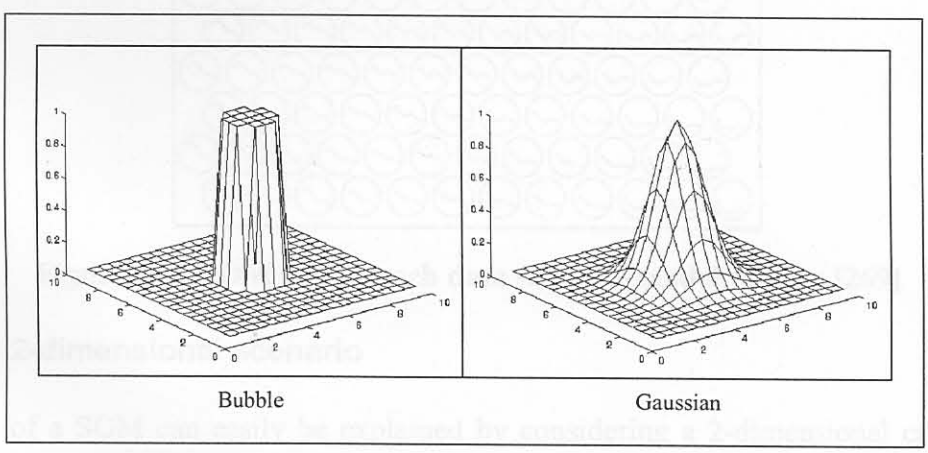


Figure H.3: Two common neighbourhood functions

H.2.6 Learning rate

The learning rate $\alpha(t)$ is a decreasing function of time. Two commonly used forms are a linear function and a function inversely proportional to time, such as:

$$\alpha(t) = \frac{A}{(t+B)} \tag{H.4}$$

where *A* and *B* are selected constants.

H.2.7 Training phases

Training is usually performed in two phases. In the first phase, relatively large initial α values and neighbourhood radii are used. In the second phase both the α values and the neighbourhood radii are small from the beginning. If the linear initialisation procedure is used, the first training phase can be skipped.

H.2.8 Batch-algorithm

Another variant of the basic SOM training rules is the batch algorithm. In this case the whole training set is used at once and after which the map is updated. The algorithm usually converges after a couple of iterations, and is much faster to calculate in MATLAB than the normal sequential algorithm.

H.3 Examples of SOMs

The SOM is best explained with simple examples. Two examples are discussed here in order to clarify the use of SOMs. For more information on SOMs, see references [221,198,199,269-274].

H.3.1 Speech data

The SOM in Figure H.4 was calculated with speech data (Finnish) [269]. The model vectors are shown on each neuron. Note that similar patterns are arranged close to one another on the map. This technique can be used for voice or word recognition.

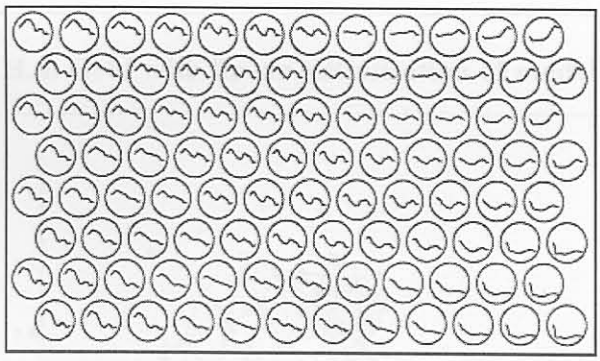


Figure H.4: SOM with speech data showing model vectors [269]

H.3.2 Simple 2-dimensional scenario

The principles of a SOM can easily be explained by considering a 2-dimensional case. Say, for instance, that a data set exists which clearly displays two clusters, as shown in Figure H.5. The one cluster may correspond to an error (*err*) situation, and the other to an acceptable (*OK*) region, for example:

$$\begin{aligned} W &= [x\ y] \\ W &= [0.3\ \ 0.3] - err \\ W &= [0.7\ \ 0.7] - OK \end{aligned} \tag{H.5}$$

Figure H.6 shows the SOM after linear initialisation of the model vectors, on a rectangular grid. Figure H.7 shows the SOM after training, which resembles a 'net' folding over the 'cloud' of data. The neurons move closer to one another where the data is dense.

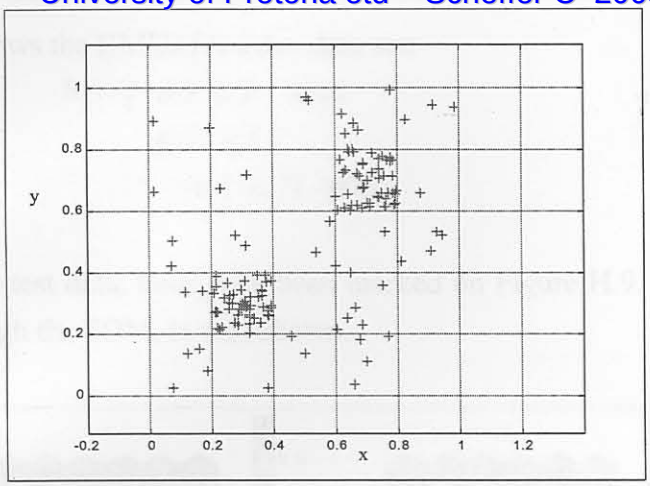


Figure H.5: Data with two clusters

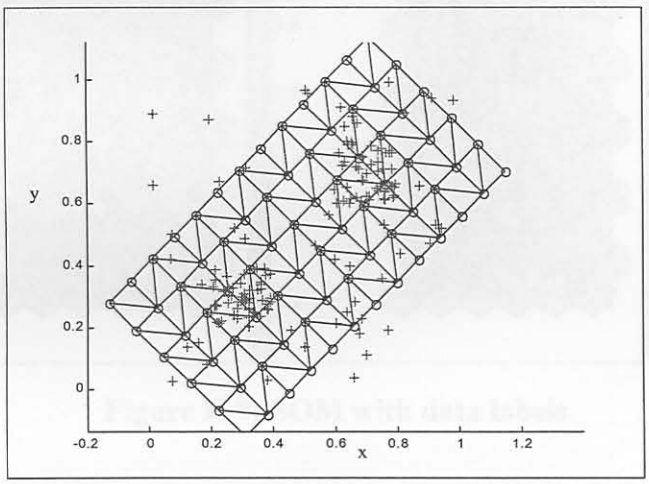


Figure H.6: Grid with linear initialisation of model vectors

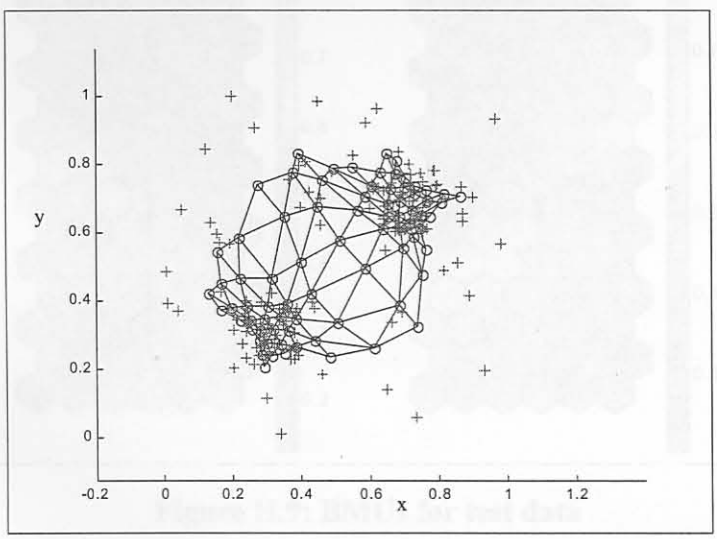


Figure H.7: Grid after training

Figure H.8 shows the SOM for variables *x* and *y*. Note that the SOM is only one entity, but a graph in the direction of each variable (dimension) can be shown. Also note that the colours of the neurons are an indication of the value of the variable, as shown on the colour map axis on the figure. The labels of the training data are also shown in Figure H.8, and it is clear that the *'err'* and *'OK'* regions have been

identified. Figure H.9 shows the BMUs for a test data set:

$$W = [\ 0.3\ \ 0.3\ \ (err)$$
$$0.5\ \ 0.5 \qquad (H.6)$$
$$0.7\ \ 0.7]\ \ (OK)$$

The three BMUs for this test data, have each been marked on Figure H.9. A trajectory, showing how the test data moves through the SOM, is also shown.

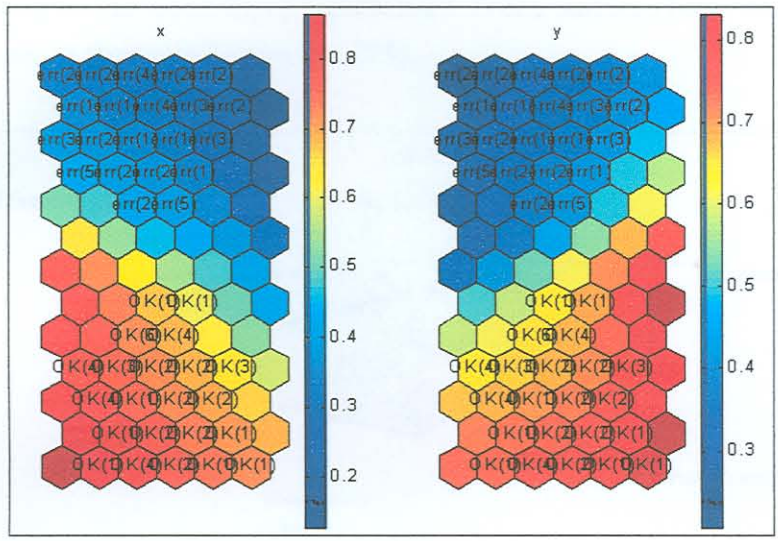


Figure H.8: SOM with data labels

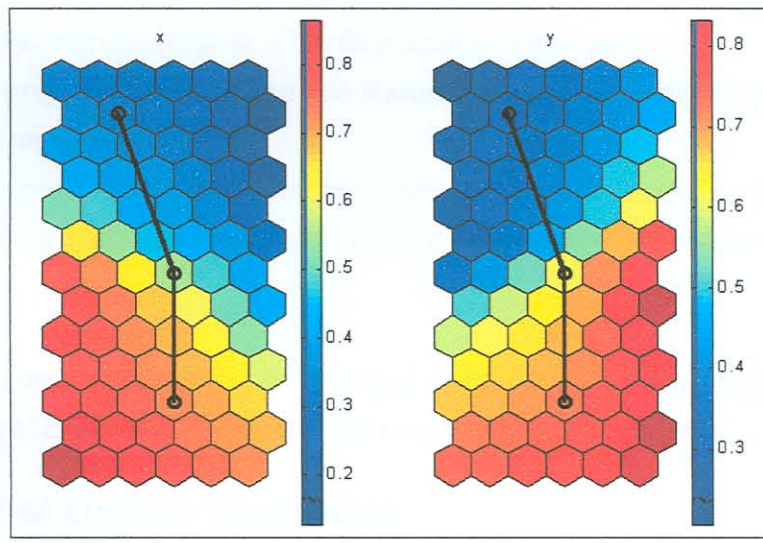


Figure H.9: BMUs for test data

I. Surface roughness analysis

I.1 Surface texture

A surface that is nominally smooth and flat will always reveal some roughness, which may vary from fine to coarse depending on the finishing operation used. Some surfaces have roughness and waviness, and may also be curved, as shown in Figure I.1 [275].

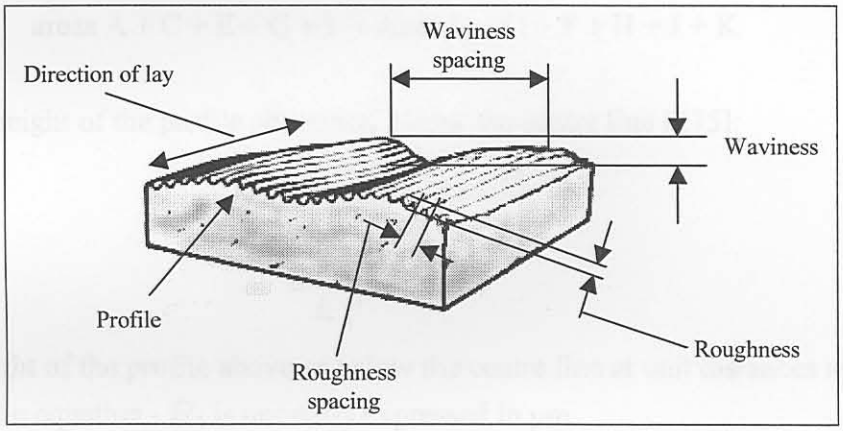


Figure I.1: Roughness and Waviness

Figure I.2 illustrates the two components of surface texture, which are:

- Roughness: The irregularities in the surface texture that is inherent to the production process, excluding waviness and errors of form.
- Waviness: This is the component of surface texture upon which roughness is superimposed. Waviness may result from such factors such as machine or workpiece deflections, vibrations, chatter or heat treatment.

Each pattern is characterised by the lay (the principal direction of the predominant surface pattern), the spacing of the principal crests, and in height (with respect to a reference line).

I.2 Assessment of surface roughness

The standard method for assessing surface texture is based on traversing a stylus across the surface to produce an electrical signal, which can generate the surface profile on a chart or an average reading on a meter. Roughness average, R_a, is defined as the arithmetical average of the profile above and below the reference line throughout the prescribed sampling length. This is illustrated in Figure I.2. Surface roughness values are normally assessed as mean results of several sampling lengths taken consecutively along the surface.

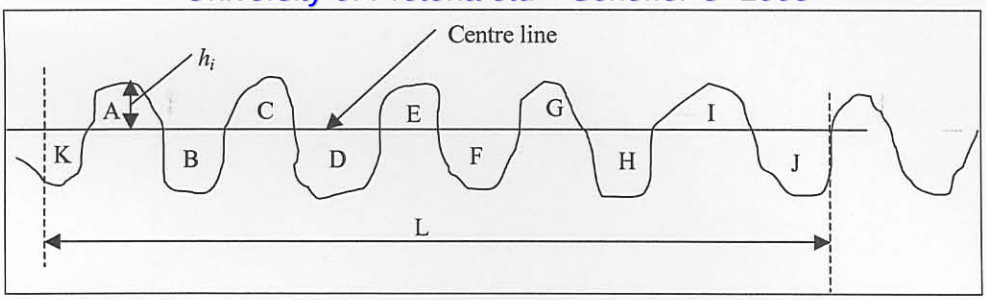


Figure I.2: Definition of the centre line

Over a length of surface L, the centre line is a line such that the sum of the areas embraced by the surface profile above the line is equal to the sum of those below the line, thus:

$$\text{areas } A + C + E + G + I = \text{areas } B + D + F + H + J + K \tag{I.1}$$

R_a is the average height of the profile above and below the centre line [275]:

$$R_a = \frac{h_1 + h_2 + h_3 + \ldots h_n}{L} = \frac{1}{L}\int_0^L |h|\, dL \tag{I.2}$$

where h is the height of the profile above or below the centre line at unit distances apart. The units of L is not added into the equation - R_a is normally expressed in μm.

R_z is the rms value for R_a:

$$R_z = \sqrt{\frac{1}{L}\int_0^L h^2(x)\,dx} \tag{I.3}$$

These are the basic parameters used for surface roughness assessment. The same type of parameters can also be derived to quantify the waviness of the surface profile.