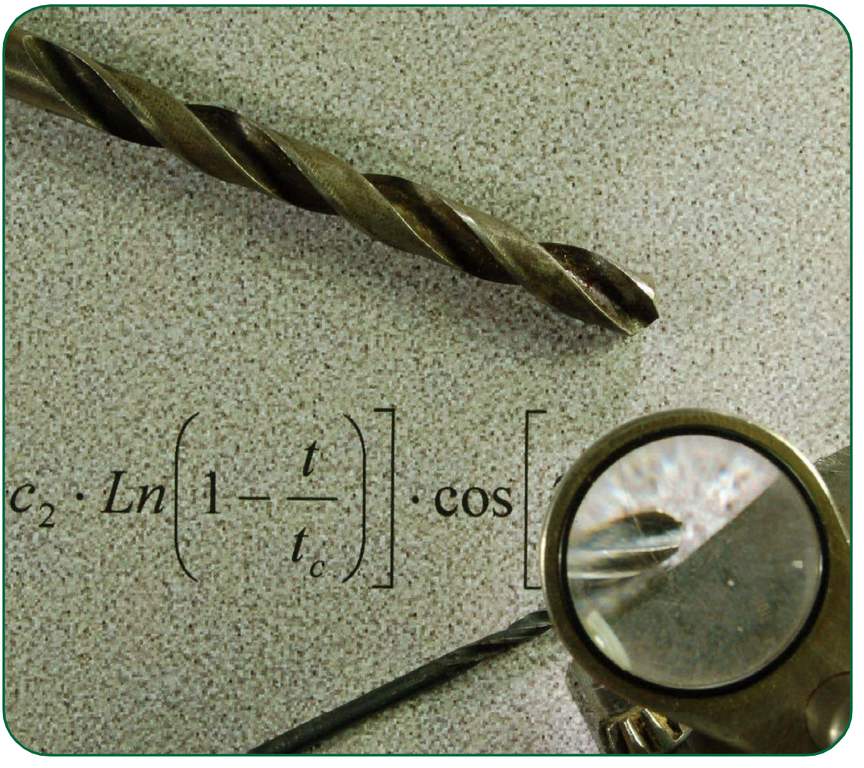
ESPOO 2005 VTT PUBLICATIONS 590



Erkki Jantunen

Indirect multisignal monitoring and diagnosis of drill wear

VTT PUBLICATIONS 590

Indirect multisignal monitoring and diagnosis of drill wear

Erkki Jantunen
VTT Industrial Systems

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Mechanical Engineering for public examination and debate in Auditorium 216 (Otakaari 4) at Helsinki University of Technology (Espoo, Finland) on the 20th of January, 2006, at 12 noon.

ISBN 951–38–6692–0 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 951–38–6693–9 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 3, PL 1000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 4374

VTT, Bergsmansvägen 3, PB 1000, 02044 VTT
tel. växel 020 722 111, fax 020 722 4374

VTT Technical Research Centre of Finland, Vuorimiehentie 3, P.O.Box 1000, FI–02044 VTT, Finland
phone internat. +358 20 722 111, fax +358 20 722 4374

VTT Tuotteet ja tuotanto, Metallimiehenkuja 6, PL 1702, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 7077

VTT Industriella system, Metallmansgränden 6, PB 1702, 02044 VTT
tel. växel 020 722 111, fax 020 722 7077

VTT Industrial Systems, Metallimiehenkuja 6, P.O.Box 1702, FI–02044 VTT, Finland
phone internat. +358 20 722 111, fax +358 20 722 7077

Technical editing Leena Ukskoski

Otamedia Oy, Espoo 2005

Jantunen, Erkki. Indirect multisignal monitoring and diagnosis of drill wear. Espoo 2005, Technical Research Centre of Finland, VTT Publications 590. 80 p. + app. 110 p.

Keywords drill wear, condition monitoring, signal analysis, polynomial regression analysis, fuzzy logic, diagnosis

Abstract

A machine tool utilisation rate can be improved by an advanced condition monitoring system using modern sensor and signal processing techniques. A drilling test and analysis program for indirect tool wear measurement forms the basis of this thesis. For monitoring the drill wear a number of monitoring methods such as vibration, acoustic emission, sound, spindle power and axial force were tested. The signals were analysed in the time domain using statistical methods such as root mean square (rms) value and maximum. The signals were further analysed using Fast Fourier Transform (FFT) to determine their frequency contents. The effectiveness of the best sensors and analysis methods for predicting the remaining lifetime of a tool in use has been defined. The results show that vibration, sound and acoustic emission measurements are more reliable for tool wear monitoring than the most commonly used measurements of power consumption, current and force. The relationships between analysed signals and tool wear form a basis for the diagnosis system. Higher order polynomial regression functions with a limited number of terms have been developed and used to mimic drill wear development and monitoring parameters that follow this trend. Regression analysis solves the problem of how to save measuring data for a number of tools so as to follow the trend of the measuring signal; it also makes it possible to give a prognosis of the remaining lifetime of the drill. A simplified dynamic model has been developed to gain a better understanding of why certain monitoring methods work better than others. The simulation model also serves the testing of the developed automatic diagnostic method, which is based on the use of simplified fuzzy logic. The simplified fuzzy approach makes it possible to combine a number of measuring parameters and thus improves the reliability of diagnosis. In order to facilitate the handling of varying drilling conditions and work piece materials, the use of neural networks has been introduced in the developed approach. The scientific contribution of the thesis can be summarised as the development of an automatically adaptive diagnostic tool for drill wear detection. The new

approach is based on the use of simplified fuzzy logic and higher order polynomial regression analysis, and it relies on monitoring methods that have been tested in this thesis. The diagnosis program does not require a lot of memory or processing power and consequently is capable of handling a great number of tools in a machining centre.

Preface

This research was carried out at the Technical Research Centre of Finland (VTT) during 1994–2005. Publications I and III are linked to the Predictive Intelligent Machine and Machining Monitoring Sensors (Pimms), BRST-CT98-5429 project. Publications II and V are part of the FMS Maint System, EUREKA MAINE EU 744 project. Publications IV, VI and VII relate to Multiple Intelligent Diagnostics for Machinery (MindMan), and Publication VIII to On-Line Multi-Sensor Diagnostic Analysis for Maintenance using Neural Networks (Neural Maine), EUREKA Project EU 1250. These projects have been funded by the EU, the National Technology Agency of Finland (Tekes), Finnish industry and the Technical Research Centre of Finland. The financial support is gratefully acknowledged.

Professor Mauri Airila, whom I would like to thank for his advice and support, supervised this thesis. I would also like to thank Professor Kenneth Holmberg for his advice and encouragement.

I would like to express my thanks to the pre-examiners Professor Tony Hope and Professor Thomas N. Moore for their kind comments.

A great number of researchers, whose names are acknowledged in the papers that form part of this thesis, were involved in the above projects and influenced this work, for which I wish to express my gratitude. Special thanks are due to Mr. Harri Jokinen for the discussions and planning, Mr. Antti Poikonen for installing the measuring system and programming the data acquisition program, Mrs. Natalia Siren for programming the fuzzy logic and regression analysis functions, and Mrs. Adelaide Lönnberg for checking the English language of the thesis.

Finally I would like to thank my wife Sinikka and our children Taru, Arto and Eero for their patience and support.

December 8th, 2005

Erkki Jantunen

Contents

List of symbols

Letters

$c_{1\ldots10}$	constants
a	regression coefficient
b	regression coefficient
b_l	defines the relation between the increasing part and the stable part of the thrust force when drilling one hole
c	damping, regression coefficient
ClassMean	mean value of class
d	constant
e	exponent in regression analysis
E	Young's modulus
f	exponent in regression analysis, feed
f_0	natural frequency of the drill
f_n	natural frequency for bending
F_0	force at the natural vibration frequency of the drill
F_{dh}	shape function that takes into account the unstable features of drilling a hole
F_{dp}	drilling process force
ϕ_{ge}	angular geometrical error due to tolerance in manufacturing
F_{nrpm}	force that takes into account the harmonic components of the drilling speed
F_{rnd}	random force (noise)
$F_{rpm1,rpm2}$	forces that take into account geometrical differences in cutting lips
ϕ_{wd}	difference in wear of the two cutting lips of the drill
F_x	horizontal drill force
g	exponent in regression analysis
h	influence of wear
H_B	Brinell hardness
HighHigh	higher limit of class
HighLow	high limit after which the class descends from unity
i	class index, counter for hole number, index in regression analysis
I	moment of inertia
j	size of class
k	shape of class, stiffness
K_n	coefficient that depends on the vibration mode
l	length of the drill

LowHigh	lower limit where class reaches unity
LowLow	lower limit of class
m	mass
μ	mean value
n	number of samples, order of harmonic component
p	factor used for emphasising the most recent data in regression analysis
q	constant that defines how much weight old data is given in regression analysis
ρ	density
σ	standard deviation
S	cross-sectional area
t	time
t_c	total lifetime of the drill
t_d	time it takes to drill one hole
ω	angular speed of rotation
x	coordinate axis
y	monitored parameter

Acronyms

AD	analogue to digital conversion
AE	acoustic emission
ARMA	autoregressive moving average
ART	adaptive resonance theory
FEM	finite element method
FFT	fast Fourier transform
FLVQ	fuzzy learning vector quantization
HMM	hidden Markov model
HOS	higher order spectrum
HSS	high speed steel
LVQ	learning vector quantization
MARSE	measured area under the rectified signal envelope
PC	personal computer
PSD	power spectral density
RAMV	ratio of the absolute mean value
RCE	restricted Coulomb energy
rms	root mean square

List of publications

This dissertation consists of a summary and eight appended publications I-VIII.

Publication I	Jantunen, E. 2002. A Summary of Methods Applied to Tool Condition Monitoring in Drilling. International Journal of Machine Tools and Manufacture, Vol. 42, pp. 997–1010. ISSN 0890-6955
Publication II	Jantunen, E. & Jokinen, H. 1996. Automated On-Line Diagnosis of Cutting Tool Condition. International Journal of Flexible Automation and Integrated Manufacturing, 4 (3 & 4), pp. 273–287. ISSN 1064-6345
Publication III	Jantunen, E. 2001. The Applicability of Various Indirect Monitoring Methods to Tool Condition Monitoring in Drilling. International Journal of Comadem, Vol. 7, No. 3, July 2004, pp. 24–31. ISSN 1363-7681 (also published in COMADEM 01. September 4–6, Manchester, UK, ISBN 0 08 0440363)
Publication IV	Jantunen, E. 2004. Dynamic Effects Influencing Drill Wear Monitoring. Proceedings of the MFPT 58th Meeting, Ed. H.C. Pusey, S.C. Pusey & W.R. Hobbs, April 25–30, Virginia Beach, USA. Pp. 51–60.
Publication V	Jantunen, E., Jokinen, H. & Milne, R. 1996. Flexible Expert System for Automated On-Line Diagnosis of Tool Condition. Proceedings of a Joint Conference, Technology Showcase, Integrated Monitoring, Diagnostics & Failure Prevention, MFPT 50th Meeting, Joint Oil Analysis Program Technical Support Center, University of Wales, Ed. H.C. Pusey & S.C. Pusey, Mobile, Alabama, USA, April 22–26. Pp. 259–268.

Publication VI Jantunen, E. 2003. Prognosis of Wear Progress Based on Regression Analysis of Condition Monitoring Parameters. Finnish Journal of Tribology, Vol. 22/2003, 4, pp. 3–15. ISSN 0780-2285 (also published in COMADEM 03 August 27–29, Växjö, Sweden. ISBN 91-7636-376-7)

Publication VII Jantunen, E. 2006. Diagnosis of Tool Wear Based on Regression Analysis and Fuzzy Logic. IMA Journal of Management Mathematics, Vol. 17, No 1, January, pp. 47–60. ISSN 1471-6798

Publication VIII Jantunen, E. 2000 Flexible Hierarchical Neuro-Fuzzy System for Prognosis. Proceedings of COMADEM 2000, 13th International Congress on Condition Monitoring and Diagnostic Engineering Management. Ed. H.C. Pusey & Raj B.K.N. Rao, December 3–8, Houston, USA. Pp. 699–708. ISBN 0-9635450-2-7

Author’s contribution

The author was responsible for the monitoring methods, signal analysis and simulation program approach in publication II. In publication V the author was responsible for the fault tree and symptom tree database definition and the definition of data acquisition, signal analysis, regression analysis and simulation module.

1. Introduction

1.1 Background and motivation

Tool wear and failure monitoring has aroused interest among many researchers and research organisations. The background and motivation for this interest is that tool condition monitoring is considered important for the following reasons:

- Cost effective unmanned production is only possible in practise if there is a reliable method available for tool wear monitoring and breakage detection. For example, based on a recent study it has been claimed that in machining centres tool maintenance and tool monitoring cause most of the stoppages during unmanned operation [Kuhmonen 1997].
- Tool wear influences the quality of the surface finish and the dimensions of the parts manufactured. The quality of the surface finish and the dimensions are linked to the above mentioned unmanned operation, i.e. if this is not monitored or the quantity of tool wear is not monitored, the unmanned machining might lead to poor quality.
- The economical tool life cannot be fully benefited from without efficient methods for tool wear monitoring because of the variation in tool life. This factor is not economically as important as the above two during drilling as far as the cost of tools is considered, but nevertheless economically meaningful when the costs of production are studied in detail.
- Where sudden tool failures are to be avoided, tool changes need to be made based on conservative estimates of tool life. This does not take into account sudden failures and at the same time leads to an unnecessarily high number of tool changes, because the full tool life is not benefited from and valuable production time is therefore lost.

1.2 Research question

In order to overcome the challenges described in the previous chapter, condition monitoring and diagnosis of tool wear is needed. This then leads to the research question: *How can the wear of the cutting tools of a machine tool be monitored* *and diagnosed in a practical and reliable manner?* Tool wear monitoring is difficult because so many factors affect the signals collected, i.e. tool type, cutting depth, cutting speed, feed rate and work piece material. Also in a cutting process many factors can cause distortion in the measured signals, e.g. cutting fluid, changes in the environment, chip formation which is a very dynamic process, and the material and geometry which are not necessarily homogeneous. In addition to the technical boundary conditions described above, the developed solution has to be easy and fast to configure for different environments, since otherwise it would not be used. The solution also needs to rely only on a limited number of transducers of an acceptable price level, so that the solution can be economically extremely well justified. If it is not clear that it will save money, industry will not make the investment. Also, in the end a diagnostic system has to be so easy to use that no special skills are required for taking it into use and interpreting the results.

1.3 Objectives of the research

The main goal of this thesis is to develop tools for practical monitoring and diagnosis of drill wear. For this purpose a number of sub-goals have to be fulfilled. It is necessary to discover which indirect monitoring methods are best for drill wear. It is also necessary to identify which signal analysis techniques work best for this purpose. For practical reasons the diagnosis has to be made automatic, which leads to the use of artificial intelligence and search for a suitable approach. In addition the diagnosis has to be reliable, i.e. the use of a number of signals is tested in order to be able to handle noise in the measurement signals. For practical reasons the automatic diagnosis approach has to be easy to configure in various environments. Due to the large number of tools in an industrial environment, there is a need to develop an approach for handling the great amount of data collected. A method for handling the varying process conditions also needs to be developed. Finally the goal is to be able to predict or make a prognosis of the remaining life time of the drill in order to enable uninterrupted unmanned use of machining tools.

1.4 Contents of the thesis

The thesis is divided into seven further chapters as follows:

Chapter 2 reviews the current state-of-the-art study of drill wear monitoring and diagnosis. Commonly used indirect monitoring methods are described. The most common signal analysis techniques are presented. Following these the diagnosis methods commonly used for drill wear monitoring are discussed.

Chapter 3 describes the test and measuring arrangement together with the test program.

Chapter 4 summarises the results of laboratory tests done with various measuring methods and signal analysis techniques. This chapter also attempts to explain why some measuring signals are better than others and, similarly, why some analysis techniques work better than others.

Chapter 5 presents an extremely simplified dynamic model of the drill and the drilling forces and especially how wear influences these forces. The model is used for artificially producing vibration data. The model provides further understanding about the reasons why certain measuring signals together with certain analysis techniques work better than other methods.

Chapter 6 discusses two possible approaches to automating the diagnosis of drill wear by flexible expert systems. Methods of automatic adjustment of the diagnosis of the tool condition are given special emphasis, as well as how the reliability of the diagnosis can be improved by combining a number of analysed parameters. Also are covered the practical aspects of data management in an industrial environment.

Chapter 7 discusses the findings of the thesis in different areas, i.e. the measuring methods, signal analysis techniques and the diagnosis based on artificial intelligence techniques.

Chapter 8 concludes this thesis and provides some suggestions for future research.

1.5 Scope of the research

The thesis covers all commonly used indirect monitoring methods such as drilling force and vibration, and tries to provide an understanding of which of these methods work best in drill wear monitoring. The direct tool wear monitoring methods that measure tool wear as such are not studied here. The reason is that although many attempts have been made to develop such monitoring methods, they still seem to be too complicated and costly for practical purposes. Similarly the work covers commonly used signal analysis techniques in condition monitoring, and tries to establish their suitability for drill wear monitoring. Neither new measuring nor signal analysis techniques are developed. However, problems related to the noise of measuring signals and the influence of cutting parameters are given a lot of consideration. Also a lot of emphasis is given to the consideration of how the drill wear monitoring and diagnosis can be made easy or automatic in practice even though there are so many factors that influence the monitoring, i.e. cutting process parameters such as drill size, drilling speed and feed and also the influence of the work piece material. For this purpose regression analysis techniques are studied together with fuzzy logic and the hierarchical structure of the diagnostic program.

Although tool wear monitoring in principle has similar challenges for a number of tool types and it could be argued that the same approach could be used, this work concentrates only on drilling, which is the most widely used machining method and which has some specific features that tend to make it more difficult to monitor. These challenges are e.g. the discontinuous nature of the drilling process, the great variation in tool size, the difficulties in positioning the measuring sensors and the complexities of modelling the drilling process.

The actual drilling process and drill wear as a physical phenomenon are not covered in this work, i.e. only the indirect monitoring signals are studied. Similarly the developed simulation model does not try to model the drilling forces in such a way that these could be used for the machining process. Instead, it merely tries to mimic the features of the measuring signals based on some characteristics of the drilling process.

The aim of the developed approach is not to differentiate between types of drill wear such as chisel, corner, crater, flank and land wear. The purpose is simply to detect whether the drill starts to get so worn that is should be changed. In addition tool breakage, which is the typical failure mode of smaller size drills, is not covered.

1.6 Scientific contribution of the thesis

The scientific contribution of the thesis can be summarised as the development of an automatically adaptive diagnostic tool for drill wear detection. The new approach is based on the use of simplified fuzzy logic and higher order polynomial regression analysis, and it relies on monitoring methods that have been tested in this thesis. The diagnosis program does not require a lot of memory or processing power and is thus capable of handling a great number of tools in a machining centre. The work consists of:

- Extensive testing of monitoring methods and signal analysis techniques. Evaluation of the best combination of monitoring methods and signal analysis techniques for drill wear monitoring.
- A simplified simulation model has been developed which can be used to produce data with features similar to real data, thus the model helps to understand why certain analysis techniques work and others do not. Especially the importance of natural vibration modes of drills and the influence of drill size on these becomes apparent with the model. This model can be used in the definition, training and testing of an automatic diagnostic tool based on artificial intelligence.
- The development of higher order polynomial regression functions with a limited number of terms which can be used for filtering the monitoring signals, i.e. they remove individual peaks from the measuring signals. The regression functions also reduce the amount of data that needs to be saved, i.e. only the summary terms of the regression functions need to be saved in order to be able to follow the trend of the monitoring signals.
- The introduction of a term into the regression functions, which controls the amount of emphasis that older data is given, compared to the current data. This feature makes the regression function fast enough to react to the rapidly developing increase of monitoring signals at the end of the

drill life. It also enables to some extent the feature that monitoring signals can adapt to changes in cutting parameters or to a change of work piece material.

- The regression functions can be used to give a prognosis of the remaining lifetime of the drill at the end of the drill life. In theory, this kind of prognosis could in fact be done fairly early assuming constant cutting conditions and homogeneity of the work piece material. However, in practice the warning of the end of the drill's lifetime is given in terms of a few percent of the total lifetime prior to the final end.
- The development of an automatic diagnosis method based on the use of multiple signals and a simplified fuzzy logic approach.
- The use of hierarchy in the diagnostic approach in order to make it possible to combine signals and parameters from a number of sources, such as the tool wear monitoring parameters and cutting process parameters.

2. Drill wear monitoring

Successful tool wear monitoring requires that a number of technical tasks are understood and handled. The wear process must be understood in order to be able to use proper monitoring signals and signal analysis techniques. Diagnostic methods that can analyse the state of the tool automatically must also be understood. Because of the complexity of the problem many different types of approaches have been developed and tested. There exist a few good summaries and reviews of what has been published in the technical literature in this field, such as those by e.g. Rehorn et al. [2004] and Byrne et al. [1995]. Dimla et al. [1997] give a review of neural network solutions and include information about the sensor signals used. In an older review, Cook [1980a] lists both direct and indirect methods that have been used for tool wear monitoring and provides literature references. Also the somewhat older review by Tlusty & Andrews [1983] focuses on sensors used in unmanned machining. Li & Mathew [1990] give a good summary of wear and failure monitoring techniques that have been used in turning, which is the most widely studied machining process as regards tool condition monitoring [Jantunen 2001]; it is probably the easiest to monitor because the work piece rotates rather than the tool. There is also a database [Teti 1995] of references related to tool condition monitoring, which inspired the compilation of the database reported by Jantunen [2001]. Publication I gives a more thorough summary and publication III discusses the benefits of various measuring signals and signal analysis techniques. Drill wear is also covered in publication VII.

2.1 Drill wear

Tool wear, and especially drill wear, is a rather complicated phenomenon. Drilling operations differ significantly from turning and face milling for several reasons [Rehorn et al. 2004]. The major difference is the fact that drilling is a complex three-dimensional material removal operation, unlike the relatively simple cases of orthogonal and oblique cutting. Drills have vastly different geometries compared with turning and face milling tools. They are usually much longer than a turning cutter and have far less cross-sectional area than a face milling cutter. Drilling operations are different in that they require the full immersion of the tool, rather than operating on the periphery or surface as is the case in face milling and end milling. Altogether seven different types of drill wear can be recognised [Kanai & Kanda 1978]: outer corner wear, flank wear (actually two types), margin wear, crater wear, chisel edge wear and chipping at the lip. Because of adhering material many of these wear types are in practice difficult to measure; therefore the outer corner wear has been used as a measure of drill wear since it can be easily and reliably measured [Kanai & Kanda 1978]. It is not within the scope of this work to try to measure directly or to increase the understanding of what happens when a drill gets worn. Instead, it is recognised that in principle drill wear is an accelerating process that takes place at the outer margin of the flutes of the drill due to intimate contact and elevated temperatures at the tool work piece contact [Thangaraj & Wright 1988]. Thangaraj & Wright [1988] explain that there is a period of initial wear, then a period of moderate wear and in the third phase a period of excessive wear. Due to production variations, a new drill is typically slightly asymmetric. Accordingly, the two corners of the drill point wear gradually while the maximum wear alternates from one cutting edge to the other. This alternating process continues until both lips have zero clearance at the margin. The drill then adheres to the work piece and breaks if the cutting process is not stopped in time. In addition, chip flow creates significant friction between the cutter and the work piece inside the drill hole. These frictional forces can significantly change the dynamics of the system and they can cause the cutter to break [Rehorn et al. 2004]. Drills, like other cutters, can fail due to either breakage or excessive wear. Based on tests it has been determined that drills of a diameter less than 3 mm tend to fail by fracture, while larger tools will fail through excessive wear [Thangaraj & Wright 1988]. In tests reported in the literature there is often great variation in the wear development of the tested drills, as in the tests with 160 drills reported by Kanai & Kanda [1978].

Drill wear is a highly complex phenomenon, and in the published literature no model exists that could describe it well enough to form a basis for drill wear monitoring. There are studies that describe the principles of tool wear in machining, such as those reported by Zhang et al. [2001] or Bhattacharyya & Ham [1969] who develop an approach to model flank wear. This model discusses the influence of various wear modes (adhesion, abrasion) and the influence of temperature but it does not look at the dynamics. It should be noted that this type of study usually concentrates on turning, which is a much more stable process than drilling. Material has been published on how to evaluate the lifetime of a tool, for example by Cook [1980b]. There are models that can be used to calculate the drilling forces, e.g. in the work by Williams [1974] or Watson [1985a, 1985b, 1985c, 1985d] for the estimation of static force components. In the references the geometry of the drill is taken into account sector by sector and a computer program to calculate the feed and the torque is presented. Liu [1987] presents a model to calculate the thrust and torque of multifacet drills as a function of drill geometry based on the summation of terms calculated for a number of segments. Chandrasekharan [1996] takes into account the drill geometry and his model is capable of predicting the drilling forces in the different phases of drilling a hole (tool entry, cutting lips only, entire drill). Also the rotational effects can be modelled. Following the principles shown by Chandrasekharan [1996], Yang et al. [2002] introduce dynamics into their model. Many of the references studied in this thesis show how important the dynamics are in drilling and how the dynamic response increases as a consequence of drill wear. Rotberg et al. [1990] show the most important vibration modes. They suggest that the spikes in vibration monitoring of drills are generated when the drill tends to stick in the work piece for a very short instant (stick slip) and as a consequence the drill tends to unwind. In this phenomenon both torsional and compressive stresses are included. As the twist increases, the drill releases and continues cutting and hence the impulsive nature of vibration is introduced. It is suggested that this phenomenon becomes increasingly severe as wear develops.

2.2 Monitoring methods

A great variety of monitoring methods have been used and tested for tool wear monitoring. In principle there are two possible approaches, i.e. direct and indirect methods. Direct methods measure tool wear directly, which means that these methods actually measure tool wear as such. Unfortunately these direct methods that can be based on visual inspection or computer vision etc. have not become economically or technically advanced enough for use in industry, therefore they are not studied here. Instead of wear, indirect monitoring methods measure something else which must be a function of wear. Publication I gives a summary of indirect monitoring methods that have been applied to tool condition monitoring in drilling. The following chapters give a brief description of the most widely used monitoring methods and try to explain why these methods can be expected to work. A brief description is given of the most commonly used measuring methods, signal analysis techniques and fault diagnosis approaches.

2.2.1 Torque, drift force and feed force

Measuring of cutting forces is very popular in all types of cutting processes. In the summary given in publication I the measurement of feed force is the most popular method used in drill wear monitoring tests. The second most popular method is to monitor torque. It is logical to monitor the cutting forces since they increase as a function of wear as reported e.g. by Lin & Ting [1995], Pan et al. [1993] and Subramanian & Cook [1977]. In theory, drift force would not work in the case of twist drills with two cutting lips, since these two cancel the influence of each other and the forces are in equilibrium and thus no indication of drill wear should be seen. However, due to production tolerances the cutting lips are not exactly identical and a drill is slightly asymmetrical. Therefore, it only wears at one lip until the height of both lips is equal [Barker et al. 1993, Braun et al. 1982, El-Wardany et al. 1996]. The second lip, which is now sharper, starts to cut and this process of alternating the cutting lip continues until neither lip has any more clearance at the margin. Although the measurement of cutting forces has been a very popular and successful monitoring method in laboratory tests, there is a drawback related to their use in normal production. The measurement of cutting forces is not easily arranged between the tool, tool holder and spindle. A force or torque transducer is relatively big and possibly makes the change of tools more complicated.

Aatola et al. [1994] gain the best indication of drill wear with feed force and torque measurements, but at the same time they suggest that the big and heavy force and torque transducer used in the tests might have had an adverse influence on the measured vibration. Another option is to make the measurement of drilling forces from the other direction, i.e. below the table where the work piece is positioned. Unfortunately this kind of measurement chain is somewhat longer so that the forces are measured further away from the drill.

Von Nedeß & Himburg [1986] show the dynamic effects including the influence of the machine tool and the machining process on drilling and their influence on the feed force and torque. They point out that the drill wear causes a much higher increase of the dynamic components compared to the increase of the static forces. König & Christoffel [1980] have reached a very similar conclusion, i.e. the dynamic components of thrust force and especially torque are considered good indicators of drill wear. In the same reference torque is also considered good in indicating the risk of tool fracture, whereas thrust force is considered to indicate the actual tool breakage better when it has already happened. Also Christoffel & Jung [1981] explain how drill wear can be monitored indirectly with the dynamic components of thrust force and torque. They also explain the self-exciting nature of the dynamics. Brinksmeier [1990] points out the importance of being able to measure the dynamic changes of torque signal in order to monitor drill wear and fracture. For measuring the higher frequency content in a torque signal, a new sensor based on eddy current technology is introduced. However, the tested version of the new sensor is relatively big and not suitable for monitoring drills with a smaller diameter. Brinksmeier [1990] predicts that the size of the sensor can be reduced, enabling a wider size range of drills to be monitored.

Li et al. [1992] verify that the dynamic components of feed force and torque give a clearer indication of tool wear than an increase in the average level. In this case an attempt is also made to define the rules of how different wear modes (chisel, flank and corner) can be distinguished from each other together with the capability of detecting tool breakage. The dynamic influence in thrust force and torque is also emphasised by König & Christoffel [1982]. With a drill diameter of 8 mm they demonstrate how big the change is in the spectrum of thrust force at a frequency of 1050 Hz. It is also pointed out how great the difference is in the roundness and shape of the drilled hole of a sharp drill and a worn drill, the difference being linked to the radial vibration of the drill.

McPhee et al. [1995] emphasise to the frequency content of the drilling power measured using a dynamometer. The drills in question are coated. It is noted and measured in the study that the dynamometer has a remarkable influence on the vibration response of a drill. In that study the most interesting frequencies with a 6 mm diameter drill are around 800 Hz which is related to the dynamics of the dynamometer, and around 2250 Hz which is considered to be linked to the drill. It is concluded that frequency analysis may assist in distinguishing between jamming and failure.

Lenz et al. [1978] have studied the influence of wear on drift forces. In their study, however, the feed force and torque do not give a similar indication. The results seem to support the idea that during drilling, the cutting moves from one lip to another as discussed previously.

2.2.2 Vibration and sound

Vibration is the most widely used measuring method in condition monitoring of rotating machinery. However, it has not been as popular in drill wear monitoring, possibly due to the amount of noise in a typical cutting process. Vibration measurement is easily arranged, since an accelerometer can easily be installed close to the spindle bearing and no modifications of the machine tools or the work piece fixture are needed [El-Wardany et al. 1996]. There is no effect on stiffness and damping properties of the drilling system and the sensor can also be mounted on the table close to the cutting action [Abu-Mahfouz 2005]. Abu-Mahfouz [2005] points out that accelerometers, when properly shielded, have good resistance against coolants, chips, electromagnetic and thermal influences. It is logical to expect vibration measurements to react to tool wear, because if in a dynamic system such as the machine tool the cutting forces increase, the dynamic response will also increase. As explained in the previous chapter, the drift forces can be used for monitoring drill wear, and these forces are also the cause of increasing vibration as a function of wear. Unfortunately there are a number of drawbacks related to vibration monitoring. Besides the influence of tool wear, the vibration signal is influenced by the work piece material, cutting conditions and machine tool structure.

Abu-Mahfouz [2003] has used vibration measurement to detect drill wear and also to differentiate between different types of wear, i.e. chisel, crater, flank, edge and outer corner wear. Narayanan et al. [1994] concentrate the diagnosis of drill bit wear on higher frequencies around 10 kHz. From their results it seems clear that the best indication of drill bit wear is seen at these frequencies. However, the geometrical details of the tool and tool holder are not reported and there is no explanation of the reasons why these frequencies are the best for drill bit wear monitoring. Also Barker et al. [1993] used vibration acceleration for monitoring the wear of drill bits, which were for drilling holes into electronic circuits.

Similarly to vibration, also sound can be used for drill wear monitoring. Mechanical vibration of the machine tool, tool holder and tool is partly transferred to airborne vibration, i.e. sound. Consequently the same information observed from vibration signals can be obtained from sound measurements recorded with a microphone. Sound measurements, although very easy to perform, have not been widely used, probably because they are affected by noise to an even greater extent than vibration measurements. In the tests covering a number of monitoring methods reported in publication II, vibration monitoring was the most effective method.

2.2.3 Acoustic emission and ultrasonic vibration

In addition to mechanical vibration up to 20 kHz, a higher frequency range has been used for monitoring drill wear. Vibration measurements in the frequency range 20–80 kHz are in the literature called ultrasonic vibration [Hayashi et al. 1988]. The use of ultrasonic vibration has been justified by pointing out that at lower frequencies structural vibrations are dominant, and that higher frequencies suffer from the joints commonly found in machine tools; thus ultrasonic vibrations are especially suitable for e.g. drill breakage detection. There are a few other studies, such as those by Kutzner & Schehl [1988], König et al. [1992] and Schehl [1991], which describe the results with ultrasonic vibration measurements, but the technique has not been widely used. It should also be noted that in these studies the emphasis is on such a low frequency range (most of the information was obtained at frequencies below 60 kHz) that although König et al. [1992] and Schehl [1991] describe it as acoustic emission, some others would call it ultrasonic vibration.

It is interesting to note that results reported with drills with very small diameters from 1 to 3 mm are good with this technique. König et al. [1992] point out that with such small drills the spindle current does not work, cutting forces do not give as good an indication as acoustic emission, and especially with the smallest drill diameters it is not possible to predict the upcoming tool breakage; acoustic emission does, however, give some indication even with such small drills. In another study [König et al. 1989] the same research team recommends the use of the frequency range 5–40 kHz.

König et al. [1992] discuss the advantages of using acoustic emission in monitoring drill wear, especially that of small drills. However, their signal analysis technique of a band passing the signal in the frequency range 1–5 kHz actually means that this kind of measurement is normally defined as mechanical vibration, although the used sensor is capable of measuring higher frequencies up to those defined as acoustic emission. Waschkies et al. [1994] suggest the use of an average value of acoustic emission measured in a wide frequency range of 0.1–1 MHz for drill wear monitoring.

2.2.4 Spindle motor and feed drive current

Spindle motor current is in principle related to measuring torque, although the measuring chain is longer. Similarly, measuring the feed drive current can be considered identical to measuring thrust force, although again through a longer measuring chain. Since they are so easy to measure, both the spindle motor current and feed drive current have been used relatively widely in test, e.g. by Adamczyk [1998], Li [1999], Ramamurthi & Hough [1993] and Routio & Säynätjoki [1995]. Li [1999] reports good results with spindle current and feed force current monitoring of breakage of small drills. The tested drills have a diameter of 1–4.5 mm, i.e. they are so small that the breakage is the typical failure mode [Thangaraj & Wright 1988].

Ramamurthi & Hough [1993] use the spindle motor current and feed motor current for tool wear detection with good results. In this case these signals are used together with thrust force measurement, which is used to predict tool failure. One of the purposes of their study was actually to test whether the current sensor would be sufficient for drill wear monitoring, since it is cheaper and easier to use than other measuring methods. From this it is concluded that if wear is not diagnosed then tool failure is predicted or vice versa, i.e. in this case the combination of two measuring techniques improves the reliability of the diagnosis.

Kim et al. [2002] predict the flank wear of a twist drill based on measured spindle motor power. The developed theory starts from the model reported by Williams [1974]. The cutting torque is divided into three components, i.e. lip, chisel and margin components. Of these only the lip component depends on the flank wear of the drill. This dependency is shown to be remarkable. In the tests the accuracy of predicting the drill wear for a drill with a 4 mm diameter was 0.02 mm for flank wear, whereas the flank wear criterion requiring drill replacement was 0.18 mm. Because of the structure of the model, also the effect of the feed rate change can be handled.

Adamczyk [1998] emphasises to the disengagement phase of the drilling process. In this phase both the feed drive current and spindle current were highly correlated with the flank, corner and margin wear of a drill with a 10 mm diameter. In fact the correlation was higher with current measurements than with acceleration measurements. The results reported by Routio & Säynätjoki [1995] on the use of spindle power for drill wear monitoring are not encouraging.

2.3 Signal analysis techniques

Various signal analysis techniques have been used in the context of drill wear monitoring. It is very important what kind of signal analysis technique is used. In principle the signal analysis tries to identify the meaningful part of the signal that is giving an indication of wear, and to remove the noise, i.e. parts of the signal that do not contain or show a wear-related trend. The used signal analysis method should be quick to perform, because during drilling the wear progresses very rapidly towards the end of the tool life, as explained in Chapter 2.1 of this thesis. In a case where drill wear is monitored in a machine tool where a great number of tools might be used, the amount of data that needs to be saved in relation to signal analysis is of some importance. Thus if a lot of information needs to be saved in order to follow the trend in parameters calculated with the signal analysis as a function of wear, the hardware must have sufficient data storage capability. The following chapters give a short introduction to the most important signal analysis methods and how they have been used in the reported literature. Publication I gives a more thorough presentation of the current use of various signal analysis methods in drill wear monitoring.

2.3.1 Time domain signal

The time domain signal is the first thing that is seen when a measurement is performed. Typically today the measurement is performed using a computer with an AD card or with some measuring equipment that performs the AD conversion. Already at this stage the frequency at which data is gathered influences the result, i.e. if data is gathered at a lower frequency than what the transducers can measure, this actually means that information at high frequencies is not properly treated. It is not practical to save the raw time data for long periods of time and for a number of tools. Typically some statistical parameters are calculated from the time domain raw data, and these parameters are then saved and used for diagnosis of tool wear. When calculating the statistical parameters the choice of sample length influences the results. The root mean square (rms), arithmetic mean, standard deviation and kurtosis are examples of time domain statistical parameters. Formulae for calculating these parameters are found in a number of books, e.g. that by Press et al. [2002] which also gives the computer code in C++ to calculate the most typically used parameters.

Noori-Khajavi [1992], Noori-Khajavi & Komanduri [1993] and Noori-Khajavi & Komanduri [1995a] used mean value and variance with force and torque transducers. In these tests no correlation with drill wear was found in the time domain. Lin & Ting [1995] have used average values of thrust force and torque. The test material was used for developing a model to calculate the force and torque as a function of drill feed, diameter and wear. The authors conclude that the models can be used for wear estimation. Liu & Anantharaman [1994] used average, peak, rms values and the area of thrust and torque with success. Radhakrishnan & Wu [1981] used mean, peak and standard deviation values of thrust force and torque signals. In these tests the standard deviation, which in practice is the same as the rms value, proved to be the best indicator of wear.

Thangaraj & Wright [1988] calculated the mean, standard deviation and maximum values of thrust force sampled at a low frequency of 40 Hz for each hole. With this kind of approach the maximum value gives the best indication of wear. The results with mean, minimum and maximum values of cutting forces reported by Valikhani & Chandrashekhar [1987] are not promising. The statistical parameters were measured for each hole. It is noted that the fluctuation of forces increases with drill wear, which could give grounds for drill wear monitoring. Tansel et al. [1992] report good results in monitoring the breakage of micro-size drills using average and standard deviation values of thrust force. In this case the statistical parameters were studied in four different segments of drilling a hole. Ramamurthi & Hough [1993] used a number of statistical parameters in connection with spindle and feed motor current together with thrust force. The statistical parameters are the rms value (spindle motor current), mean value (thrust force) and rms value of the high pass filtered signal (feed force current) together with a parameter that indicates the increase in each of these as a function of drilling time/wear.

Schehl [1991] used band pass filtered rms values of acoustic emission with success. König et al. [1992] suggest the use of band pass filtering of acoustic emission together with the use of a rectifier. The technique has the advantage that in this way acoustic emission signal can be gathered at a relatively low frequency, which makes the measuring and analysis equipment much cheaper. Routio & Säynätjoki [1995] have used maximal stable values of feed force, torque, spindle and feed drive current. In these tests the indication of wear and tool failure was observed very late because the analysed signals were almost constant until they rose very dramatically in the last hole. El-Wardany et al. [1996] use the kurtosis value, which is an indicator of peakedness of the signal, together with a new parameter called ratio of the absolute mean value, for analysing vibration for drill wear and failure monitoring successfully. Kutzner & Schehl [1988] suggest the use of a band passed high frequency vibration signal for monitoring small diameter drills. The basic idea is that the rotational natural frequency should lie in this frequency range.

2.3.2 Fast Fourier transform

Fast Fourier transform (FFT) is a means to determine the frequency content of a measured signal. The principles of FFT can be found e.g. in a book written by Randall [1977]. Basically, the idea of looking at the frequency content of a measured signal is based on the concept that at some frequencies wear influences the signal more than at some others; thus FFT serves as a means to eliminate meaningless information and emphasise more meaningful information instead. Braun et al. [1982] discuss the effectiveness of using FFT in the development of a trend index for sound signal monitoring together with the use of an enveloping technique. El-Wardany et al. [1996] use FFT to calculate the power spectrum and also cepstrum. The power spectrum is used for monitoring the drill wear of large drills with a drill diameter of 6 mm. The cepstrum with statistical parameters explained in the previous chapter, are used for detecting the tool breakage of smaller size drills with a drill diameter of 3 mm.

Valikhani & Chandrashekhar [1987] have, alongside the statistical functions explained earlier, also used the power spectrum successfully to monitor tool wear based on the drift force. However, they indicate that the amount of test material is limited and suggest further testing. Govekar & Grabec [1994] used a relatively small number of points, 256 in the time domain instead of the typical 2048, for FFT when measuring torque and feed force. The reason for this choice is apparently the use of neural networks (self organising map) as the following diagnostic tool in the approach.

Noori-Khajavi [1992], Noori-Khajavi & Komanduri [1993] and Noori-Khajavi & Komanduri [1995a] report that use of the power spectral density (PSD) function gave better results in drill wear monitoring than the statistical parameters described in the previous chapter. The PSD function was calculated for thrust force, torque and strain measurement in two horizontal directions. In this case relatively low frequencies from 50 Hz to 300 Hz gave the best results. No individual frequencies were considered; instead the change of area under the PSD plots was used. Barker et al. [1993] tested higher order spectral (HOS) functions calculated for vibration for drill wear detection, and compared these with the normal power spectrum approach. With the tested material the HOS approach gave a higher detection rate of drill wear, although at the same time the false alarm rate also increased.

2.3.3 Other analysis techniques

Envelope detection is one method of signal analysis that has become popular especially in rolling bearing fault detection. Envelope detection is a means of looking at the signal energy contents in a certain frequency range. Typically this range is rather high, i.e. of the order of 10 kHz, and the idea is that by using band pass filtering it is possible to concentrate on the information in this range. Braun et al. [1982]] and Braun & Lentz [1986] suggest the use of envelope detection or a somewhat further developed signal analysis technique which can pick up the information at higher frequencies for drill wear monitoring using sound signal measurements. Hayashi et al. [1988] used envelope detection of high frequency vibration (20 kHz – 80 kHz) together with a statistical parameter called the clipped running mean, i.e. a running mean from which some higher peaks that pass a certain threshold value have been clipped away. Together with this parameter, the number of occurrences of values that are higher or lower than certain limits that have been calculated in relation to the clipped running mean are followed. These then give an indication of tool breakage.

Drilling a hole is not a stable process in that the measured signals vary from the beginning to the end of drilling a hole. Quadro & Branco [1997] recognise five stages and two of these are considered best for monitoring drill wear using acoustic emission. In this study acoustic emission is studied using the measured area under the rectified signal envelope (MARSE). One approach that can be used in signal analysis is autoregressive modelling. Radhakrishnan & Wu [1981] use the autoregressive moving average (ARMA) model for modelling the thrust force and surface waviness. The approach is suggested for use in on-line monitoring of drill wear.

Wavelet transform is another method that can be used to extract meaningful information from the measured time signal. The principles of wavelet analysis can be found e.g. in a book written by Newland [1993]. When compared to FFT, which only gives information in the frequency domain, or the time domain parameters, which only contain information in the time domain, a wavelet can be considered to include both of them, i.e. information in the time-frequency domain. Li [1999] used wavelet transform for drill breakage detection based on AC servo motor current measurements of all four axis motors. The drill size in the tests was small, from 1 mm to 4.5 mm in diameter. However, the diagnosis was passive, i.e. there was no warning prior to actual breakage. Tansel et al. [1993] used wavelets to diagnose a severely damaged micro drill from a new one. The monitored signal was thrust force. Again there is no indication whether a warning was obtained prior to the drill being severely damaged. Hiebert & Chinnam [2000] used wavelets to analyse the thrust force and the torque. Some of the wavelet parameters were used as input into a neural network, which aimed to diagnose the degradation of drill bits. The reliability of the method is discussed and it is noted that since many degradation signals increase in slope as they approach failure, the accuracy of failure predictions should increase when approaching the critical limit.

Abu-Mahfouz [2003] combines and also compares, in the case of a vibration acceleration signal, the effectiveness of statistical time domain parameters such as mean, variance, skewness and kurtosis, together with parameters calculated using discrete harmonic wavelet transform and the eight highest peaks calculated with the Burg power spectral density function. In the approach, different types of wear can be detected and in that study the parameters calculated with the wavelet transform proved to be superior compared to the other methods.

2.4 Fault diagnosis systems

Today machining processes are usually automatic and unmanned. However, various types of problems or faults in the process necessitate manual intervention. Tool wear and breakage is one of the factors that prohibit fully automatic production in three shifts. If tool wear and breakage monitoring is used, in practice it needs to be automatic, i.e. the system used for tool monitoring needs to be able to diagnose the condition of the tool automatically, which means that some sort of artificial intelligence is involved.

Tönshoff et al. [1988] define the components that are needed in a tool wear monitoring system: sensor, signal conditioning, model and strategy. The three first components are covered in the previous chapters. Strategy means that different actions are taken based on the monitored signals. A monitoring system only gives an indication or alarm if the signals reach a certain level. A diagnostic system tries to find a functional or causal relation between the failures in machining and their origin. Adaptive control systems automatically adapt machining conditions according to a given strategy. Tönshoff et al. [1988] also point out the advantages and challenges of multi-sensor systems, and how they bring more information. At the same time the importance of building multi-model systems is explained. It is claimed that the use of more sensors and models results in a more reliable and more flexible supervising process and increases the feasibility of better control.

Ertunc et al. [2001] employed Hidden Markov Models (HMM), which have successfully been used in speech recognition, for drill wear detection based on thrust force and torque. In the approach three different stages of the tool were recognised, i.e. sharp, workable and dull. It is suggested that different models should be defined for different cutting conditions since these influence the results. In addition to the HMM approach, Ertunc & Oysu [2004] tested a so-called phase plane method. They report that one of the benefits of this approach is its simplicity, since the thrust force is plotted as a function of the torque, and if the tool is in a normal condition the plotted results stay within a predefined rectangle. The authors state that even though the method is very simple, it does give satisfactory criteria for monitoring tool wear.

Liu et al. [2000] report the results of using a polynomial network for predicting corner wear in drilling operations. The input parameters are cutting speed, feed rate, drill diameter, torque and thrust force. The development of a polynomial network is rather straightforward, but it means that the network is first trained with suitable data. Liu et al. [2000] had 27 training cases and eight test cases. It is concluded that the use of thrust force gives a more reliable indication than the use of torque. The difference between the predicted corner wear and measured corner wear was less than 10% with the test data.

2.4.1 Predefined limits / rule based systems

The simplest way to automate the diagnosis of tool condition is to use predefined limits for the measured signals and parameters calculated from those signals. This means that if a parameter value exceeds the limit given to it, the tool is considered worn. The approach can be made more reliable by combining the information from various sensors and/or calculating a number of parameters of these signals. This information can be combined with the information from the cutting process parameters, e.g. using the so-called rule based approach in building rules, i.e. the knowledge base, so that a number of conditions need to be fulfilled simultaneously. One example could be that if the drill diameter is more than 4 mm and less than 5 mm and the drilling speed is … and … etc then …. Erdélyi & Sántha [1986] describe the principles of this kind of approach in general for a production cell. Publication V addresses the principles of this type of approach in greater detail for tool wear monitoring.

The use of sophisticated analysis methods can be seen as one attempt to make the use of predefined limits more reliable and possibly more general. If the parameter that is used to detect tool wear is insensitive to other factors, such as the cutting speed, it is easier to build rules that define the condition of the tool. This highlights one drawback of the rule based approach. If many different types of tools are used in the machine tool, it might be very laborious to build a rule based expert system that can detect tool wear and warn of the upcoming breakage. However, the rules might also be very simple for each machining state/tool and one way to define the limits is simply to define them manually for each tool type.

Another possibility to make the definition of limits more general is to use trending, which means that the parameter values are saved when the tool is in good condition and the limits are defined at the beginning for the relation of the current measurement to the measured value. For example, Thangaraj & Wright [1988] use the gradient of the thrust force and state that the proposed control system does not require considerable tuning for operation under a wide range of cutting conditions. Another example is given by El-Wardany et al. [1996], who perform the more sophisticated analysis only when a certain parameter reaches a predefined value compared to the initial value. Also Lechler [1988] discusses the definition of limit values and how they can be used for tool wear and fracture monitoring with various force and strain based sensors. They point out how important it is for the personnel to have sufficient training.

Adamczyk [1998] suggests a relatively simple combination of rules based on standard deviation values of the feed drive and spindle current for the stable and transient phase of drilling. Basically, if a simple condition is fulfilled in both conditions the drill is considered worn. Adamczyk [1998] shows a simple procedure for combining information from three different sensors (two current and one accelerometer). Takata et al. [1986] present some results with the pattern recognition technique, which is based on speech recognition. The signal measured and analysed with a sound sensor forms a 16 x 16 time/frequency pattern which can be used for defining the cutting state and detecting a broken tool. Tönshoff et al. [1988] describe the principles of building a rule based approach that relies on the information from three different types of sensor: force, temperature and vibration.

Li et al. [1992] use a simple rule set based on the relationship between the current value and the average value of feed force, torque and their dynamic components. One of the advantages of the approach is that there is no need for training or definition of the limits; instead they are calculated for each of the monitored tools. The rule set can also distinguish between various types of drill wear. However, the approach has been developed based on only four tested drills, which unfortunately raises the question of how general the results actually are.

2.4.2 Fuzzy logic

The rules in rule based systems are usually crisp but they can also be fuzzy, i.e. not exact. The principles of fuzzy logic can be found e.g. in a book by Rao & Rao [1993]. Li & Wu [1988] categorise drill wear into four fuzzy classes: initial, small, normal and severe. In this approach fuzzy limits are defined based on an algorithm used for clustering thrust force and torque data. When the data is analysed, the result is not crisp but shows membership to each of the four classes. The approach works, although only two test cases are shown. In the approach, only the parameters (rms value) related to thrust and torque are used and a so-called c-mean algorithm is used for defining the relationship between the tool conditions and the measured parameters. Du et al. [1995] describe the c-mean algorithm in a more general form together with other possible approaches to linking together the measured parameter values and state of the tool. Xiaoli & Zhejun [1998] used this kind of approach for monitoring tool wear during boring. The monitored seven parameters in this case were from wideband AE measurements which had been treated using wavelet transform. The seven parameters were actually a set chosen from 16 frequency bands. The authors conclude that the proposed approach can give a high success rate over a wide range of cutting conditions.

Du et al. [1995] justify the use of fuzzy classification by claiming that for dealing with uncertainties inherent in the metal cutting processes, fuzzy systems offer the advantage of providing systematic means for describing the relationship between tool condition and various process signatures. Fuzzy logic can also be used in connection with neural networks for pre-processing input data into the network and/or post-processing the output of the network [Rao & Rao 1993].

Li & Tso [1999] develop regression models for spindle motor current and feed motor current as a function of cutting variables, i.e. cutting speed, feed rate and drill diameter, for various flank wear states. Using fuzzy classification it is then possible with the test data to predict the membership in three different wear states. The number of definition cases for development of the regression functions is 12. In this set eight cutting speeds are used together with five feed rates and three drill diameters. The number of test cases is also 12. The result is considered good since the grade of membership function associated with the relevant flank wear states is always close to unity. However, although the results in the paper are considered good, the relatively small number of test cases compared to the number of input parameters raises some questions about the generalised nature of the methodology.

Li et al. [2000] used fuzzy logic together with neural networks. In this case drill wear is monitored using vibration acceleration. The rms value in five separate frequency bands between 0 and 2500 Hz are used as input features. Drill wear is categorised in five different classes: initial wear, normal wear, acceptable wear, severe wear and failure. It is concluded that a fuzzy relationship between the tool condition and monitoring may be identified by using a fuzzy neural network. However, the recognition rate for initial wear is reported to be 52% and for severe wear 68%. Drill failure and air cutting have been recognised at a rate of 100%.

2.4.3 Neural networks

Neural networks have become very popular in industry because of their classification and optimisation capabilities [Dimla et al. 1997]. Neural networks can be seen as an attempt to automate the process of building a diagnostic system. In principle neural networks can be trained to model non-linear dependencies of manufacturing process parameters and parameters which indicate tool wear and failure. The principles of neural networks can be found e.g. in a book by Rao & Rao [1993]. Dimla et al. [1997] critically examine 37 approaches that have been tried with different types of neural networks in order to diagnose tool wear and breakage in various types of machining processes. The success rate is tabulated based on references. Some of the main conclusions by Dimla et al. [1997] are: The most widely tested neural network approach is a so-called multilayer perception (MLP) network. MLP networks are particularly suitable for high-speed real time applications. In many cases more than one feature has been extracted from one sensor and this is criticised as not really being a multi-sensor approach. Although most of the references claim to be on-line solutions they actually seem to be off-line networks, which have not been tested in a real production environment. In most cases the data has been sampled using only one set of cutting conditions. A tool condition monitoring system needs to be able to handle various cutting conditions.

Liu & Ko [1990] built a simple network comprising two input features and one output. Drill wear was classified into five categories. The inputs were peak to peak acceleration and the percentage increase of the thrust force. They concluded that an on-line recognition level of over 85% can be reached. The limited number of tests did not include variation of cutting process parameters. The same data was used to develop a two-category linear classifier for drill wear detection in studies by Liu [1987] and Liu & Wu [1990]. In this case a success rate greater than 90% is reported for drill wear monitoring in one drilling process condition.

Liu & Anantharaman [1994] tested the influence of the number of hidden layers. In the cases tested the number of input features was nine based on thrust force, torque and one process parameter. It is concluded that artificial neural networks can distinguish between a worn and a usable drill with 100% reliability and also accurately distinguish the average flank wear even under different drilling conditions. However, the authors have not included documented material of the variation of cutting conditions. They compare different versions of the number of neurons in the hidden layer and also a modified version with adaptive activation-function slopes. This modified neural network is reported to converge to a solution much faster than a conventional feedforward network.

Liu et al. [1998] introduced the influence of drill size, feed rate and spindle speed together with the same thrust force and torque parameters used earlier in the neural network solution. They report that the network can reach up to 100% reliability for on-line detection of drill wear states and that it is feasible to recognise the drill wear states on-line even if the drill size, feed rate and spindle speed have changed. However, it should be noted that there was no variation of the work piece material and that the total number of tests was seven, in which five different drill sizes, six feed rates and five spindle speeds were used, which would suggest that the number of test cases was rather small compared to the number of influencing parameters.

Noori-Khajavi [1992], Noori-Khajavi & Komanduri [1993] and Noori-Khajavi & Komanduri [1995b] use neural networks for sensor signal integration. This is done based on torque, feed and drift force signals. Noori-Khajavi [1992] shows that it is not advantageous to integrate information from these because they are equally good and contain the same information of drill wear. Govekar & Grabec [1994] use a self-organising neural network. Torque and feed force spectra are further treated so that the low frequency information below 200 Hz is left out and the information at higher frequencies is combined into 30 representative bands. They conclude that the approach is promising. The effect of cutting process parameters is not covered.

Tansel et al. [1992] tested a different kind of neural network called a restricted coulomb energy (RCE) network for drill wear diagnosis in micro drilling. The theory of RCE network is explained in their report. The drilling of each hole is divided into four segments and the average and standard deviation of feed force is used as the input features, i.e. altogether eight inputs. The RCE network recognized tool failure with an accuracy of over 90%. The processing parameters were not varied, although it is pointed out that the feed force varied a lot from test to test. The same test data as in the previous reference has been tested in connection with another type of neural network based on adaptive resonance theory (ART) [Tansel et al. 1993]. In this case the input features were calculated using wavelet transform of the feed force. The approach was tested with two network structures, one with 22 input features and the other with six. The approach with a higher number of input features gave a better indication, only one error in 61 cases, but was slower. Again there was no variation of process parameters.

Tsao [2002] tested two types of neural network solution for flank wear prediction of a coated drill based on maximum values of thrust force and torque. The two neural network methods were radial basis function network (RBFN) and a modified RBFN called adaptive RBFN (ARBFN). With a training set of 18 cases and a set of nine test cases good results were obtained. In the prediction the maximum drill wear error was only 0.4% which is a remarkable result. It should be noted that together with the variation of spindle speed and feed rate, also the drill coating deposition was varied. One thing that is clearly noticeable in the measured data is that the results are very consistent, i.e. the relation between the maximum thrust force and torque with the drill wear is very similar in all three cases for all the varied input parameter combinations, which would indicate that possibly very simple methods could give good results with the measured data.

Fu & Ling [2002] have developed a very basic neural network for the detection of breakage of micro drills. The solution is based on torque signal together with such parameters as the drill diameter, feed and spindle speed. The maximum and average values of torque were used. The approach works with very small drills but is passive in the sense that detection is made only after drill breakage has occurred, which is much easier than making a prognosis of breakage beforehand. There are benefits related to this late detection, although not as remarkable as in the case of prognosis.

Brophy et al. [2002] report the results of a project in which the network developed was based on input from a spindle power signal. In this case a network was developed to detect abnormalities in drilling. The spindle power was treated in the first stage with principal component analysis (PCA) to get the input features for the neural network. After a training phase of 3 weeks the neural network was tested in real production for 3 months. The authors report that the network draws similar conclusions to those of an experienced operator.

Abu-Mahfouz [2003] used a multiple layer neural network to detect drill wear and to differentiate between different types of wear such as chisel, crater, flank, edge and outer corner wear based on a vibration acceleration signal. From acceleration signal statistical time domain parameters together with wavelet based parameters and parameters of Burg power spectral density function were calculated. In the study, different types of architectures of the neural network were tested and also the process parameters, i.e. speed and feed, were varied. The reported results are promising. The percentage of correct predictions was around 80 to 90 when differentiating between the various artificially introduced wear types, and 100 when detecting drill wear. Based on the same measured signal and analysed parameters as described above Abu-Mahfouz [2005] reports the results of two other neural network approaches, namely learning vector quantization (LVQ) and fuzzy learning vector quantification (FLVQ), in detecting drill flank wear. Again the reported results are good with success rates of 86% with LQV and 88.8% with FLVQ. Also in this case the process parameters are varied. The test material was based on drilling tests in dry conditions covering the total tool life [Abu-Mahfouz 2005].

3. Machining tests

The complete test and measuring set-up and the test program are described in detail in publication II. In this chapter the main characteristics of the set-up and the drilling program are described briefly.

3.1 Test set-up

A horizontal-type machining centre was used in the drilling tests for tool condition monitoring. The main specification of the machining centre is shown in Table 1.

Table 1. Specification of the machining centre in the tests.

Machine tool	Niigata EN40B	Spindle nose	NT No. 40 for BT
Control unit	Fanuc 11 MA	Number of tools	30 tools
Controlled axis	4 axis (X, Y, Z, and B)	Spindle speed	15–6000 $^{1}/_{min}$
Table size	400 x 400 mm	Main motor power	11/7.5 kW

3.2 Test program

The twist drill sizes investigated in the tests were: diameter 3.3 mm, 5.0 mm, 6.8 mm, 8.5 mm and 10.2 mm. The drill material was HSS and the work piece material was Fe52. The total number of tested drills was 26. A description of the drilling parameters and monitoring methods is given in publication II.

3.3 Measuring arrangement

In the drilling tests the tested measuring methods included vibration, sound, acoustic emission (200 kHz and 800 kHz centre frequencies and also 100–1000 kHz frequency range), spindle power and current, z-servo current, force measured from guideways, feed force and torque with a dynamometer and 3-axis table dynamometer. In the tests the measuring signals were recorded with a 14 channel instrument tape recorder and analysed afterwards in the laboratory. The measuring configuration was varied during the measurements due to the limitations of the tape recorder, i.e. the number of channels used (12) was not sufficient for recording all the possible signals simultaneously. A more thorough description with a graphical presentation of the measuring arrangement is given in publication II.

4. Signal analysis

A detailed description of the signal analysis methods and results is given in publication II. Some results are also shown in publications III, VI and VII. Due to the great amount of test data, an automatic analysis program for PC was used. The data recorded with an instrument data recorder was analysed overnight with a PC equipped with an AD card. A mathematical programming toolbox MatLab was used for the signal analysis. The signal analysis was done both in the time domain (statistical parameters) and in the frequency domain (FFT analyses). Prior to the signal analysis the data was cleaned of irrelevant signals, i.e. data recorded during rapid movements of the tool prior to actual drilling. Regression analysis was used to rank the different methods used in the tests.

4.1 Statistical analyses

For all of the recorded measuring signals (12 sensors), except for the tachometer pulse used for recording the running speed of the tool, altogether eight statistical parameters in the time domain were calculated. These were: arithmetic mean, root mean square (rms), mean deviation, standard deviation, skewness, kurtosis, maximum and minimum. All of these time domain parameters are easy and fast to calculate [e.g. Press et al. 2002]. Usually they contain the whole frequency content of the measured signals and are therefore rather sensitive to noise, i.e. there is a lot of variation in the measured values. In the case of vibration signals, low-pass filtering was also tested to improve them. In drill wear monitoring the best results with statistical parameters were obtained with the root mean square and mean deviation of low-pass filtered horizontal vibration (cf. publication II). Figure 1 shows an example of the analysed root mean square value of a low-pass filtered horizontal vibration signal in drill wear monitoring.

4.2 FFT analyses

Fast Fourier transform (FFT) was used in the case of dynamic monitoring signals (vibration, force/torque, spindle motor power and sound) that were expected to contain frequency dependant information. A sample and hold card was used together with the normal AD card in order to analyse data simultaneously from four channels. A MatLab mathematical package was used for programming the tested functions. FFT based functions including autocorrelation, spectrum, 1/3 octave spectrum, 1/1 octave spectrum, cepstrum and liftered spectrum were tested for one signal at a time. For simultaneous analysis of more than one signal at a time, the tested functions were frequency response, coherence, coherent output power, cross-correlation, signal to noise ratio, Scot and multi-signal frequency response and partial coherence. In the signal analysis a Hanning window [Randall 1977] was used, as well as time and spectrum domain averaging. In order to save space, normally only the 20 highest amplitudes of each function were saved together with the corresponding frequency. As seen from the tabulated lists in publication II, it makes little difference whether the analysis is based on one or more signals. Due to the large number of analysis functions and analysed parameters, a procedure based on regression analysis was developed for further analysis of FFT based functions in order to define which of the measuring signals and analysis functions could be expected to work best for diagnosing drill wear. Of all the functions analysed with FFT the best results in drill wear monitoring were obtained with a horizontal vibration spectrum.

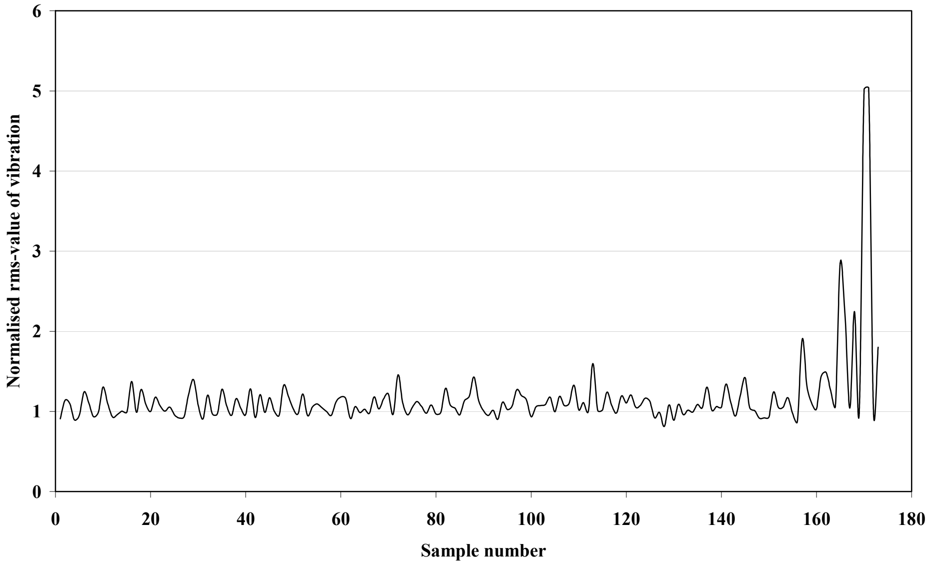


Figure 1. Normalised rms value of vibration (perpendicular to drill axis) for a 10.2 mm twist drill.

4.3 Regression analyses

In order to define which analysis functions work best for drill wear monitoring, regression analysis was used for parameters calculated using both statistical and FFT based signal analysis methods. The four regression analysis functions (1st, 2nd and 3rd order polynomials and a logarithmic function based on the idea developed and reported by Jantunen & Poikonen [1993]) used to rank the signal analysis results are described in publication II. The idea behind the ranking of monitoring parameters was that the coefficient of determination calculated in the regression analysis could be used to define the ranking order of the measuring signals, analysis functions and parameters. Of all the measured signals the best results were gained with horizontal vibration. However, it can be said that the difference is not big and other measuring signals such as sound, force and acoustic emission also worked well. A more detailed discussion of the applicability of various monitoring methods is given in publication III. In publication II the conclusion is that for practical purposes it could be beneficial to use more than one measuring method in order to get rid of false alarms.

The development of a higher order polynomial regression function with a limited number of terms is described in detail in publications VI and VII. The principle of why a regression analysis technique can be expected to help in monitoring and diagnosis of drill wear is explained in detail in publication VI. Basically the idea is simply to mimic the development of the wear curve, which in the case of tools typically develops exponentially towards the end of the tool life. A higher order polynomial regression function with a limited number of terms is defined in its general form by the following equation:

$$y(t) = a \cdot t^{e} + b \cdot t^{f} + c \cdot t^{g} + d \tag{1}$$

where y(t) is the monitored parameter as a function of time. The parameter can be either a statistical time domain parameter such as root mean square (rms) value or an amplitude value at a specific frequency if FFT has been used. In the equation a, b and c are regression coefficients and t is time. The exponents e, f and g define the degree of the function and there is also a constant d in the function. With a proper choice of exponents e, f and g Equation 1 can also be used to define the 1st, 2nd and 3rd order polynomials (with the 3rd order d also becomes a regression coefficient). As shown in publication VI Equation 1 also mimics quite closely the behaviour of the logarithmic regression function, with the difference that with Equation 1 the total lifetime of the drill does not need to be known. The principles of the solution for regression coefficients can be found e.g. in the book by Milton & Arnold [1995].

For emphasising the most recent data, a factor to be used when calculating the summary terms in regression analysis is introduced:

$$p_i = q^{(n-i)} \tag{2}$$

where n is the current total number of samples, i is the index in the calculation of the summary terms, and q is a constant that defines how much weight the earlier terms are given when all the terms in the calculation of the summary terms are multiplied by p. The most important reason for the introduction of factor q is that regression analysis functions tend to become very stable, i.e. they do not react to current data very rapidly if they have been used for some time with similar data. This lack of response is contradictory to what was presented in chapter 2 concerning the rapid development of wear towards the end of the tool life, hence the introduction of factor q is needed.

Figure 2 shows the same data as in Figure 1, analysed using a higher order polynomial regression function with the following parameter values: e = 9, f = 6, g = 3, d =1 and q = 0.99.

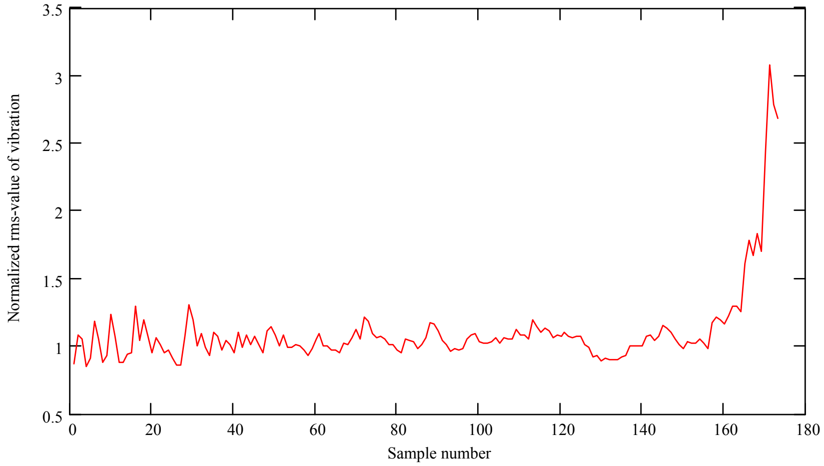


Figure 2. Normalised rms value of vibration (perpendicular to drill axis) for a 10.2 mm twist drill analysed using a higher order polynomial regression function.

5. Simulation model

The development of artificial drilling forces that are influenced by drill wear and a simplified dynamic model that can be used for producing vibration simulation data are described in detail in publication IV. The purpose of developing this simple model was to get a better understanding of the dynamics that influence the drilling process, especially what could happen when a drill is worn. The simulated signals can be used in the testing and training of automatic diagnosis tools. The drilling force model is not supposed to predict the absolute level of drilling forces correctly, consequently it is not of use in the adjustment of machining processes.

5.1 Drilling force model

The artificial drilling force model was developed for calculation of the horizontal drilling force, i.e. the force perpendicular to the axis of the drill. This is also known as the drift force. In principle it should be zero when there are two cutting lips in a drill, since these cancel the influence of each other. However, for a number of reasons the force is not zero in practise, as discussed in chapter 2 of this thesis. As explained in chapter 2, Yang et al. [2002] have treated the dynamics and especially the horizontal vibration of a drill due to the imbalance of forces in their model, which gave the idea for the development of the following simplified model, which is described in more detail in publication IV.

The developed model tries to introduce excitation forces perpendicular to the drill axis at frequencies which might be seen in reality, and also a term is introduced which is a function of drill wear. The simplified horizontal force is calculated according to the following formula:

$$F_x(t) = F_{rpm1}(t) + F_{rpm2}(t) + F_{nrpm}(t) + F_{rnd}(t) + F_0(t) \tag{3}$$

The first two terms in the formula, F_{rpm1} and F_{rpm2}, try to take into account the possible geometrical differences between the two cutting lips and are defined as follows:

$$F_{rpm1}(t) = F_{dp}(t) \cdot \left[c_1 - c_2 \cdot Ln\left(1 - \frac{t}{t_c}\right)\right] \cdot \cos\left[2 \cdot \pi \cdot \omega \cdot t + \phi_{ge} + \phi_{wd} \cdot \sin\left(\omega \frac{t}{c_3}\right)\right] \tag{4}$$

$$F_{rpm2}(t) = F_{dp}(t) \cdot \left[c_1 - c_2 \cdot Ln\left(1 - \frac{t}{t_c}\right)\right] \cdot \cos\left\{ 2 \cdot \pi \cdot \omega \cdot t + \pi \cdot \left[1 + \phi_{wd} \cdot \sin\left(\omega \frac{t}{c_4}\right)\right]\right\} \quad (5)$$

where c1 … c4 are constants, t_c is the total lifetime of the drill, ω is the angular speed of rotation, ϕ_{ge} is the angular geometrical error due to the tolerance in manufacturing the drills, ϕ_{wd} is the difference in wear of the two cutting lips of the drill and F_{dp} is a drilling process force that scales the size of the forces and is defined as follows:

$$F_{dp}(t) = c_5 \cdot H_B \cdot f \cdot F_{dh}(t) \quad (6)$$

where c_5 is a constant, H_B is the Brinell hardness of the work piece material and f is the feed per revolution. The influence of the work piece hardness and the feed follows the statistical model presented by Subramanian & Cook [1977]. However, two terms that take into account the influence of the geometry and wear have been left out since the model described here does not try to predict the cutting forces. It should be noted that the statistical model [Subramanian & Cook 1977] deals with torque and thrust force and the model in this study deals with the horizontal drilling force, which can be estimated to be a function of the thrust and torque [e.g. Yang et al. 2002].

Here the term F_{dh} takes into account the unstable nature of the drilling process, i.e. in the beginning the forces increase when a hole is started, reaching a stable level when the cutting lips of the drill have fully reached the work piece material. F_{dh} is defined as follows:

$$F_{dh}(t) = \frac{t - i \cdot t_d}{\frac{t_d}{b_1}} \qquad \text{if} \qquad i \cdot t_d \le t < i \cdot t_d + \frac{t_d}{b_1} \quad (7)$$

$$F_{dh}(t) = 1 \qquad \text{if} \qquad i \cdot t_d + \frac{t_d}{b_1} \le t \le i \cdot t_d + t_d \quad (8)$$

where t is time, i is a counter for the hole number, t_d is the time it takes to drill one hole and b_1 is a coefficient that defines the relation between the increasing part and the stable part of the thrust force.

The term F_{nrpm} is supposed to describe a number of harmonic components that are multiples of the drilling speed and that can originate from such sources as the bearings and the electric motors of the machine tool in question:

$$F_{nrpm}(t) = \sum_{n=3}^{11} \left\{ F_{dp}(t) \cdot \left[\frac{c_6}{n} - \frac{c_7}{n} \cdot Ln\left(1 - \frac{t}{t_c}\right) \right] \cdot \cos(n \cdot 2 \cdot \pi \cdot \omega \cdot t) \right\} \tag{9}$$

where c6 and c7 are constants, n defines the order of the harmonic component, $F_{dp}(t)$, ω and t_c as defined above.

In order to make the simulation produce signals that also contain random noise, the term F_{rnd} is introduced:

$$F_{rnd}(t) = rnd(c_8) - \frac{c_8}{2} \tag{10}$$

where c_8 is a constant and rnd denotes the MathCad program function [Mathsoft 2002] that produces an equally distributed random number between 0 and c_8.

One phenomenon that can quite clearly be seen and understood is the influence of vibration on the drilling forces, i.e. since the drill is vibrating perpendicular to its axis the drilling forces are also a function of this. The phenomenon can be seen, for example, in the paper by Yang et al. [2002]. The influence of vibration at the natural frequency of the drill is taken into account by the term F_0.

$$F_o(t) = \cos(2 \cdot \pi \cdot f_o \cdot t) \cdot F_{dp}(t) \cdot \left[c_9 - c_{10} \cdot Ln\left(1 - \frac{t}{t_c}\right) \right] \tag{11}$$

where c_9 and c_{10} are constants, t_c is the total tool lifetime, F_{dp} is the drilling force as defined above, and f_o is the first natural frequency of the drill and tool holder calculated using the following formula [Thomson 1972]:

$$f_o = \frac{1}{2 \cdot \pi} \cdot \sqrt{\frac{k}{m}} \tag{12}$$

where m is the mass of the drill and tool holder, and k is the stiffness of the structure. Assuming the drill is a straight round bar that is fixed at one end, the formulae for calculating the natural bending and torsional frequencies can be found e.g. in the book by Young [1989]. For bending, the formula for natural frequency can be written in the following way:

$$f_n = \frac{K_n}{2 \cdot \pi} \sqrt{\frac{E \cdot I}{\rho \cdot S \cdot l^4}} \tag{13}$$

where K_n is a coefficient that depends on the vibration mode, E is Young's modulus, I is the moment of inertia, ρ is the density of the material, S is the cross-sectional area and l is the length of the drill. Making assumptions about the effective diameter and length of a drill, the influence of the drill diameter on the natural frequency can be calculated according to Equation 13. Figure 3 shows the approximate frequency of the first and second bending modes together with the first rotational natural frequency of a drill as a function of drill diameter. As Figure 3 shows, there is a strong dependency of the drill diameter, i.e. the smaller the drill diameter is the higher is the natural frequency. The calculation formula for the natural frequency of the torsional vibration mode is also found in the book by Young [1989]. It should be noted that the torsional natural frequencies are quite a lot higher (more than 10 times) than those of the lowest bending modes.

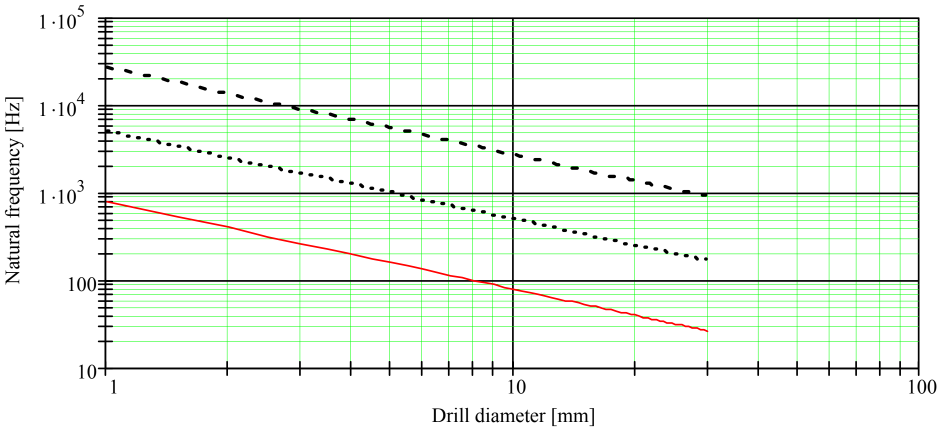


Figure 3. Approximate frequency of the first (lowest line) and second (intermediate line) natural bending vibration modes together with the first (highest line) rotational natural frequency of a drill as a function of drill diameter.

All of the terms described above, except for the random force, are described to some degree as functions of a term, which could be called the wear influence shown in the following formula:

$$h = Ln\left(1 - \frac{t}{t_c}\right) \tag{14}$$

where t is the time and t_c is the total lifetime of the drill. All the terms except for the random term in equation 3 are scaled by a term that takes into account the Brinell hardness of the work piece and the feed of the drill. The influence of drilling separate holes is also included in these terms, i.e. when a new hole is started the forces start from zero again except for the random term.

It is quite apparent in the above development of the simplified simulation model that the model is not a physical one that could correctly predict the horizontal forces in a drilling process. There are many constants in the formulae which were chosen by trial and error when judging the predictions. However, the model can easily be used for producing test data for the development of a diagnostic approach for the automatic diagnosis of drill wear. Based on literature references the model includes terms that could be expected to influence the drilling process but their size as such and relation to each other has no justification through testing.

5.2 Dynamic model

The simplified dynamic model has been developed following the principles presented by Yang et al. [2002]. In the model it is assumed that the tool and tool holder can be modelled as a beam that is rigidly supported at one end and that the excitation force influences at the other end. In their approach Yang et al. [2002] performed the study with two degrees of freedom, i.e. with two differential equations which gave the basis for the iterative calculation of the excitation force. In the present study a model with only one degree of freedom is used and the excitation force is assumed to take into account the influence of the rotating route that the drill travels in the hole during the drilling process. The following basic differential equation describes the dynamic model [Thomson 1972 and Yang et al. 2002]:

$$m \cdot x'' + c \cdot x' + k \cdot x = F_x(t) \tag{15}$$

where m is the mass of the vibrating tool and tool holder, c is the damping, k is the stiffness, and $F_x(t)$ is the dynamic horizontal drilling force as defined in the previous chapter.

Figure 4 shows an example of the calculated vibration acceleration response, together with the excitation force for holes two, three and four. Figure 5 shows the corresponding acceleration response together with the excitation force for the last three holes when the drill was defined as having broken right after the 60th hole. In the examples the following values of input parameters have been used: c = 1.21 Ns/m, c_1 = 20, c_2 = 400, c_3 = 2, c_4 = 1.7, c_5 = 1, c_6 = 0.04, c_7 = 0.08, c_8 = 0.5, c_9 = 0.02, c_{10} = 0.04, b_1 = 4, f = 0.2 mm/rev, f_o = 84.539 Hz, k = 395 N/mm, m = 1.4 kg, t_c = 240.001 s, t_d = 4 s, ϕ_{ge} = 0.00013 rad, ϕ_{dw} = 0.00027 rad and ω = 10 rad/s. The mass, damping and stiffness in this example are the same as in the example given by Yang et al. [2002]. The calculated standard deviation of the vibration acceleration during the simulated drilling of the last hole is about seven times that during the drilling of the first holes.

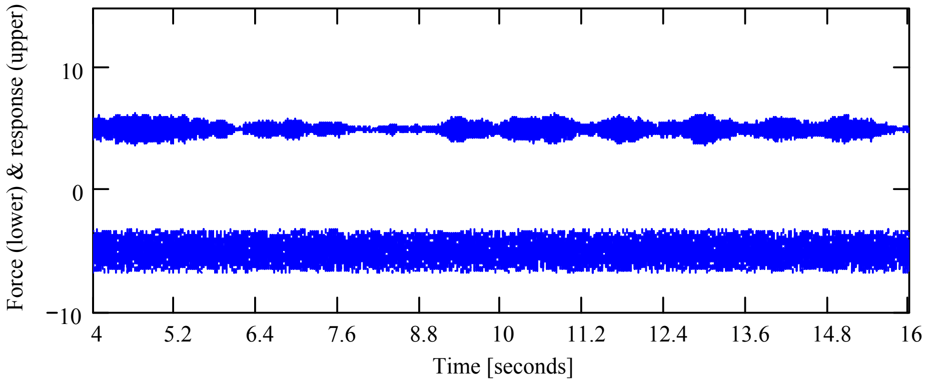


Figure 4. Excitation force (lower curve) and vibration response (upper curve) for holes two, three and four.

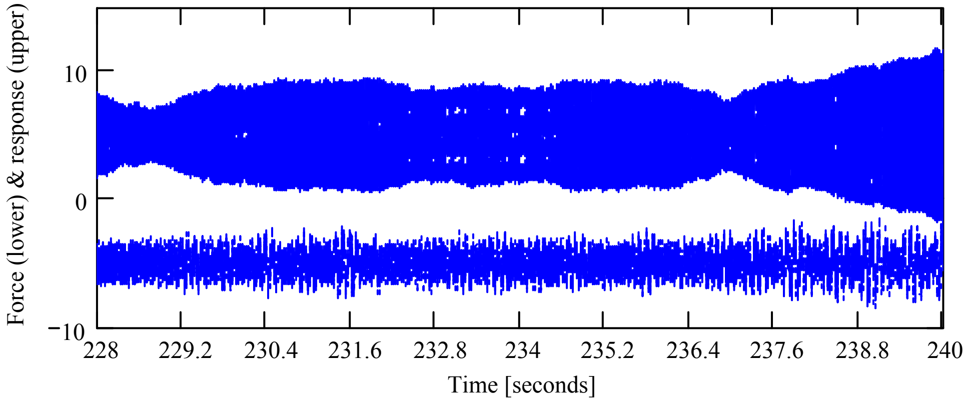


Figure 5. Excitation force (lower curve) and vibration response (upper curve) for the last three holes.

6. Diagnosis of tool wear

In order to enable the unmanned use of machine tools and flexible manufacturing systems, the diagnosis of tool wear needs to be made automatic. In practice this means that some kind of artificial intelligence is needed. Also important is easy configuration for a range of machine tools. The principles of an expert system based approach are described in detail in publication V. The advantages of regression analysis are discussed in publication VI. In publication VII, regression analysis techniques are combined with fuzzy logic. The possibilities of a hierarchical neuro-fuzzy approach that combines information from various sources are described in publication VIII.

6.1 Expert system

Assuming that the diagnosis of drill wear can be based on diagnostic rules such as: "If the amplitude of some parameter increases beyond a predefined limit the drill is worn," it is possible to build rule based expert systems that can be used for the diagnosis of drill wear. The main practical problem with this kind of an approach is the time it takes to describe all the rules. For example, if there is variation in the measuring signals and parameters used for diagnosing wear, a lot of work is needed to redefine the expert system for the specific case it will be used in, or if a generic system is developed it will be very complicated. In the developed approach the basic idea is to use a fault tree database interface program for defining the faults to be monitored, such as drill wear, and describe the corresponding condition monitoring methods (symptoms) using a symptom tree database interface program. After defining the fault and corresponding symptoms that can be used to diagnose the fault, the user starts a rule synthesiser program. The rule synthesiser translates the contents of the fault and the symptom databases into an expert system rule code for the computer performing the monitoring task. In this automatic code writing process, the rule synthesiser takes one page at a time from the symptom tree and from it writes a module onto the expert system code. The procedure is shown in Figure 6.

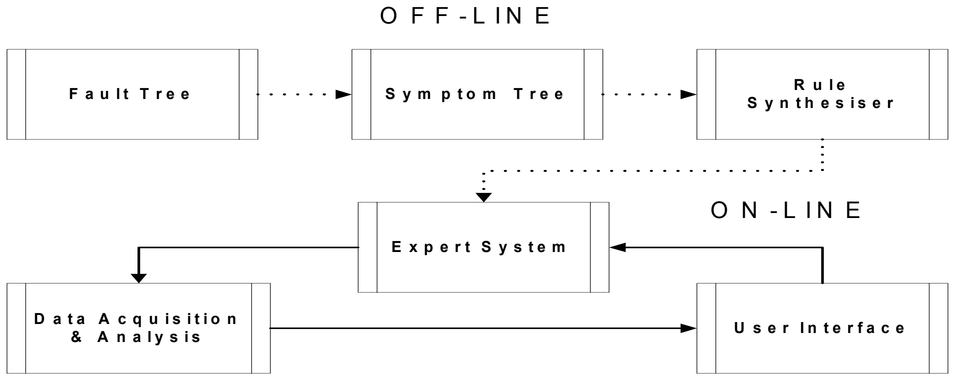


Figure 6. Principles of the new approach to expert system rule generation.

6.1.1 Fault tree

In the fault tree user interface window and the corresponding table of the database, the machine tool is defined by a chain of subcomponents. The approach is general, i.e. it can handle various types of faults and also various types of rotating machinery. In the case of drill wear, the subcomponent chain is defined as follows: component = machine tool, subcomponent = spindle, subsubcomponent = tool holder, subsubsubcomponent = drill. For this chain 18 different types of fault can be defined, e.g. such as worn out. Since the fault tree is part of a database, all the typical features of a relational database program such as find, copy etc. are available.

6.1.2 Symptom tree

In the following step of building an expert system for each of the faults, a symptom tree database definition is performed. The definition of symptoms that define a fault include the following: status of the machine tool (power on, hydraulics on etc.), machining information (spindle rotating, machining etc.) and condition monitoring information (signal, sensor, time criticality, analysis method, averaging, alarm limit etc.). The definitions include all the necessary information for defining the data collection through an AD card and also all the necessary information for performing signal analysis using a collection of mathematical subroutines. When FFT is used to calculate e.g. the power spectrum or other analysis functions, only the so-called cursor values, i.e. the 20 highest peaks of the analysed functions, are saved to keep the size of the database reasonable. Again all the features of a typical relational database are available. Since the above definitions are done for each tool type included in the wear monitoring program, the editing functions are important to make the amount of work manageable.

6.1.3 Rule synthesiser

The idea of the rule synthesiser is to automate the laborious writing of expert rules for different types of machine tools using a variety of tools. In principle, all the necessary information is saved in the fault and symptom tree database tables. The rule synthesiser takes the information from the symptom tree database table and automatically generates the computer program code containing the needed expert system rules. The rule synthesiser works by processing each rule specification in the symptom tree database, then breaking each rule into several function calls. The rule synthesiser also builds the links between these function calls in a logical order so that the data can go through the steps of data acquisition, signal processing, feature extraction and testing against the specified limits. In addition, the rule synthesiser automatically combines rules into groups corresponding to each fault defined in the fault tree, e.g. all the rules needed to detect a worn-out drill of a specific size e.g. 10.2 mm.

6.1.4 Fault manager

The purpose of the fault manager module is to combine the information based on various sensors and analysis functions into the final conclusion. Typically in a cutting process there are a number of changes taking place in the measured signals. These can be due to changes in the cutting parameters or variation in the work piece material etc. In order to handle this it is suggested that a number of measuring signals and analysis functions are used. The rule synthesiser can build the rules for each of these features used in the expert tool. Development of the analysed features with time is saved using regression analysis techniques, thus only the summary terms of the regression functions need to be saved. The fault manager then follows these features and their reliability based on the coefficient of determination of the regression analysis functions, and calculates the sum of the coefficients of determination of those analysis functions that have triggered the predefined threshold limit. The final conclusion of whether a tool is worn is then based on comparison of the sum of coefficients of determination.

6.2 Fuzzy classifier

Fuzzy classification is one possible way to automate the diagnosis of tool wear as described in chapter 2 of this thesis. The development of the approach of using simplified fuzzy classification following the principles shown by Rao & Rao [1993] in the diagnosis of drill wear is explained in publication VII. The idea is that in the beginning, when a tool is in good condition, some of the early data is used for defining the fuzzy classification limits for the analysed parameters of the monitored signals. In the developed approach the number of classes has been limited to eight, class two meaning that the tool is in good condition and class eight that it is completely worn. Class one has been reserved for lower values of the monitored parameter, which possibly mean that the cutting conditions are different from those when the limits were defined.

The classes are defined using the mean and standard deviation of the measured signal. These statistical parameters are typically used when so-called health indexes are calculated [Williams et al. 1994] or alarm limits are defined in condition monitoring standards such as the PSK 5705 Standard [2004]. In the developed approach the classes are defined using the following definitions: The mean value of each class (class index i = 1..8) is defined according to the following formula:

$$ClassMean_i = (i-2) \cdot j \cdot \sigma + \mu \tag{16}$$

where j is a coefficient defining the size of the classes, k is a coefficient that defines the shape of the classes, and μ is the mean value and σ the standard deviation of the first measured parameters. The upper and lower limits of the classes are defined as follows:

$$LowLow_i = ClassMean_i - j \cdot (1+k) \cdot \sigma / 2 \tag{17}$$

$$LowHigh_i = ClassMean_i - j \cdot (1-k) \cdot \sigma / 2 \tag{18}$$

$$HighLow_i = ClassMean_i + j \cdot (1-k) \cdot \sigma / 2 \tag{19}$$

$$HighHigh_i = ClassMean_i + j \cdot (1+k) \cdot \sigma / 2 \tag{20}$$

Figure 7 shows an example of fuzzy classification of the rms value of vibration. In this example the basic signal is the same as that used in the analysis of data in Figure 1 and Figure 2. In the example, the 20 first values analysed have been used for defining the mean and standard deviation in the above equations. The values used are j = 1 and k = 0.5.

The results of fuzzy classification can be used as input for a neural network as shown in the following chapter. The use of fuzzy logic in pre-processing the input data for a neural network follows the principles presented by Rao & Rao [1993].

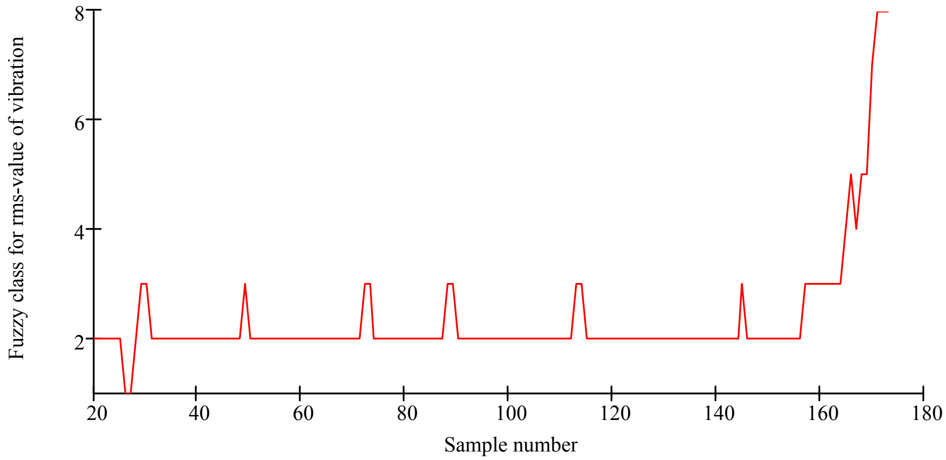


Figure 7. A result of fuzzy classification of the rms value of vibration.

6.3 Hierarchy

Publication VIII describes the principles of building a flexible hierarchical neuro-fuzzy system for prognosis. The basic idea is simply to use a hierarchy, i.e. to build a bigger and a more complicated model using sub-models, as seen in Figure 8. In the most simplified level a higher level conclusion is drawn based on a number of monitoring parameters analysed. In this approach the maximum number of parameters in a sub-model is limited to eight, i.e. the conclusion at the lowest level is based on eight parameters. The choice of eight as the maximum is based on numerical and logical reasons. It is relatively easy to handle models of this size and eight is a multiple of two, which can be handled with three bits. At sub-model level the idea is to define the condition of the monitored tool or, more generally, the condition of a machinery part.

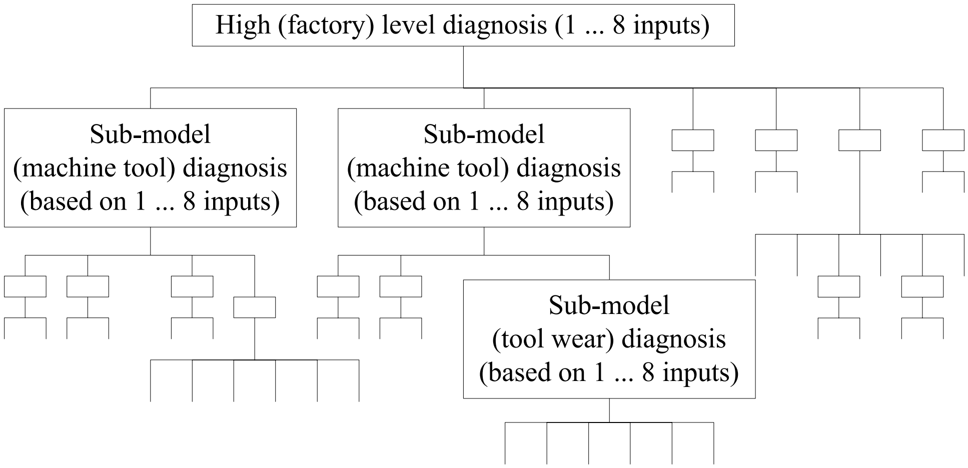


Figure 8. Structure of the hierarchical neuro-fuzzy system.

Table 2 shows the principal idea of the hierarchical approach. In case of drill wear monitoring, typically two measuring signals such as vibration and acoustic emission could be used. If four statistical parameters are calculated from these two signals, this actually fills the lowest level sub-model and should be the basis for defining that the drill is worn. At any level of the hierarchical approach the decision making process is similar, since at the maximum there are always eight inputs and only one output representing the conclusion. In the case of drill wear monitoring, it is important also to be able to define when drilling is taking place so that the signal analysis is only carried out when it is relevant to do so. In the case of flexible manufacturing systems, the hierarchical model could be used in such a way that various tool types could have various sub-models and also the condition of the machine tool could be followed using some sub-models dedicated to various types of relevant wear models, such as spindle bearings and tool changers etc. However, this is beyond the scope of this thesis.

Table 2. Principles of the hierarchical approach.

Model level	**Fault**	**Parameters to handle**	**Technique for classification**	**Technique for conclusion**
Low level (linked to wear of component which can be monitored)	E.g. drill wear (bearing fault etc.)	Parameters analysed from monitoring signals	Fuzzy logic based on values of mean and standard deviation, allows manual interpretation	Simple logic: highest wins, if two signals indicate, etc., Could be neural net if statistical data for training
Machine level(s), possible to handle various process states, different tool types etc.	Machine needs maintenance, tool change etc.	Conclusion from lower level	Values from lower level	Usually logical, corresponds to rule based approach
High level (factory level)	Production can be followed (are there any faults)	Conclusion from lower level	Values from lower level	Usually logical

7. Discussion

7.1 Measured signals and signal analysis

In the reported tests and in the available literature, which is covered in chapter 2 of this thesis, there is a lot of variation in how well different measuring signals and analysis techniques have worked. This is due to the wide variety of factors influencing the results. The size of drills and the work material together with the drilling process parameters influence the measured signals. For example, it is easy to understand that the smaller the drill, the smaller the forces and the higher the frequencies are at which one could expect the greatest variation to take place. This tendency can be well understood in the light of dynamic simulation. The natural frequencies of a drill and the drill holder increase with a decrease in the diameter of the drill. This then actually means that a combination of measuring and signal analysis techniques that works well with a certain size of drills does not work as well with others.

Feed force and torque have been used a lot in laboratory tests with drills and some good results have been reported. However, it is difficult to measure forces at very high frequencies and this is one reason why good results with small drills have been reported with measuring techniques such as high frequency (ultrasonic) vibration measurements capable of sensing these higher frequencies. Also when smaller drills are used this influences the analysis techniques that should be used. These should be simple and quick enough to react to the quick changes and they should not be too demanding on the analysis equipment. As a consequence, measuring motor currents works better with large drill diameters because the drilling forces are higher; their portion of the total signal is then higher and also the measuring chain might be able to react quickly enough, but when the drill sizes are smaller the opposite is true.

The results reported in publications II and III apply to drills of moderate size, i.e. about 5 mm and more in diameter for the reasons stated above. The measuring equipment and signal analysis techniques that have been used with dynamic signals such as force, vibration and sound cover the frequencies of interest, i.e. rotational frequency and the lowest natural frequencies of the tested drills. However, already the somewhat more complex analysis based on FFT was occasionally rather slow with the equipment being used to analyse the test signals. It could be claimed that although the amplitudes at certain frequencies gave a better indication of drill wear, there is a risk that since the drills in these tests (as in many tests reported in the literature) wear very quickly at the end of their life, this phenomenon could be missed between analysis rounds.

In the literature even more complicated approaches than FFT, such as autoregressive modelling [Radhakrishnan & Wu 1981], are suggested for diagnosing drill wear. It would seem that this kind of method is not very generic, i.e. the models work for one specific drill size and work piece material, but they would need to be trained for new combinations and this would be very time consuming and laborious, even though it would apparently work in a fixed case.

Publication II lists the best measuring and analysis functions for drill wear monitoring. Vibration, acoustic emission, sound and some of the force measuring techniques were the best methods. Publication III shows good examples of vibration, acoustic emission and sound measurements analysed in the time domain. All of these measuring techniques can be considered acceptable for on-line use in a real production environment, in the sense that the necessary sensors can be mounted relatively easily to a typical machine tool and they do not influence the production. In publication II, statistical time domain parameters such as root mean square, mean deviation and maximum were listed as the best in drilling tests. However, as explained above, especially the drill size has a great influence on what would be the optimum measuring arrangement and signal analysis technique, thus the results shown in this thesis should not be over-generalized. To overcome the challenges brought about by drill size, it is suggested that the smaller the drills are, the higher the frequencies should be that are included in the measuring and analysis chain. Also, since the proportion of signals from the drill decreases with decreasing drill size, it raises the question of how close to the drill the sensors should be able to measure. In other words, the closer to the drill the sensors can measure, the higher the proportion of the drill signal is of all the signals that the sensor can measure. In order to overcome this problem of low signal levels with smaller drills more sophisticated signals analysis might be needed than is the case with medium and bigger size drills which introduce higher signal levels from drilling.

The higher order statistical parameters such as kurtosis and skewness were especially sensitive to variation in the tests, therefore they were not as good as the above-mentioned parameters. It is logical that the minimum value is not a very good parameter for drill wear monitoring, because the lowest values in the measuring signals resulted from some disturbance in the measuring procedure.

7.2 Simulation model

In theory, assuming the static drilling forces can be calculated as explained in chapter 2, and knowing how the cutting forces introduce wear into the drill, and also knowing how the drill dynamics influence the cutting forces and vice versa, it would be possible to build a dynamic drill wear model. This kind of model would also need to have probabilistic features in order to introduce differences between the cutting lips, which is one of the important factors that influence the vibration response of a worn drill. As stated earlier, this type of model does not seem to exist today and the simulation studies presented in publication IV are very far removed from this kind of approach.

The approach suggested in publication IV and covered briefly in chapter 5 of this thesis simply tries to show and test the possible influence of various artificial dynamic loads, which would increase with a similar trend to that seen in laboratory tests, and then to hide this trend behind noise and see how the used analysis functions work in this type of scenario. The model presented by Yang et al. [2002] is much cleverer in the way it calculates real forces and torque, taking into account the dynamic influence caused by the fact that drills do not drill straight but vibrate and consequently move from one edge to the other. However, the model only vibrates if it is given an initial push from equilibrium, and the only dynamic influence taken into account is then vibration due to the natural vibration modes of the drill and the unstable forces introduced by this vibration. In publication IV a number of dynamic excitation forces are introduced into the model; these are not derived from laboratory tests or theory, but are the results of a trial and error approach in the sense that with a suitable combination of parameters and logically chosen excitation forces, the final result resembles that seen in the tests when vibration is considered. It is also important to remember that the influence of wear has been introduced into the excitation forces as a function of the term defined by Equation 14, and consequently this term defines the influence of wear throughout the simulation model. The simulation is also very limited in the sense that it could be expected that different types of wear, e.g. chisel, corner, flank and margin wear, introduce different kinds of vibration spectra [El-Wardany et al. 1996], but this is not covered at all in the model.

It should also be remembered that when the simulation model described in chapter 5 and publication IV is limited to the first radial vibration mode, Rotberg et al. [1990] point out on the basis of measurements that the most important vibration mode in drill wear monitoring is the torsional vibration mode coupled with the axial vibration mode. The natural frequencies of these vibration modes are higher than for the radial modes. Also when studying the drilling process it is somewhat unclear how much support the drilled hole actually gives in a radial direction when there is no support in the torsional direction. However, in principle the situation seems to be similar for all of the vibration modes. Wear introduces higher dynamic loads and consequently the vibration increases at the first natural modes (the second mode in a radial direction might be more easily excited than the first, due to the supporting effect of the hole) in all possible directions. This means that the behaviour could be expected to be similar in all directions, and in fact in reality all of these vibration modes are combined. Although the calculation procedure is similar, it becomes more demanding the higher the natural frequencies are, and in this sense the modelling in the radial direction is easiest to perform. Again the findings presented by Rotberg et al. [1990] point out how far from reality the simulation described in chapter 5 really is, although it is claimed that the principles and the trends could be similar in reality as are the indications in measured parameters.

It should be noted that the simplified simulation shown in publication IV with MathCad [Mathsoft 2002] takes about an order of magnitude longer than the wear process of a typical twist drill because of the high frequency range. It could be deduced that the introduction of a real drill geometry by performing the calculation over a number of sections would multiply the calculation effort by hundreds if not thousands.

With this kind of simplified model, with a one-degree-of-freedom model the vibration at the natural frequency is very dominant. However, this tendency of some frequencies to dominate the spectrum is similar to what was measured in the reported tests, and in some cases this phenomenon is used in signal analysis using band pass filtering [e.g. Kutzner & Schehl 1988]. It could also be claimed that the simulation model supports the idea that the measuring technique and analysis function used should be able to handle the frequency range where the torsional and radial natural frequencies of a drill installed in a drill holder lie. This simulation model supports what was said in the previous chapter about the influence of drill size. With small drill diameters the frequency range goes beyond the capabilities of normal vibration measuring equipment, i.e. the frequencies for a drill with a 1 mm diameter might be of the order of 25 kHz [Kutzner & Schehl 1988]. The simulation model also supports what has been claimed about the best statistical indicators of tool wear, but this proof should be treated as uncertain because the input, i.e. forces introduced into the model, certainly have an affect what the produced signal looks like.

7.3 Regression analysis

As pointed out in chapter 2 there are a number of problems related to automatic diagnosis of tool wear in practice. The measured signals are noisy because of the nature of the cutting process and there may be sharp peaks in the signal, which may not indicate anything. The absolute values of the analysis parameters are usually not meaningful because there is so much variation due to the variation in tool size, the cutting parameters, work piece material etc. Instead it is important to notice the trend in the parameters analysed. However, this could mean that a lot of information would need to be saved. The use of the higher order polynomial regression function with a limited number of terms as described in chapter 4 and in more detail in publications VI and VII provides a solution to the problems described above: The higher order polynomial regression function smoothens sudden individual peaks and picks up the trend in the analysed parameter. Since the regression function mimics the shape of wear development, the function can also be used to give a prognosis of the upcoming tool failure. When regression functions are used, the trend in a signal is saved. One of the benefits of regression functions is that in order to save the information they contain, only a very limited number (nine) of summary terms need to be saved.

There are also possible drawbacks related to regression functions. One is that they may be slow to react to changes if a stable situation has continued for a long time. In the proposed approach, the idea behind introducing a weighting term is to solve this problem and keep regression functions quick enough to respond. Another factor that influences this is the order of the function, and for this reason the use of relatively high order functions is suggested. A drawback of higher order polynomial functions is that they may behave in very strange ways, i.e. they tend to become unstable with noisy data. The use of a limited number of terms helps in this respect because with this limitation the functions actually behave like a third order function, with the difference that now the changes can take place more rapidly.

It could be argued that higher order polynomial regression functions tend to increase the relative error. However, this is not really linked to the higher order polynomial functions but rather to the nature of the problem. Wear tends to develop very quickly towards the end of the tool life so there is no way of avoiding this, i.e. any prediction technique/function would suffer from the same problem of the relative error increasing. The use of the weighting function, i.e. that the current data is emphasized at the cost of older data, provides some help in this respect and makes the prognosis more reliable than if all the data had equal weight. Introduction of the weighting function can in some cases also make it possible for the approach to adapt to small changes caused by a change of cutting parameters. However, this is something that should be tested more thoroughly. The polynomial regression function does filter out some of the unwanted variation of the measured parameters, i.e. short peaks due to noise in the signals, and in this way makes the analysis more robust which is important in a machining environment. Naturally, if smoothing of the time-series data had been the sole target of the data manipulation, a much more simplified function would have been available, such as that described by Williams et al. [1994]. Their study gives examples of the use of moving average or exponential smoothing in condition monitoring. The biggest difference between the approach suggested in this thesis and those very simple methods is that the simple methods do not give a prognosis of the forthcoming trend of the monitored parameter. Due to this restriction, simple smoothing techniques do not react as quickly to changes of the monitored parameter.

The results of publications VI and VII suggest that relatively high values of the parameters of the regression function such as $e = 9$, $f = 6$ and $g = 3$ give good results. Typically q can have a value of e.g. 0.99 if the process is stable with frequent measurements. The lower the value is, the more the last measurements are emphasized. In fact, using lower exponent values such as $e = 3$, $f = 2$ and g =1 together with a lower value of q = 0.9 would give a result quite close to that obtained with the higher exponent values mentioned earlier. However, if there is enough data the higher values result in a regression function that more closely resembles the wear function described in publications VI and VII.

Certainly, instead of the higher order polynomial regression function quite a number of other functions could also be tested. One of the simplest possible solutions would be the exponential function. In simple format this kind of function also includes three unknowns, i.e. the exponent, a parameter used to multiply the exponent term, and a coefficient. Based on the definition an exponent function could be rather sensitive and possibly not as well suited to prognosticating as the polynomial regression function. However, it has not been within the scope of this thesis to widely compare different possible regression functions. It is accepted at this stage that the polynomial function is suitable for the defined task and it is left to further studies to suggest and compare other possible functions.

7.4 Expert system

There are a number of advantages in building an expert system as suggested in chapter 6.1 of this thesis. It is not necessary to write a lot of expert system code manually that could handle a huge number of tools. It is easy to make changes or add information thanks to the practical user interface. Unfortunately there are also disadvantages in this approach. The amount of work is still relatively high and demanding, i.e. the user must know what to do and how to define limits for the various signals, and this need for professional manpower makes the whole approach unpractical for everyday use. In addition the size of the final program will be extensive, but this is not possibly so meaningful today because of the improvement of processing power.

7.5 Fuzzy classification

The simplified fuzzy classification has been introduced into the approach in order to make diagnosis of tool wear automatic. The same approach can be applied using both fuzzy limits and crisp limits. In both cases the conclusion can be shown in eight classes and it can be argued whether the simplified use of fuzzy limits actually brings any benefits. One argument is that in reality the limits are fuzzy, and thus the use of fuzzy limits is closer to reality. Another argument is that the use of fuzzy limits could make the following step more robust if neural nets were used. The reason for this is that the use of fuzzy limits brings some variation to the inputs of the neural net.

The diagnosis examples shown in Publication VII are based on the use of two measuring signals, i.e. vibration and acoustic emission, and the final conclusion of drill wear is in most cases based on the simple rule that at least two parameters must give an indication of drill wear. In publication VII different parameter values and principles in making the final conclusion are tested. The conclusion is that relying on more than one statistical parameter makes the diagnosis more stable, and that conservative values (small values of j) should be preferred when the fuzzy limits are defined. The use of small values of j actually means that the upcoming tool failure is seen too early rather than too late. It should be noted that there is a remarkable difference in using fuzzy classification in such a simplified manner as was done in publication VII, compared to that shown e.g. in the paper by Du et al. [1995]. The more sophisticated (normal) way of using fuzzy limits could reveal a much improved connection between the various parameters and improve the reliability of the conclusion. However, the problem is that this relationship would have to be trained prior to the use of the approach, which again is a very severe limitation if an automatic approach is the final goal.

7.6 Automatic diagnosis

Many of the approaches that have been developed for tool wear diagnosis and are reported in the literature rely on training and a definition phase in order to work properly. This is also true for the rule based approach described in chapter 6.1 of this thesis. In normal production, the need for training and the definition phase might be very problematic if a great number of tools are used in different machining conditions with varying work piece materials. The following development phase based on the use of regression analysis techniques and fuzzy logic does not suffer from this as much. A number of parameters have to be defined, but when this has been done for the production environment these could be kept the same for a number of tools, and the definition of limits for diagnosing tool wear should take place automatically. There are also limitations to the suggested approach. The first measurements are used for defining the limits, and if the tool fails during that period the diagnosis system does not provide any help. This restriction does not apply in cases where the tool type and cutting parameters are kept constant, i.e. there is historical information of similar cases and thus the same limits can be used that were defined earlier and have proved to work.

Naturally, significant questions related to the suggested approach remain open. Although the approach works with laboratory data from medium and large size drills, does it really work in real life in normal production where the environment is much more demanding? There are external disturbances influencing the signals and there is variation in the work piece materials etc. Is the approach really so easy to define that it attracts users? Will there be too many mistakes in the diagnosis, so that users do not rely on the system? The only way to get answers to the above is to test the system in real production. This has not been done to date, but hopefully the opportunity will come to test and gain experience of the capabilities of the suggested approach in real production.

8. Conclusion

There exists a great potential to improve the machine tool utilisation rate with an advanced condition monitoring system using modern sensor and signal processing techniques. A comprehensive cutting test procedure was carried out with drills. Based on the tests, different measuring methods and analysis techniques together with their benefits and disadvantages have been discussed. Especially vibration measurements and methods that are closely related to it, i.e. sound and acoustic emission, seem to be potential and practical methods that could be recommended for everyday use in production. The importance of natural vibration modes of the drill and tool holder is apparent in the light of tests and the simplified simulation carried out. The use of higher order polynomial regression analysis functions with a limited number of terms is suggested for filtering the measured data and saving it in a compact form, which is especially beneficial when the number of monitored tools is high. An automatic diagnosis approach has been developed based on simplified fuzzy logic. The approach can be linked to a wider context, e.g. monitoring a complete machine tool through the proposed hierarchical structure. However, even though the results with laboratory data are promising, there are no test results from a real production environment. It should also be noted that the current results apply to medium and large size drills, and unfortunately the diagnosis of wear and breakage of small size drills is more demanding. The proposed approach is unable to detect what kind of wear is taking place, i.e. it does not differentiate chisel, corner, crater, flank or land wear from each other.

Based on the research reported in this thesis and the above conclusions, some suggestions can be made for further work:

- First of all, wider testing of the developed approach both in the laboratory and in the industry is suggested. In these tests the benefits of the higher order regression analysis function could be tested more thoroughly, including mathematical optimisation of the order of the function and the emphasis of current data, i.e. the variation of parameter q. These tests could also include testing of the whole automatic diagnosis approach in the prediction of the remaining lifetime of the tool. Furthermore, the tests might also help to widen the scope of the approach so that it could also be used for monitoring small size drills.

- One further step in gaining a better understanding of drill wear monitoring could be the development of a real physical wear model for drill wear. This could include the statistical treatment of material variation both in the drill and in the work piece, leading to a natural variation of the wear of the cutting lips. The model could also aim to differentiate between various wear types. This kind of model would probably have to be built using the finite element method (FEM) for modelling. However, it should be noted that even the very simplified model presented in this thesis could be used more widely in the development of automatic tool wear monitoring, diagnosis and prognosis.
- Assuming that all the above-mentioned testing gave positive results, one further task that would then have to be carried out is the development of an automatic tool monitoring information database for practical and easy handling of the numerous tools in a real production environment.
- Further work could also be done in testing the same approach in diagnosing and predicting the condition of machinery components suffering from a similar type of exponentially increasing wear, such as rolling bearings. Although the first version of the hardware capable of performing all the tasks presented in this thesis has been built, a further version could be developed that would include a better capability of signal amplification and filtering and improved automatic adjustment of parameters.

References

Aatola, S., Heikkinen, H., Hemming, B., Jantunen, E., Jokinen, H. and Poikonen, A. 1994. Värähtelymittauksia laservibrometrillä [Vibration Measurements with a Laser Vibrometer]. Espoo: Technical Research Centre of Finland, VTT Tiedotteita – Research Notes 1613. 61 p. + app. 2 p. ISBN 951-38-4728-4 (In Finnish.)

Abu-Mahfouz, I. 2003. Drilling Wear Detection and Classification using Vibration Signals and Artificial Neural Networks. International Journal of Machine Tools & Manufacture, Vol. 43, pp. 707–720.

Abu-Mahfouz, I. 2005. Drill Flank Wear Estimation Using Supervised Vector Quantization Neural Networks. Neural Computing & Applications, to be published, currently available in Internet.

Adamczyk, Z. 1998. Transient States in Drilling Process as a Source of Tool Wear Knowledge for Intelligent Tool Condition Monitoring System. IX Workshop on Supervising and Diagnostics of Machining Systems Manufacturing Simulation for Industrial Use. Prace Naukowe Instytutu Technologii Maszyn i Automatyzacji Politechniki Wroclawskiej, Vol. 69, No. 31, pp. 195–204.

Barker, R.W., Kluthe, G.A. and Hinich, M.J. 1993. Monitoring Rotating Tool Wear Using Higher-Order Spectral Features. Journal of Engineering for Industry, Vol. 115, Transactions of the ASME, pp. 23–29.

Bhattacharyya, S.K. and Ham, I. 1969. Analysis of Tool Wear – Part I: Theoretical Model of Flank Wear. Journal of Engineering for Industry, Transactions of the ASME, pp. 790–798.

Braun, S. and Lenz, E. 1986. Machine Tool Wear Monitoring. Mechanical Signature Analysis, Theory and Applications. Academic Press Ltd. Pp. 321–342.

Braun, S., Lenz, E. and Wu, C.L. 1982. Signature Analysis Applied to Drilling. Journal Mechanical Design, Vol. 104, Transactions of the ASME, pp. 268–276.

Brinksmeier, E. 1990. Prediction of Tool Fracture in Drilling. Annals of the CIRP, Vol. 39, No. 1, pp. 97–100.

Brophy, B., Kelly, K. and Byrne, G. 2002. AI-based Condition Monitoring of the Drilling Process. Journal of Materials Processing Technology, Vol. 124, pp. 305–310.

Byrne, G., Dornfeld, D., Inasaki, I., Ketteler, G., König ,W. and Teti, R. 1995. Tool Condition Monitoring (TCM) – The Status of Research and Industrial Application. Annals of the CIRP, Vol. 44, No. 2, pp. 541–567.

Chandrasekharan, V. 1996. A Model to Predict the Three-Dimensional Cutting Force System for Drilling with Arbitrary Point Geometry. Ph.D. thesis, University of Illinois at Urbana-Champaign.

Christoffel, K. and Jung, W. 1981. Überwachungseinheit für die Bohrbearbeitung. Industrie Anzeigner, Vol. 103 (62), pp. 198–199.

Cook, N.H. 1980a. Tool Wear Sensors. Wear, Vol. 62, pp. 49–57.

Cook, N.H. 1980b. Prediction of Tool Life and Optimal Machining Conditions. Wear, Vol. 62, pp. 223–231.

Dimla, D.E. Jr., Lister, P.M. and Leighton, N.J. 1997. Neural Network Solutions to the Tool Condition Monitoring Problem in Metal Cutting – A Critical Review of Methods. International Journal of Machine Tools & Manufacture, Vol. 37, No. 9, pp. 1219–1241.

Du, R., Elbestawi, M.A. and Wu, S.M. 1995. Automated Monitoring of Manufacturing Processes, Part 1: Monitoring Methods. Journal of Engineering for Industry, Vol. 117, pp. 121–132.

El-Wardany, T.I., Gao, D. and Elbestawi, M.A. 1996. Tool Condition Monitoring in Drilling Using Vibration Signature Analysis. International Journal of Machine Tools & Manufacture, Vol. 36, 6, pp. 687–711.

Erdélyi, F. and Sántha, C. 1986. Monitoring Tasks on Boring and Milling Production Cells. Computers in Industry, Vol. 7, pp. 65–71.

Ertunc, H.M. and Oysu, C. 2004. Drill Wear Monitoring using Cutting Force Signals. Mechatronics, Vol. 14, pp. 533–548.

Ertunc, H.M., Loparo, K.A. and Ocak, H. 2001. Tool Wear Condition Monitoring in Drilling Operations Using Hidden Markov Models (HMMs). International Journal of Machine Tools & Manufacture, Vol. 41, No. 9, pp. 1363–1384.

Fu, L. and Ling, S.-F. 2002. Neural Network Based On-line Detection of Drill Breakage in Micro Drilling Process. Proceedings of the 9th International Conference on Neural Information Processing (ICONIP'02), Vol. 4, pp. 2054–2058.

Govekar, E. and Grabec, I. 1994. Self-Organizing Neural Network Application to Drill Wear Classification. Journal of Engineering for Industry, Vol. 116, No. 3, Transactions of the ASME, pp. 233–238.

Hayashi, S.R., Thomas, C.E. and Wildes, D.G. 1988. Tool Break Detection by Monitoring Ultrasonic Vibrations. Annals of the CIRP, Vol. 37, No. 1, pp. 61–64.

Hiebert, S.F. and Chinnam, R.B. 2000. Role of Artificial Neural Networks and Wavelets in on-line Reliability Monitoring of Physical Systems. IEEE, pp. 369–374.

Jantunen, E. 2001. A Short List of Some Tool Wear Monitoring Related References from the Late 20th and Early 21st Century. Report AVAL73-011048, Espoo.

Jantunen, E. and Poikonen, A. 1993. Dynamics in Monitoring Gear Faults. Proceedings of the 47th of Mechanical Failures Prevention Group, Virginia Beach, Virginia, USA.

Kanai, M. and Kanda, Y. 1978. Statistical Charasteristics of Drill Wear and Drill Life for the Standardized Performance Tests. Annals of the CIRP, Vol. 27, No. 1, pp. 61–66.

Kim, H.Y., Ahn, J.H., Kim, S.H. and Takata, S. 2002. Real-time Drill Wear Estimation Based on Spindle Motor Power. Journal of Materials Processing Technology, Vol. 124, pp. 267–273.

Kuhmonen, M. 1997. The Effect of Operational Disturbances on Reliability and Operation Time Distribution of NC-machine Tools in FMS. Ph.D.

Thesis, Lappeenranta University of Technology, Research Papers 59. 125 p. ISBN 951-764-162-1

Kutzner, K. and Schehl, U. 1988. Werkzeugüberwachung von Bohrern kleinen Durchmessers mit Körperschallsensoren. Industrie Anzeigner, Vol. 110 (82), pp. 32–33.

König, W. and Christoffel, K. 1980. Sensoren für die Bohrbearbeitung. Industrie Anzeigner, No. 100, Vol. 12, pp. 29–33.

König, W. and Christoffel, K. 1982. Werkzeugüberwachung beim Bohren und Fräsen. Industrie Anzeigner, Vol. 104 (96), pp. 36–38.

König, W., Ketteler, G. and Klumpen, T. 1992. Process Monitoring in Turning, Milling, Drilling and Grinding Processes. Proceedings of the 8th IMEKO International Symposium on Technical Diagnosis, Dresden, Federal Republic of Germany. Pp. 433–446.

König, W., Kutzner, K. and Schehl, U. 1989. Körperschall als Basis der Prozeßüberwachung. Industrie Anzeigner, 11, pp. 18–21.

König, W., Kutzner, K. and Schehl, U. 1992. Tool Monitoring of Small Drills with Acoustic Emission. International Journal of Machine Tools & Manufacture, Vol. 32, No. 4, pp. 487–493.

Lechler, G. 1988. Zur Werkzeugüberwachung beim Bohren und Drehen. VDI-Z, 130, No. 2, pp. 39–41.

Lenz, E., Mayer, J.E. and Lee, D.G. 1978. Investigation in Drilling, Annals of the CIRP. Vol. 27, No. 1, pp. 49–53.

Li, D. and Mathew, J. 1990. Tool Wear and Failure Monitoring Technique for Turning – A Review. International Journal of Machine Tools & Manufacture, Vol. 30, No. 4, pp. 579–598.

Li, G.S., Lau, W.S. and Zhang, Y.Z. 1992. In-Process Drill Wear and Breakage Monitoring for a Machining Centre Based on Cutting Force Parameters. International Journal of Machine Tools & Manufacture, Vol. 32, No. 6, pp. 855–867.

Li, P.G. and Wu, S.M. 1988. Monitoring Drilling Wear States by a Fuzzy Pattern Recognition Technique. Journal of Engineering for Industry, Vol. 110, No. 2, Transactions of the ASME, pp. 297–300.

Li, X. 1999. On-Line Detection of the Breakage of Small Diameter Drills Using Current Signature Wavelet Transform. International Journal of Machine Tools & Manufacture, Vol. 39, Issue 1, pp. 157–164.

Li, X. and Tso, S.K. 1999. Drill Wear Monitoring Based on Current Signals. Wear, Vol. 231, pp. 172–178.

Li, X., Dong, S. and Venuvinod, P.K. 2000. Hybrid Learning for Tool Wear Monitoring. The International Journal of Advanced Manufacturing Technology, Vol. 16, pp. 303–307.

Lin, S.C. and Ting, C.J. 1995. Tool Wear Monitoring in Drilling Using Force Signals. Wear, 180 (1–2), pp. 53–60.

Liu, H.S., Lee, B.Y. and Tarng, Y.S. 2000. In-process Prediction of Corner Wear in Drilling Operations. Journal of Materials Processing Technology, Vol. 101, pp. 152–158.

Liu, T.I. 1987. Development of Crankshaft Multifacet Drills and On-Line Drill Wear Monitoring. Ph.D. Thesis, University of Wisconsin–Madison.

Liu, T.I. and Anantharaman, K.S. 1994. Intelligent Classification and Measurement of Drill Wear. Journal of Engineering for Industry, Vol. 116, Transactions of the ASME, pp. 392–397.

Liu, T.I. and Ko, E.J. 1990. On-Line Recognition of Drill Wear via Artificial Neural Networks. Monitoring and Control for Manufacturing Processes, PED. Vol. 44, ASME, Dallas, TX. Pp. 101–110.

Liu, T.I. and Wu, S.M. 1990. On-line Detection of Drill Wear. Journal of Engineering for Industry, Vol. 112, Transactions of the ASME, pp. 299–302.

Liu, T.I., Chen ,W.Y. and Anantharaman, K.S. 1998. Intelligent Detection of Drill Wear. Mechanical Systems & Signal Processing, Vol. 12, No. 6, pp. 863–873.

Matlab software package. MathWorks Inc. http://mathworks.com/

Mathsoft. 2002. Mathcad 11 User's Guide. Mathsoft Engineering & Education Inc. ISBN 1-57682-297-4

McPhee, S.C., Strafford, K.N., Wilks, T.P. and Ward, L.P. 1995. Evaluation of Coated Twist-drills using Frequency-domain Analysis. Surface & Coating Technology, Vol. 71, pp. 215–221.

Milton, J.S. and Arnold, J.C. 1995. Introduction to Probability and Statistics. Principles and Applications for Engineering and the Computing Sciences, 3rd ed. McGraw–Hill. ISBN 0-07-042623-6

Narayanan, S.B., Fang, J., Bernard, G. and Atlas, L. 1994. Feature Representations for Monitoring Tool Wear. IEEE. Pp. 137–140.

Newland, D.E. 1993. An Introduction to Random Vibrations, Spectral and Wavelet Analysis. 3rd ed. Longman Scientific & Technical. ISBN 0-470-22153-4

Noori-Khajavi, A. 1992. Frequency and Time Domain Analyses of Sensor Signals in a Drilling Process and Their Correlation with Drill Wear. Ph.D. Thesis, Oklahoma State University, Stillwater, OK.

Noori-Khajavi, A. and Komanduri, R. 1993. On Multisensor Approach to Drill Wear Monitoring. Annals of the CIRP, Vol. 42, No. 1, pp. 71–74.

Noori-Khajavi, A. and Komanduri, R. 1995a. Frequency and Time Domain Analyses of Sensor Signals in Drilling, Part 1: Correlation with Drill Wear. International Journal of Machine Tools and Manufacture, Vol. 35, No. 6, pp. 775–793.

Noori-Khajavi, A. and Komanduri, R. 1995b. Frequency and Time Domain Analyses of Sensor Signals in Drilling, Part 2: Investigation of the Problems Associated With Sensor Integration. International Journal of Machine Tools & Manufacture, Vol. 35, No. 6.

Pan, H., Chen, Y. and Orady, E. 1993. Monitoring Methods of Tool Wear in a Drilling Process. Proceedings of the 2nd Industrial Engineering Research Conference. Pp. 380–384.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. 2002. Numerical Recipes in C++. The Art of Scientific Computing. 2nd ed. Cambridge University Press. ISBN 0-521-75033-4

PSK 5705 Standard. 2004. Condition Monitoring, Vibration Measurement, Planning of Condition Monitoring, PSK Standardisointiyhdistys ry.

Quadro, A.L. and Branco, J.R.T. 1997. Analysis of the Acoustic Emission during Drilling Test. Surface & Coating Technology, Vol. 94–95, No.1–3, pp. 691–695.

Radhakrishnan, T. and Wu, S.M. 1981. On-Line Hole Quality Evaluation for Drilling Composite Material Using Dynamic Data. Journal of Engineering for Industry, Vol. 103, Transactions of the ASME, pp. 119–125.

Ramamurthi, K. and Hough, C.L. Jr. 1993. Intelligent Real-Time Predictive Diagnostics for Cutting Tools and Supervisory Control of Machining Operations. Journal of Engineering for Industry, Vol. 115, Transactions of the ASME, pp. 268–277.

Randall, R.B. 1977. Application of B&K Equipment to Frequency Analysis. Brüel & Kjær. ISBN 87 87355 14 0

Rao, V.B. and Rao, H.V. 1993. C++ Neural Networks and Fuzzy Logic. Management Information Source, Inc. ISBN 1-55828-298-X

Rehorn, A.G., Jiang, J. and Orban, P.E. 2004. State-of-the-Art Methods and Results in Tool Condition Monitoring: A Review. International Journal of Advanced Manufacturing Technology. Published online 11 August 2004.

Rotberg, J., Braun, S. and Lenz, E. 1990. Vibration-Based Drill Wear Monitoring. Manufacturing Review, Vol. 3, No. 1, pp. 60–65

Routio, M. and Säynätjoki, M. 1995. Tool Wear and Failure in the Drilling of Stainless Steel. Journal of Materials Processing Technology, Vol. 52, Issue 1, pp. 35–43.

Schehl, U. 1991. Werkzeugüberwachung mit Acoustic-Emission beim Drehen. Fräsen und Bohren, Aachen.

Subramanian, K. and Cook, N.H. 1977. Sensing of Drill Wear and Prediction of Drill Life. Journal of Engineering for Industry, 101, pp. 295–301.

Takata, S., Ahn, J.H., Miki, M., Miyao, Y. 1986. A Sound Monitoring System for Fault Detection of Machine and Machine States. Annals of the CIRP, Vol. 35, No. 1, pp. 289–292.

Tansel, I.N., Mekdeci, C., Rodriguez, O. and Uragun, B. 1993. Monitoring Drill Conditions with Wavelet Based Encoding and Neural Network. International Journal of Machine Tools & Manufacture, Vol. 33, No. 4, pp. 559–575.

Tansel, I.N., Rodriguez, O. and Mekdeci, C. 1992. Detection of Tool Breakage in Microdrilling Operation with RCE Neural Networks. PED, Vol. 47, No. 1, ASME, pp. 83–88.

Teti, R. 1995. A Review of Tool Condition Monitoring Literature Database. Annals of the CIRP, Vol. 44, No. 2, pp. 659–666.

Thangaraj, A. and Wright, P.K. 1988. Computer-assisted Prediction of Drill-Failure Using In-Process Measurements of Thrust Force. Journal of Engineering for Industry, Transactions of the ASME 110, pp. 192–200.

Thomson, W.T. 1972. Theory of Vibration with Applications. 4th edition, Prentice Hall, New Jersey. ISBN 0-13-915323-3

Tlusty, J. and Andrews, G.C. 1983. A Critictal Review of Sensors for Unmanned Machining. Annals of the CIRP, Vol. 32, No. 2, pp. 563–572.

Tsao, C-C. 2002. Prediction of Flank Wear Different Coated Drills for JIS SUS 304 Stainless Steel using Neural Network. Journal of Materials Processing Technology, Vol. 123, pp. 354–360.

Tönshoff, H.K., Wulfsberg, J.P., Kals, H.J.J., König, W. and van Lutterveld, C.A. 1988. Developments and Trends in Monitoring and Control of Machning Processes. Annals of the CIRP, Vol. 37, No. 2, pp. 611–622.

Valikhani, M. and Chandrashekhar, S. 1987. An Experimental Investigation into the Comparison of the Performance Characteristics of TiN an ZrN Coatings on Split Point Drill Using the Static and Stochastic Models of the Force System as a

Signature. The International Journal of Advanced Manufacturing Technology, 2 (1), pp. 75–106.

Von Nedeß, C. und Himburg, T. 1986. Automatisierte Überwachung des Bohrens. VDI-Z, Bd. 128, Nr. 17, pp. 651–657.

Waschkies, E., Sklarczyk, C. and Schneider, E. 1994. Tool Wear Monitoring at Turning and Drilling. Non-Destructive Characterization of Materials VI. Pp. 215–222. ISBN 0-306-44816-5

Watson, A.R. 1985a. Drilling Model for Cutting Lip and Chisel Edge and Comparison of Experimental and Predicted Results. I – Initial Cutting Lip Model. International Journal of Machine Tool Design and Research, Vol. 25, No. 4, pp. 347–365.

Watson, A.R. 1985b. Drilling Model for Cutting Lip and Chisel Edge and Comparison of Experimental and Predicted Results. II – Revised Cutting Lip Model. International Journal of Machine Tool Design and Research, Vol. 25, No. 4, pp. 367–376.

Watson, A.R. 1985c. Drilling Model for Cutting Lip and Chisel Edge and Comparison of Experimental and Predicted Results. III – Drilling Model for Chisel Edge. International Journal of Machine Tool Design and Research, Vol. 25, No. 4, pp. 377–392.

Watson, A.R. 1985d. Drilling Model for Cutting Lip and Chisel Edge and Comparison of Experimental and Predicted Results. IV – Drilling Tests to Determine Chisel Edge Contribution to Torque and Thrust. International Journal of Machine Tool Design and Research, Vol. 25, No. 4, pp. 393–404.

Williams, J.H., Davies, A. and Drake, P.R. 1994. Condition-Based Maintenance and Machine Diagnostics. Chapman & Hall. ISBN 0-412-46500-0

Williams, R.A. 1974. A Study of the Drilling Process. Journal of Engineering for Industry, Vol. 98, pp. 1207–1215.

Xiaoli, L. and Zhejun, Y. 1998. Tool Wear Monitoring with Wavelet Packet Transform-Fuzzy Clustering Method. Wear, Vol. 219, No. 2, pp. 145–154.

Yang, J.A., Jaganathan, V. and Du, R.A. 2002. A New Dynamic Model for Drilling and Reaming Process. International Journal of Machine Tools & Manufacture, Vol. 42, pp. 299–311.

Young, W.C. 1989. Roark's Formulas for Stress and Strain. McGraw–Hill International Editions. ISBN 0-07-100373-8

Zhang, M.Z., Liu, Y.B. and Zhou, H. 2001. Wear Mechanism Maps of Uncoated HSS Tools Drilling Die-cast Aluminum Alloy. Tribology International, Vol. 34, pp. 727–731.

PUBLICATION I

A summary of methods applied to tool condition monitoring in drilling

In: International Journal of Machine Tools and Manufacture 2002. Vol. 42, pp. 997–1010.
Reprinted with permission from the publisher.

Pergamon

International Journal of Machine Tools & Manufacture 42 (2002) 997–1010

INTERNATIONAL JOURNAL OF
MACHINE TOOLS & MANUFACTURE
DESIGN, RESEARCH AND APPLICATION

A summary of methods applied to tool condition monitoring in drilling

Erkki Jantunen *

Technical Research Centre of Finland, PO Box 1702, FIN-02044 VTT, Finland

Received 31 May 2001; received in revised form 18 March 2002; accepted 22 March 2002

Abstract

This paper presents a summary of the monitoring methods, signal analysis and diagnostic techniques for tool wear and failure monitoring in drilling that have been tested and reported in the literature. The paper covers only indirect monitoring methods such as force, vibration and current measurements, i.e. direct monitoring methods based on dimensional measurement etc. are not included. Signal analysis techniques cover all the methods that have been used with indirect measurements including e.g. statistical parameters and Fast Fourier and Wavelet Transform. Only a limited number of automatic diagnostic tools have been developed for diagnosis of the condition of the tool in drilling. All of these rather diverse approaches that have been available are covered in this study. In the reported material there are both success stories and also those that have not been so successful. Only in a few of the papers have attempts been made to compare the chosen approach with other methods. Many of the papers only present one approach and unfortunately quite often the test material of the study is limited especially in what comes to the cutting process parameter variation, i.e. variation of cutting speed, feed rate, drill diameter and material and also workpiece material. © 2002 Published by Elsevier Science Ltd.

Keywords: Tool wear; Drilling; Monitoring methods; Signal analysis; Diagnostic tools

1. Introduction

Tool wear and failure monitoring has raised quite a lot of interest among researchers and has consequently been studied in a number of research projects by a number of research organisations. The reason for the interest is that tool condition monitoring is considered important for the following reasons:

- Unmanned production is possible only if there is a method available for tool wear monitoring and tool breakage detection.
- Tool wear influences the quality of the surface finish and the dimensions of the parts that are manufactured.
- The economical tool life cannot be benefited from without a means for tool wear monitoring because of variations in tool life.
- Today tool changes are made based on conservative estimates of tool life which does not take into account sudden failures and at the same time leads to an unnecessarily high number of changes because the full lifetime of tools is not taken into account and consequently valuable production time is lost.
- As a consequence of the above, automated production control is not really possible without a means for tool wear monitoring.

The economical values involved in modern manufacturing are very high because of the high investments in the manufacturing equipment and naturally it would be in the interest of the industry to benefit from the equipment in an optimal way including automated production with high availability.

In principle, the tool wear monitoring methods can be classified in two categories, i.e. direct and indirect methods. With direct methods it is possible to determine tool wear directly, which means that these methods really measure tool wear as such. In spite of the many attempts direct methods such as visual inspection or computer vision etc. have not yet proven to be very

* Fax: +358-9-460627.
E-mail address: erkki.jantunen@vtt.fi (E. Jantunen).

0890-6955/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0890-6955(02)00040-8

attractive economically nor technically. In this paper only indirect measuring methods such as torque or vibration are covered. In fact the paper tries to cover all the indirect methods in drilling that have been found in the literature search that has been carried out.

There are differences in how well certain monitoring methods work depending on the purpose they are used for in tool condition monitoring. Some of the methods are more effective for detection of a sudden failure and some are more suited for tool wear monitoring. In this respect there is even more distinction in the suitability of the different signal analysis methods. It could even be claimed that the most effective and reliable methods for tool wear monitoring are so slow in practise that they are not suitable for the detection of sudden failures. Again the paper tries to cover both aspects when individual analysis techniques are discussed.

Drill wear is a progressive process which takes place at the outer margin of the flutes of the drill due to the intimate contact and elevated temperatures at the tool workpiece contact [1]. However, under constant cutting conditions drill failure is a stochastic process. The reasons for varying drill life are the inhomogeneities in the workpiece and drill materials, the irregularities in the cutting fluid motion and the unavoidable asymmetry introduced during the grinding of the cutting edges.

Similarly, as in the case of measuring methods, quite a number of signal analysis techniques have been tested for tool wear monitoring. In machining there are many disturbances and even the process as such can be run using different process parameters and hence signal analysis is really needed in order to be able to separate the wanted information from the rest of the "noise".

During the recent years quite a lot of effort has been spent on developing methods for automatic diagnosis of tool wear because automation of diagnosis is also needed in order to facilitate automatic production systems. Especially different types of neural networks have gained a lot of interest. The attempts to make the diagnosis automatic are also covered in this paper.

2. Measuring methods

A summary of the monitoring methods that various researchers have studied is shown in Table 1. In addition to the methods that have been tested and described in each reference, the possible coverage of the effect of cutting speed and feed rate is shown in Table 1, i.e. the table shows whether the researchers have tried to cover the effect of cutting conditions to the measured signals and calculated parameters. One reason for measuring cutting speed and feed rate is the use of these as parameters in adaptive control systems, e.g. [2].

Torque, drift and feed force together with strain measurement are all measures of cutting forces and are treated together in the subsequent study. In Table 1 the reported strain measurements are tabulated in the applicable force category because strain as such is linked to the force: force transducers actually measure strain which then is transposed to force. Spindle motor and feed drive current are closely related to the forces, i.e. they too measure the same cutting forces and phenomena, although through a longer measuring chain where also other factors influence the signals. Again spindle motor and feed drive current are treated together in the subsequent studies.

Vibration, sound, ultrasonic vibration and acoustic emission are actually all vibration measurements, although the frequency range in each of these differs and, in addition to that, sound is airborne vibration when all the others are mechanical vibrations of the structure. The frequency range in vibration measurements is typically from about 1 Hz to about 10 kHz (or 20 or 16 kHz is used as a limit [3]); in sound measurements the range is from 20 Hz to 20 kHz, which is the range a young person can hear; in ultrasonic vibration the frequency range is from 20 kHz to about 80 kHz [4]; and acoustic emission starts where ultrasonic vibration ends up and ranges to about 1 MHz. Again all the vibration related techniques are treated together. In some cases the measured vibration frequencies do not fall into the limits defined above and if this is the case then both categories are marked with "×". This is the case e.g. for vibration and ultrasonic vibration which have both been marked when the band-passed frequency is from 0.5 to 40 kHz, as it is in [5].

2.1. Torque, drift force and feed force

It is very logical to monitor forces in a cutting process in order to follow the development of cutting tool wear. It is generally known that cutting forces increase as tool wear increases [6]. This is due to the increase of friction between tool and workpiece. In drilling it is possible to monitor torque, drift forces (lateral forces affecting the workpiece) and the feed (thrust, z-axis) force. All of these have been monitored in Ref. [7]. The idea behind monitoring torque and feed force is very clear, i.e. it is expected that these forces change as the tool gradually wears. The thrust force has been used as the only measured signal in [1,8–10]. The simultaneous monitoring of thrust force and torque is rather common (see e.g. [2,6,11–19]) and special electronics have been developed for this purpose [11].

Drill wear as such differs to some extent from the wear of other cutting tools. Due to production tolerances a drill is slightly asymmetric, therefore it only wears at one lip until the height of both lips is equal [7,20]. The second lip, which is now sharper, starts cutting. This alternating process continues until neither lip has no more clearance at the margin. In the end the drill sticks

Table 1
Summary of monitoring methods that have been studied for tool condition monitoring in drilling

Reference number	Torque	Drift force	Feed force	Vibration	Sound	Ultrasonic vibration	Acoustic emission	Spindle motor current	Cutting speed	Feed drive current	Feed rate
[1]			×						×		×
[2]	×		×						×		×
[3]	×	×	×	×		×			×		×
[4]				×		×			×		
[5]			×	×		×			×		×
[6]	×		×						×		×
[7]	×	×	×		×				×		×
[8,9]			×								
[10]			×								
[11]	×		×								
[12]	×		×								
[13,36–38]	×	×	×	×	×		×	×	×	×	×
[14]	×		×								
[15]	×		×								
[16]	×		×						×		
[17]	×		×						×		×
[18]	×		×					×	×	×	×
[19]	×		×					×			×
[20]				×					×		×
[21]	×	×	×						×		×
[22–24]	×	×	×								
[25]			×					×	×	×	×
[26]	×	×	×								
[27]	×										
[28]				×							×
[29]			×	×	×				×		
[30,31]			×	×							
[32]							×				
[33]				×				×		×	
[34]								×	×	×	×

into the workpiece and breaks if the cutting process is not stopped. Assuming this kind of wear progress gives reason to monitor the drift forces. In a series of tests [21] no consistent change of feed force or torque was observed but a certain change in the drift forces was recorded. This is again explained to be because first the cutting edge on one side and then on the other side wears.

The measurement of thrust force and torque have been linked to the waviness of the hole surface and especially the effect of tool wear to the waviness has been studied in [17]. In the analysis more emphasis has been given to thrust force than to the torque, i.e. thrust has been considered a more reliable indicator of tool wear.

Torque, feed force and strain of the table in two directions have been measured in [22–24]. The strain measurements actually in their function correspond to the measurement of drift forces, i.e. they serve the same purpose. Strain has also been measured in [25], but in this case located in the spindle and corresponding to the measurement of thrust force. Torque, drift and feed force have been also measured in [3] and compared with the measurement of ultrasonic vibration. Also in [26] torque, drift and feed force have been measured simultaneously when comparing two different types of coatings (titanium nitride and zirconium nitride).

A new method for measuring torque is suggested in [27]. The technique is based on the measurement of eddy current. The sensor can be positioned some 0.2–0.5 mm from the drill shank. This technique is affected by the distance between the sensor and the drill shank and also the material of the drill has an effect on the measured torque. The method is suitable for both static and dynamic torque measurements and consequently suited for both wear and failure monitoring. The method has been patented in Germany.

Based on the tests with copper alloy and a model described in [19], formulas that define the thrust force and torque as a function of feed per revolution, drill diameter and flank wear have been developed and their applicability has also been tested [6]. It should be noted that the tests indicated that the increase in cutting speed over the range studied had no significant effect on work material strength, and hence it has no significant effect on cutting forces [6]. In fact the correlation of the regression formulas with the test data without the rotational speed of the spindle is very good (for feed force R^2=0.94 and for torque R^2=0.97). It is concluded that tool wear can be properly estimated knowing the thrust force and other cutting parameters, especially for larger tool wear.

Based on tests with different workpiece material hardness, formulas for torque and thrust force have been developed as a function of Brinell hardness of work material, diameter of the drill, feed per revolution, average flank wear and radius at the cutting edge [19]. It is concluded that the variation in drill life is significantly influenced by the workpiece hardness. It is speculated that it could be so that the presence of a few random workpieces with a high hardness may influence the drill life much more than a large number of workpieces with a low hardness. Hence, in an industrial operation, drills may fail very early or after a long time, depending on the occurrence of a few workpieces with a high hardness. This could explain the large variation in drill life observed in industrial conditions. The workpiece hardness also influences the amplitudes of thrust forces and torque occurring in a drilling operation. If the variation in thrust force, on account of changes in flank wear, is to be significant, the variation in workpiece hardness has to be held within 5% of the mean hardness value in order to be able to base the diagnosis of flank wear on the amplitude of thrust force or torque. This is very difficult to achieve in industrial castings. Hence, torque or thrust measurements for monitoring drill wear should be attempted only after a very close tolerance has been obtained in the workpiece hardness.

2.2. Vibration and sound

Vibration is widely used for condition monitoring of rotating machinery. However, vibration has not been used to the same extent in tool condition monitoring, probably because as a method it is rather sensitive to noise which is present in cutting processes. The advantages of vibration measurement include ease of implementation and the fact that no modifications to the machine tool or the workpiece fixture are required [20]. However, the disadvantages reported in the literature include dependency of the vibration signals on workpiece material, cutting conditions and machine structure.

The work of [20] deals with the development of vibration-based monitoring methods for detecting breakage of small size drills (3 mm diameter) and wear of larger size drills (6 mm diameter). Vibration is measured both in the transverse and axial direction. The vibration signals are considered to contain reliable features for monitoring drill wear and breakage for the following reasons: the vibrating drill length in the transverse and axial modes does not change during drilling, thus maintaining a rather constant mode frequency; the natural frequencies of the transverse and axial modes of the workpiece–drill system are basically insensitive to drill cross-sectional size, thus simplifying monitoring for a wide range of drill sizes; vibrations in the directions Y and Z are influenced by the torque and thrust force which are the major excitation sources in drilling.

In the tests reported in [28] three accelerometers were used each measuring in the direction of one of the three axes. In [29] both vibration and the use of sound measurements are discussed. The sound measurement and analysis is discussed in more detail in [7].

Vibration measurement together with thrust force has been used in the tests reported in [30,31]. The purpose of the tests has been to obtain signal for the development of a diagnosis tool capable of recognizing tool wear. In the tests tool wear has been recorded with a vision system.

In theory, sound measurements could be expected to give the same information as can be detected using vibration measurements because in the structural boundary the mechanical vibration of the structure or tool/workpiece contact is partly transferred to airborne vibration, i.e. sound. However, quite a number of factors influence how the mechanical vibration is transferred and how it takes place at the different frequencies. Also there is a great difference when the influence of disturbances from outside sources are compared in vibration and sound measurements. The sound measurements are more vulnerable than vibration but at the same time it should be remembered that the operators sometimes or perhaps actually rather often rely on what they hear when they define whether the tool is worn or not. In [13] both vibration and sound measurements together with a number of other methods have been tested and compared in drilling, with the result that vibration was the most effective method of all of the tested methods.

A higher frequency range from 0.5 to 40 kHz for vibration measurements has been tested with very thin drills. The reason for looking at this kind of frequency range is that the rotational natural frequencies fall into that range since for a drill of 1 mm diameter the natural frequency could be about 25 kHz and for a drill of 3 mm diameter it could be about 7 kHz [5]. In the reported examples the band-pass filtered vibration signal has given more clear indication of both tool wear and failure than the feed force signal [5].

2.3. Acoustic emission and ultrasonic vibration

The use of ultrasonic vibrations (UEs) in the frequency range from 20 to 80 kHz for tool breakage detection in various metal cutting processes including drilling has been tested [4]. The practicality of using ultrasonic vibrations is explained when compared to other vibration techniques. Acoustic emission (AE) is seen to suffer from severe attenuation and multi-path distortion caused by bolted joints commonly found in machine tool structures and restricting the mounting location of the AE transducer to somewhere very near the tool or workpiece. The lower frequency signal used for UE analysis does not suffer such severe attenuation and distortion, and so the transducer can be placed fairly far from the chip forming zone. In the low vibration frequency range, i.e. below 20 kHz, structural modes are prominent. A common strategy is to compare the amplitudes of several frequency bands in this range. Particular variation in the relative strengths of vibration in these bands indicate process abnormalities such as tool breakage or tool wear. This method shares the advantage of remote transducer placement with the UE method but unfortunately is much more sensitive to machine and tooling variations. Since structural modes change in complex ways with machine movement, loading, temperature, and tooling, this approach generally must be tuned empirically each time that the process is changed. In contrast, in the frequency range used for UE analysis the structural modes are so closely spaced that they form a so-called pseudo-continuum. There are no individual resonances to shift out of the analysis band with machine movement, loading, and so on.

The applicability of ultrasonic vibration measurement for the tool wear and failure detection has also been studied in [3]. In the reference the frequency range in question, i.e. from 10 to 70 kHz, is defined as acoustic emission and the used sensor with non-linear frequency response is considered as an AE-sensor. However, following Ref. [4], the frequency range in question is in this context defined as the ultrasonic range. In [3] ultrasonic vibration is compared with torque, feed and drift force measurement and proven to be a more effective means for tool wear and failure detection in drilling. The same sensor has also been used for measurements in the frequency range from 1 to 5 kHz which normally is considered mechanical vibration.

Acoustic emission is a phenomenon which occurs when, for different reasons, a small surface displacement of a material surface is produced [32]. This occurs due to stress waves generated when there is a rapid release of energy in a material, or on its surface.

Acoustic emission with centre frequencies of 200 and 800 kHz and also in a broader band from 100 to 1000 kHz has been tested in [13]. In the tests the 200 kHz sensor was used for tool wear and the 800 kHz sensor for tool breakage detection. The broad band sensor was used for finding the best frequency range for further investigation. Also in [32] acoustic emission was recorded in a broad band from 100 to 1000 kHz in order to monitor tool wear.

2.4. Spindle motor and feed drive current

Spindle motor current is in principle a measure of the same feature as torque, i.e. they both enlighten how much power is used in the cutting process and they both also advise about the dynamics of cutting. It is fair to claim that torque is a more sensitive way to measure than is the spindle motor current since the torque sensor is located close to the cutting tool and e.g. the dynamics of the electric motor do not influence it to the same extent that they influence the current measurement. However, measuring torque is more complicated than measuring the current of the spindle motor and therefore the measurement of the current has also been widely tested and used [13,19,33,34].

Similarly, as spindle current corresponds to torque, feed drive current corresponds to the measurement of the thrust force. Again there is some similar difference in the sensitivity of the methods as described above. The feed drive current as an indicator of tool wear and failure has been studied in [13,33,34].

Both feed drive and spindle current have also been measured in [25]. In these tests it has been possible to compare the measurement of feed current to the measurement of thrust force based on the use of strain gases. It is stated that typically, the strain gage is a better sensor than the feed motor current sensor for wear diagnosis. Nevertheless, the current sensor was used to investigate whether the cost effective and easily implementable current sensors alone would suffice.

The reported results in [18] for feed drive current and spindle power together with feed force and torque are quite similar. The measurement results show that all the quantities measured remain at an almost constant level during the entire tool life-time until the hole in which the drill totally fails. It is impossible to successfully apply these measurements as tool-monitoring methods, stopping the machining after the increase in one or several signals above a particular limit value before actual tool failure. However, the measurements can be used for tool-breakage detection where the machining operation is interrupted after tool breakage. With this system, one workpiece may be rejected because of the tool failure, but further damage is avoided.

3. Signal analysis

The kind of signal analysis methods used is of some importance. Sometimes it looks as if some researchers think that if the measured signal is acceptable then it would be possible with a clever diagnostic tool to solve everything. Unfortunately this is not the case. The diagnosis always needs to be based on reliable and meaningful information and this is where signal analysis can help by providing effective features as a basis for diagnosis. The role of signal analysis could be described as a tool which tries to pick up the meaningful information out of the mass of information. In many cases the dilemma is that the more sophisticated methods need a lot of raw signals and it takes time to collect this raw material and it also takes time to perform the calculations. Consequently, many of the most sophisticated methods are not suitable, e.g. for tool breakage monitoring. In addition, the results with a sophisticated analysis function are influenced by the cutting process, i.e. workpiece material, type of tool, feed rate and cutting speed which makes the diagnosis more demanding. On the other hand, very simplistic methods are fast to use and often not that sensitive to changes in cutting conditions. Unfortunately, at the same time they are not so sensitive to tool wear either. A summary of signal analysis methods that have been tested, used and reported in the literature for drill wear and failure monitoring is given in Table 2.

3.1. Time domain signal

The time domain signal is not very informative as such, or at least it is very time consuming to look at the raw signal in graphical format (e.g. with an oscilloscope [14]). Evaluation of the changes by measuring only the amplitude of the signal is very complicated and therefore an RMS-voltmeter is used [21]. Usually a number of statistical parameters such as root mean square (RMS), arithmetic mean, standard deviation and kurtosis are calculated and these are then used for comparison and diagnosis.

With almost all of the measuring signals the most common parameter to look at is the RMS value, which also is actually the value that is normally seen if the signal is drawn with a plotter or looked at with a voltage meter. The RMS value contains all the energy in the signal and therefore also all the noise and all the elements that depend on the cutting process. Therefore, it is not the most effective parameter but has retained its place because it is so easy to produce and understand. Besides, it does actually work when compared to other statistical parameters. In a series of tests in [13] the RMS value was compared to seven other statistical parameters, i.e. arithmetic mean, mean and standard deviation, skewness, kurtosis, maximum and minimum. The comparison showed that the RMS value is usually not the best but it is often one of the four best functioning parameters.

In the tests reported in [22–24], mean value together with the variance of the sensor signals (torque, feed and drift force) have been calculated for all of the holes. No significant changes were found in the mean and the vari-

Table 2
A summary of signal analysis methods that have been used for tool condition monitoring in drilling

Reference number	Time domain signal	Statistical parameters	Auto regressive moving average	Fast Fourier transform spectrum	Cepstrum analysis	Higher-order spectrum	Wavelet transform
[1]	×	×					
[3]	×	×		×			
[4]	×	×					
[5]	×	×					
[6]	×	×					
[7,27,29]	×	×		×			
[8]							×
[9]	×	×					
[10]	×	×		×			
[11]	×	×					
[12]				×			
[13,36–38]	×	×		×	×		
[14]	×	×					
[15]	×	×					
[16,30,31]	×	×					
[17]	×	×	×				
[18]	×	×					
[19]	×	×					
[20]	×	×		×	×		
[21]	×	×					
[22–24,39]	×	×		×			
[25]	×	×					
[26]	×	×		×			
[28]				×		×	
[32]	×	×					
[33]	×	×					
[34]							×

ance of sensor signals. Therefore, it has been concluded that the force sensor signals in the time domain do not show any correlation with drill wear.

Based on the comparison of static and dynamic components of the feed force and torque, the analysis of the process-dynamics in drilling is considered essentially a more delicate instrument to the investigation of the wear condition than the interpretation of the increase in static feed force and torque [10].

Due to the great variation in measured signals, i.e. dynamic behaviour, average values for longer test period are often used in statistical studies. For example, in [6] average values of thrust force and torque are used when developing tool wear models.

Average, peak, RMS values and the area of thrust and torque have been used as input features in the diagnostic system described in [16]. These features have been chosen because of their previous successful application for on-line monitoring and diagnosis. Furthermore, these features were justified from the researchers' experimental observations.

Mean, peak and standard deviation have been used in the analysis of thrust force and torque signals in [17]. Of the tested statistical parameters, standard deviation proved to be the best indicator of tool wear and it was the indicator that is more closely related to the change in the standard deviation of the hole surface in composite material.

Mean, standard deviation and maximum values of the thrust force have been studied in [1]. From a series of drilling experiments conducted in the laboratory, the gradient of the thrust force has been identified to be a suitable process parameter for prediction of drill failure. A Finite Impulse Response filter using a Hamming window has been designed and used to determine the gradient of the thrust force data. Experimental evidence emphasizes the correlation between thrust force and outer corner wear; it is suggested that the sharp spikes in the thrust force that are observed under failure conditions are caused by a macroscopic stick–slip phenomenon. It has been shown that the proposed approach does not require considerable tuning for operation under a wide range of cutting conditions. This would make it ideally suited for an industrial environment.

Mean value of cutting forces (torque, drift and thrust force) has been studied in [26]. Also the maximum and minimum deviations about the mean value have been studied. In the tests two different types of drill coatings were used. The mean values were much smaller with one of the coatings (zirconium nitride) than the other (titanium nitride). The recorded mean values and deviation from these values have not given a logical indication of tool wear or alarm for tool breakage.

Smoothed average and standard deviation values of thrust force have been calculated in [9] for the detection of poor operation conditions (just before breakage, breakage, and drilling with broken tool) in micro-drilling. The processing of the data is done in four segments during each drilling cycle. These studies indicated that the average force and standard deviation value must be presented together when used as input to a neural network. Also, the study indicated that the main cause for failure was not related to tool wear. Most of the time, the very thin shaft of the drill could not carry the loads and it broke. In the test cases, total drill life varied between 0.1 and 10 mm. There was no considerable difference between the force characteristics after the first and the 25th hole, except when the tool was broken or damaged.

A mix of statistical parameters is used in [25]. For spindle motor current the use of RMS has been justified in the following way. The low frequency energy of the spindle motor current is directly proportional to the cutting torque exerted by the tool on the workpiece. As the tool wears, the torque requirement increases and correspondingly the spindle motor current also increases. The RMS value of the spindle motor current thus becomes a valuable feature for wear prediction. In addition to the RMS value, the change in RMS value with respect to the first hole is also another good feature, since it indicates the temporal trend of the cutting torque. Also in the case of feed, motor current RMS value with a corresponding parameter indicating the change are used. For thrust force (strain gage) measurement the mean value again together with the corresponding indicator for the trend are used.

In [3] the emphasis is on the way the wear influences ultrasonic vibration in different frequency ranges, i.e. 10–20, 20–30, 30–40, 40–50, 50–60 and 60–70 kHz. The RMS value of the band passed signal has been used. There is variation in how well tool wear is observed in the different frequency ranges, although all the time the percentage increase in RMS value of some of the frequency ranges of ultrasonic vibration are always higher than is the case with the measured forces. An acoustic emission sensor in the frequency range from 1 to 5 kHz (normally considered vibration) has proven to be especially suitable for tool wear monitoring. Apparently there have been structural vibration modes that have their frequency in this frequency range and thus increase the signal level. Tool failure has also been clearly detected with the same sensor, though the indication is clearer at higher frequencies, e.g. from 20 to 40 kHz.

Maximum stable values are used for feed force, torque, spindle and feed drive current [18]. In the case of the spindle power and current of the Z-axis motor, the values represent the difference in the measured quantity between cutting and idle running at the corresponding rotational frequencies.

Kurtosis value is defined as the fourth central moment of a Gaussian distribution and is a measure of peakedness of the signal. Therefore, in [20], a lot of emphasis is given for this value to be used as a possible indicator for tool failure. In [20] a new parameter called ratio of the absolute mean value (RAMV) is also introduced, since kurtosis was not reliable alone due to its tendency to decrease when the number of peaks in the signal becomes high. RAMV represents the ratio of absolute mean value at the current revolution of the spindle to the absolute mean value in the beginning of the drilling process, i.e. RAMV is a normalized mean value calculated with a time constant of one revolution. The RAMV value has been used with good success for triggering of the calculation of kurtosis value together with cepstrum analysis. In the tests [20] kurtosis value was found to be insensitive to cutting conditions or changes in the workpiece hardness.

One way to further process the time domain signal is to use envelope detection. As such envelope detection can be used as a practical alternative for analysing signal containing information at high frequencies and thus making the analysis process easier [7,29]. The possible use of moments of the probability distribution of intensities and time of occurrences is also discussed and a trend index (TI) based on these is described in [7,29]. The published TI curve [7] seems to indicate tool wear but does not as such give a clear indication of when the tool should be changed.

When the envelope of a signal is calculated the process at first also involves band-pass filtering of the signal. Low-pass, high-pass and band-pass filtering can all be regarded as time domain parameters and are often used, as for example band-pass filtering of the vibration signal from 0.5 to 40 kHz in [5] in the case of thin drills in order to concentrate the analysis in the frequency range where the rotational natural frequency of the drill is expected to lie. The same approach has been used for both tool wear and failure detection.

Envelope detection together with the use of the flexible tool breakage algorithm is described in [4]. A fundamental quantity used in the signal analysis is the running mean. To establish an average signal that is not influenced by large pulses, a clipped running mean is computed each time through an algorithm loop. The running mean is a non-linearly weighted arithmetic average of the most recent samples. The clipping performs the non-linear weighting by limiting the contribution of samples larger than a certain ratio times the current mean. A suspicion test compares the most recent sample to an upper and a lower level, each of which is a multiple of the current running mean. Together with some other similar tests based on statistical parameters calculated from the time domain signal, the test forms an algorithm that is capable of detecting tool breakage.

It should be noted that tool wear monitoring in drilling is a very periodic process, i.e. drilling one hole does not usually last very long. In addition it is possible to recognize certain stages in drilling when monitoring is practical [32]. Usually in the process the drill first touches the work material and thereafter progressively drills it, with a given penetration rate. After the final depth is reached, the descending mechanism of the drilling machine is stopped and the drill keeps rotating but without drilling any further. A moment later the drill is retrieved from the hole which is then completed. Naturally, the measuring signals vary as a function of the drilling stage. In [32] the transient drilling stage (when the drill starts to penetrate into the workpiece) and the stage when the drill is stopping were found to be the best moments to monitor tool wear using the envelope value of acoustic emission.

3.2. Autoregressive moving average

Stationary stochastic process data in the form of a single, time dependent series can be mathematically modelled as an Autoregressive-Moving Average or ARMA model [17]. The modelling strategy involves fitting models in increasing order n starting from 1. The adequacy of the model may be tested using the conventional F-test. In condition monitoring the autoregressive parameters or their relations have often been used for diagnosis of faults or failures. In [17] the autoregressive model is based on the use of thrust force and torque signal and it has been used to define frequencies of modes that have then been used as the frequencies for which spectral density has been calculated. This technique has been called the Dynamic Data System (DDS) technique. With that it has been possible to get information of the contribution of each of the frequencies to the overall variance of the data. It is concluded that the dispersion analysis using the DDS technique shows a very strong correlation between the changes in the standard deviation of the lamination frequency (of composite material) component in the thrust and surface signals. This gives a direct indication of the change in the surface waviness and can be used to monitor the drill condition on-line for appropriate replacement of the drill.

3.3. Fast Fourier transform

The widely used Fast Fourier transform (FFT) provides a means to find out the frequency content of a measured signal. Assuming the wear influences the frequency contents of the measured signal, FFT then gives an inside view of this process. Many studies about the effectiveness of FFT have been reported [7]. Although the calculation of the power spectrum is a more sophisticated way of signal analysis than the calculation of many of the statistical parameters in the time domain and thus is a more powerful tool to get rid of noise and disturbances [13], it does suffer from limitations such as [20]: (a) materials such as cast iron are not homogeneous and will affect the amplitude of the vibration measured, which may cause false alarms; (b) tool damage in drilling produces a high level of transient vibrations (spikes) which are largely attenuated by the averaging procedure typically used in spectrum calculations and this makes it difficult to extract a discriminating feature to distinguish the change in the tool conditions; (c) non-uniform hardness of the workpiece material, built up edges, and micro-cracks can also cause false alarms by increasing the vibration amplitude. In order to decrease the adverse risks explained earlier, the trapezoid method has been used to calculate the area of the power spectrum between two frequencies in order to monitor tool wear with vibration [20].

The power spectral density function of torque, drift and feed force have been calculated in [26] for two different types of drill coatings tested (zirconium nitride and titanium nitride). It is concluded that the power content of the axial force and torque is significant over the entire frequency range, whereas the power content of the drift force is band limited. The power spectrum of the drift force changes from a band limited process to a wide band process when the drill is worn. The power content of the high frequencies of the cutting forces (especially the drift force component) increase as the tool approaches failure. This can be used as an index to detect the failure of the cutting tool.

Sometimes the number of points in the time domain is kept very small compared to typical values like 2048. If a small number of points is used, calculation of the power spectrum is much faster and also the frequency resolution is lower which is an advantage in the sense that even though the frequency of amplitudes in the spectrum might wander a little, they stay at the same frequency in the power spectrum. Another advantage is that the number of possible features that are used as input for a diagnosis system is in this way limited. In [12] only 256 points in the time domain have been used, which corresponds to a spectrum of 128 points in this case.

The somewhat limited 256 points in the time domain have also been used in [22–24]. The area under the power spectral density function (obtained through the Welsh method) has been studied with success. Averaging of the spectrums over a hole proved to give noisy results but this could be improved by averaging the results over a number of holes. All sensor signals, i.e. feed and drift force (strain) and torque gave similar results. Signal-to-noise analysis indicated that the power at frequencies between 50 and 300 Hz have the highest value of signal-to-noise ratio and, hence, are the most reliable frequencies. Comparison of the PSD plots showed that power at each frequency increases with increase in drill wear. Normalized PSD plots of all of the four sensor signals at different states of drill wear were coincident. This indicates that power at all frequencies increases proportionally with an increase in drill wear. Therefore, the change of area under the PSD plots was considered instead of power at one frequency, for integration decreases the error.

A number of FFT based functions such as autocorrelation, power spectrum (20 highest amplitudes, harmonics, as well as 1/3 and 1/1 octave bands), cepstrum and liftered spectrum and also two-channel functions such as cross-correlation, cross-spectrum, frequency response as well as some multi-signal frequency response function with more than two channels have been tested for tool wear monitoring in metal cutting including drilling [13]. The normal power spectrum worked well when the analysed data were fitted to a third order regression curve. Some of the two-channel functions (cross-spectrum, coherence) also proved to work well in drilling.

Cepstrum analysis is used to identify a series of harmonics or side bands in the power spectrum and to estimate their relative strength [20]. Cepstrum is calculated from the power spectrum either with inverse FFT (complex cepstrum) or taking the power spectrum of the logarithmic power spectrum (power cepstrum) [35]. In the tests [20] cepstrum analysis was performed only when a statistical RAMV indicator (explained earlier) reached a certain threshold value but the cepstrum showed larger amplitude at the frequency [35] corresponding to the time of one spindle revolution. In the tests reported in [13], cepstrum analysis also worked well in drilling and in milling. This is a logical result because both of these tool types have a number of cutting edges and when a fault starts to increase the difference between the way the cutting edges work becomes larger and consequently this is seen at the harmonics of the rotational speed of the tool which is what cepstrum then can show.

When compared to the traditional power spectrum, benefits from the use of the higher-order spectrum (HOS) features have been reported [28] in tool wear monitoring. Use of HOS features is reported to enhance monitoring performance primarily because they provide information on the strength of the non-linear and periodic component sideband structure in the received signal.

3.4. Wavelet transform

Wavelet transforms have become well known as useful tools for various signal processing applications [34]. Wavelet transform is described as a good solution in the time-frequency domain so that it can extract more information in the time domain at different frequency bands. Both continuous and discrete wavelet transforms are used for tool breakage detection using spindle and feed current signals. The test signals have been shown both in the time domain and after wavelet transform in [34], but no comparison with other methods is given. Hence it is difficult to compare whether what is seen clearly after the wavelet transform could also be seen clearly with some other analysis method, i.e. the benefits of the use of wavelet transform are not apparent.

In [8] the use of wavelet transformations together with neural networks is proposed in order to detect severe damage to micro-drills just before the breakage occurs. The use of wavelets is justified on the basis of the many weaknesses FFT has, the first being fixed resolution. The resolution of an entire frequency spectrum depends on sampling frequency and the number of data points. The second weakness is the representation of the entire spectrum, with the addition of harmonic signals, by assuming that the data window is repeated indefinitely. This assumption causes leakage problems and the transitions cannot be identified in the data window. A third weakness is the considerable noise in the transformations because of the very large degrees of freedom of the system. FFT analysis must be repeated several times and the results must be averaged to obtain smooth output. In [8], the Daubechies type wavelet system based on an orthonormal base was used. By using wavelets, the thrust force signal of the micro-drill has been simplified. The analysis indicated that the wavelet translation coefficients can represent the characteristics of micro-drilling signals with reasonable accuracy without high frequency components. The transition coefficients of all the normal micro-drills demonstrated similar patterns. The characteristics of the severely damaged micro-drills were found to be totally different. Based on these results, it is suggested that wavelets might be the perfect tool for many applications requiring automated monitoring of manufacturing operations. However, no comparison to FFT or statistical parameters has been made.

4. Diagnostic tools

The most simplistic methods of diagnosis in all monitoring is to use predefined limits, i.e. if a certain parameter in the analysis reaches a certain upper or lower limit this is an indication of a failure of the tool or worn tool. These types of fixed limits are often used by a human operator and they are also used in monitoring systems and form the basis of rule based expert systems. Quite similar systems based on fuzzy logic are based on these types of limits which then are fuzzy, i.e. they are not exactly defined and the limits in this case usually overlap. For example, the systems described in [33,36–38] use crisp limits and the systems described in [15] use fuzzy limits.

When performing diagnosis it is often more effective to be able to follow the trend in the monitored signal and parameter than just to look at the absolute value. The reason for this is that in many cases there are outside factors affecting the absolute value. In tool wear monitoring the limits for a certain parameter, e.g. vibration amplitude, are a function of the tool type, workpiece material, cutting parameters etc. Therefore, it is more effective to follow some trend in the signal, e.g. if the amplitude has increased to double or perhaps is five times what it was when the tool was sharp it can be supposed that the tool is worn. In [20] the more sophisticated analysis is only carried out when a situation occurs that a certain parameter reaches a predefined value compared to the value in the beginning. One possible disadvantage of trend analysis is the amount of data that might need to be saved in case the whole history of the signals of the tools were to be saved. The amount of data could be enormous in the case of a machine tool with a tool magazine of tens of tools. One possible solution in order to reduce the amount of data to be saved is given in [36,37]. The suggested solution relies on the use of regression analysis and the idea is to save only the summary terms of the regression function.

The use of neural nets can be seen as an attempt to automate the process of writing the diagnostic rules, i.e. if a sufficient amount of good data exists it is possible to train a net that is capable of diagnosing the condition of the tool. In principle, neural nets can be trained to model the non-linear dependencies of the measured and analysed parameters together with process parameters. If process parameters are left out of the model, either parameters that are insensitive to cutting conditions must be used or they need to depend in such a way both on process parameters and tool wear, and failure that the model works in a number of cutting conditions. Alternatively, a number of models corresponding to the possible cutting conditions need to be developed. This is in principle the same problem or limitation as described for the rule based approach. The previous statement can be rephrased in another way, i.e. if simple models based on less sophisticated parameters are used the number of models and corresponding work increases. Unfortunately, the opposite is also true, i.e. if sophisticated models which rely on sophisticated parameters are used, the time it takes to train the models increases as does the calculation time of the parameters. A summary of approaches adopted for diagnosis of tool condition in drilling is given in Table 3.

One of the earliest expert system concepts to monitor both the cutting process and the condition of the cutting tool is described in [33]. Among other things, the VILMOS-1 system is expected to monitor tool wear and tool breakage and also to protect the tools against overload.

In [36–38] a rule based expert system is described. The system consists of a number of modules: data acquisition and analysis, fault tree, symptom tree, rule synthesizer and fault manager. The system can be configured by the user through a graphical interface. The data are acquired through an AD-card using a number of measuring sensors such as vibration and acoustic emission etc.

Table 3
A summary of approaches adopted for diagnosis of tool condition in drilling

Reference number	Rule based	Fuzzy logic	Pattern recognition	Machining influence diagram model	Multi-layer perception neural network	Kohonen self-organising map	Restricted Coulomb energy networks	Adaptive resonance theory networks
[8]								×
[9]							×	
[12]						×		
[15]		×						
[16]					×			
[22,23,39]			×		×			
[25]				×				
[30]					×			
[31]			×					
[33]	×							
[36–38]	×							

Signal analysis is based on the use of statistical parameters and FFT based functions. All the data are saved in a database. The actual rules of the system are written automatically through the use of fault and symptom tree modules. The idea has been to make the system very flexible so it could be used for monitoring all kinds of machining processes with all kinds of tools. The actual diagnosis of the so-called fault manager relies on a number of parameters from a number of sensors.

A generalized Machining Influence Diagram (MID) is formulated for modelling different modes of failure in drilling [25]. A faster algorithm for this model is developed to solve the diagnostic problem in real-time applications. The MID model is utilized for diagnosing two failure modes: the drill wear, and the drill failure prediction. Each drilling operation is categorized deterministically using the machining parameters. The estimation of probability that the tool is worn is done by fusing information about wear from the two sensors: spindle and feed motor current. No sensor fusion is used for tool failure prediction since only the strain gage signal is used. The state of the drill is diagnosed. It consists of three states, "ok", "worn" and "about-to-fail". Three options are available for control: "continue", "reduce feed" and "replace tool". The cost of machining is a function of the control options and the state of the drill. The response time of the on-line system is well within the desired response time of actual production lines. The instance of diagnosis is reasonably close to the actual instance of wear. The accuracy of prediction has also been significantly promising. In cases where the drill wear is not diagnosed, the system is reported to at least predict drill failure, and vice versa. Consequently, by diagnosing at least one of the two failure modes, the system is able to prevent any abrupt failure of the drill.

The system described in [15] has two input features which are the feed force and torque and the wear of the drill is clustered in four wear states, i.e. initial, small, normal and severe. The approach is fuzzy, i.e. fuzzy limits are defined using the fuzzy C-means algorithm. With the presented two test cases used for the development of the system the approach works well. However, the approach does not take into account the effect of the cutting process into the measured parameters, i.e. the user would need to define new fuzzy limits for different types of workpiece materials and drills and also for different cutting parameters.

A two category linear classifier has been used for the detection of drill wear [31]. Sensor fusion is used for on-line drill wear detection. The indirect indexes used are the percentage increase of the peak-to-peak amplitude of vertical acceleration and the percentage increase of the drilling thrust. A two-category linear classifier is employed to distinguish the worn-out drills from those that are still usable. Flank wear area is used to categorize the drill conditions. The wear states are classified into two categories, usable and worn-out. Based on the present data a success rate greater than 90% has been obtained for on-line detection of the drill wear in one cutting process situation.

A rather simple neural network has been developed in [30] with two input features and one output. The number of neurons in the hidden layer has been varied from four to nine. Wear has been classified into five categories, i.e. initial, slight, moderate, severe and worn-out wear (with the same data as in [31]). It is concluded that neither the percentage increase of peak-to-peak amplitude of the vertical acceleration or the percentage increase of the thrust can be used for on-line classification of drill wear. However, integrating both signals yields better results. Based on the drilling tests, a success rate of over 85% can be reached for on-line recognition of drill wear using artificial neural networks. No variation of the cutting process parameters has been included, i.e. the wider applicability of the model has not been demonstrated.

The effectiveness of artificial neural networks with different numbers of hidden layer neurons together with the use of adaptive activation-function slopes have been tested in the diagnosis of tool wear in [16]. In all of the models nine input features (feed force and torque based on statistical parameters with one cutting process indicator) and one output have been used. The number of neurons in the hidden layer has been varied between 14 and 22. It is concluded that modified artificial neural networks with adaptive activation-function slopes converge much faster than the conventional feed forward neural networks. Artificial neural networks can distinguish between a worn and a usable drill on-line with 100% reliability and also accurately estimate the average flank wear even under different drilling conditions. The increase in number of neurons does not necessarily improve the accuracy of on-line drill wear measurement. A neural network with 9×14×1 architecture yielded the best results for on-line drill wear measurement. Although the reported results seem good even when changing the cutting conditions, the limitation in the presented material is that the variation in cutting conditions has not been documented and it is possible that there has not been any variation of feed rate in the tests.

The use of neural networks in the sensor integration has been studied in [22,23,39] based on torque, feed and drift force signals. In the thesis [22] it is shown that there is no point in trying to integrate the information from these signals because they all have equally good correlation with tool wear and one sensor is adequate for monitoring and controlling tool wear. It is also stated that integration of the sensor signals can introduce redundancy in the sensor integration technique and, in the presence of noise, result in the deterioration of the estimation of drill wear. Periodograms of sensor signals at different states of drill wear are mixed and therefore it is difficult to apply the clustering technique.

A self-organizing neural network has been used in the development of a diagnosis system based on the use of feed force and torque together with FFT based feature extraction [12]. The approach is regarded as a promising empirical modeller. The conclusion is made that a certain number of feature vector components or dimensionality of a dynamic system does exist by which the drilling process can be properly characterized. Also the classification error is studied with different numbers of features. The effect of the cutting parameters is not covered in this context.

The Restricted Coulomb Energy (RCE) network is a parallel neural network modelled after the human learning and classification process [9]. The architecture of the RCE network is a feed-forward arrangement. This allows the network to classify pattern signals in real time without any special hardware. The network is composed of three layers of cells: the input layer, the internal (hidden) layer, and the output layer. The feature vectors of the pattern to be learned are presented to the input layer. The nodes of the input layer are connected to every node of the internal layer. The nodes in the internal layer are connected selectively to the output nodes during the training process. The output nodes correspond to different pattern classes. The internal connections occur in such a way that the correct output cell will be fired when an appropriate pattern class is given to the system. RCE networks use two learning mechanisms. When new patterns are presented to the network, the response of the neural network is tested without any modification of the weight matrix. If the classification of the network matches the required output, the weight matrix is not changed. Otherwise, the second method is used and the influence of the exiting nodes are modified and/or a new node will be created. In the case of breakage detection in micro-drilling, eight input features (four average, four standard deviation) based on thrust force have been used. The RCE network correctly recognised normal and tool failure cases with a higher than 90% accuracy.

Adaptive resonance networks have been tested for the detection of severe micro-drill damage just before a complete tip breakage occurs [8]. According to adaptive resonance theory (ART), adaptive resonance occurs when the input to a network and the feedback expectancies match. ART2-type neural networks have been developed to realize a self-organized stable pattern recognition capability in real time. The ART2-type neural networks compare a given input with previously encountered patterns. If the input is similar to any of the patterns, it will be placed in the same category with similar patterns. On the other hand, if the input is not similar to any of the previously presented patterns a new category will be assigned to the given input. The sensitivity of the neural network is adjusted with the vigilance value. Two approaches have been tested: in the first, 22 wavelet coefficients, and in the second, six parameters were calculated from the original 24 coefficients to represent the information of the wavelet coefficients to the neural net. The direct encoding method with 22 coefficients was found to be slower but more reliable. The ART2-type neural networks required two to three times more computational time to classify the 22 wavelet coefficients than the six parameters of the secondary encoding method. However, there was only one classification error in 61 cases. The ART2 worked much faster with the parameters of the secondary encoding; but there were at least three estimation errors in any studied case.

5. Concluding remarks

A summary of the monitoring methods, signal analysis and diagnostic techniques for tool wear and failure monitoring in drilling has been given. In this context only indirect monitoring methods such as force, vibration and current measurements have been covered, i.e. direct monitoring methods based on dimensional measurement etc. are not included. In signal analysis statistical parameters calculated from the time domain signal are widely used. Fast Fourier and Wavelet Transform are more sophisticated means of signal analysis that have also been used for tool wear and breakage detection by a number of research groups. Only a limited number of automatic diagnostic tools have been developed for diagnosis of the condition of the tool in drilling. All of these rather diverse approaches that have been available in the literature are covered in this study. In the reported material there are both success stories and attempts that have not been so successful. Only in a few of the papers have attempts been made to compare the chosen approach with other methods, i.e. many of the papers only present one approach and unfortunately quite often the test material the study is based on is limited, especially when it comes to the cutting process parameter variation, i.e. variation of cutting speed, feed rate, drill diameter and material and also workpiece material.

References

[1] A. Thangaraj, P.K. Wright, Computer-assisted prediction of drill-failure using in-process measurements of thrust force, Journal of Engineering for Industry, Transactions of the ASME 110 (1988) 192–200.

[2] Y. Kavaratzis, J.D. Maiden, System for real time process monitoring and adaptive control during CNC deep hole drilling, in: Proceedings of Comadem '89, Kogan Page, London, UK, 1989, pp. 148–152.

[3] U. Schehl, Werkzeugüberwachung mit Acoustic-Emission beim Drehen, Fräsen und Bohren, Aachen, 1991.

[4] S.R. Hayashi, C.E. Thomas, D.G. Wildes, Tool break detection by monitoring ultrasonic vibrations, Annals of the CIRP 37 (1) (1988) 61–64.

[5] K. Kutzner, U. Schehl, Werkzeugüberwachung von Bohrern kleinen Durchmessers mit Körperschallsensoren, Industrie Anzeigner 110 (82) (1988) 32–33.

[6] S.C. Lin, C.J. Ting, Tool wear monitoring in drilling using force signals, Wear 180 (1-2) (1995) 53–60.

[7] S. Braun, E. Lenz, C.L. Wu, Signature analysis applied to drilling, Journal Mechanical Design, Transactions of the ASME 104 (1982) 268–276.

[8] I.N. Tansel, C. Mekdeci, O. Rodriguez, B. Uragun, Monitoring drill conditions with wavelet based encoding and neural network, International Journal of Machine Tools & Manufacture 33 (4) (1993) 559–575.

[9] I.N. Tansel, O. Rodriguez, C. Mekdeci, Detection of tool breakage in microdrilling operation with RCE neural networks, PED, ASME 47 (1) (1992) 83–88.

[10] C. Von Nedeß, T. Himburg, Automatisierte Überwachung des Bohrens, VDI-Z, Bd 128 (17) (1986) 651–657.

[11] K. Christoffel, W. Jung, Überwachungseinheit für die Bohrbearbeitung, Industrie Anzeigner 103 (62) (1981) 198–199.

[12] E. Govekar, I. Grabec, Self-organizing neural network application to drill wear classification, Journal of Engineering for Industry, Transactions of the ASME 116 (3) (1994) 233–238.

[13] E. Jantunen, H. Jokinen, Automated On-line Diagnosis of Cutting Tool Condition (Second version), International Journal of Flexible Automation and Integrated Manufacturing 4 (3-4) (1996) 273–287.

[14] W. König, K. Christoffel, Sensoren für die Bohrbearbeitung, Industrie Anzeigner 103 (100) (1980) 29–33.

[15] P.G. Li, S.M. Wu, Monitoring drilling wear states by a fuzzy pattern recognition technique, Journal of Engineering for Industry, Transactions of the ASME 110 (2) (1988) 297–300.

[16] T.I. Liu, K.S. Anantharaman, Intelligent classification and measurement of drill wear, Journal of Engineering for Industry, Transactions of the ASME 116 (1994) 392–397.

[17] T. Radhakrishnan, S.M. Wu, On-line hole quality evaluation for drilling composite material using dynamic data, Journal of Engineering for Industry, Transactions of the ASME 103 (1981) 119–125.

[18] M. Routio, M. Säynätjoki, Tool wear and failure in the drilling of stainless steel, Journal of Materials Processing Technology 52 (1) (1995) 35–43.

[19] K. Subramanian, N.H. Cook, Sensing of drill wear and prediction of drill life (I), Journal of Engineering for Industry, Transactions of the ASME 101 (1977) 295–301.

[20] T.I. El-Wardany, D. Gao, M.A. Elbestawi, Tool condition monitoring in drilling using vibration signature analysis, International Journal of Machine Tools & Manufacture 36 (6) (1996) 687–711.

[21] E. Lenz, J.E. Mayer, D.G. Lee, Investigation in drilling, Annals of the CIRP 27 (1) (1978) 49–53.

[22] A. Noori-Khajavi, Frequency and time domain analyses of sensor signals in a drilling process and their correlation with drill wear, PhD Thesis, Oklahoma State University, Stillwater, OK, 1992.

[23] A. Noori-Khajavi, R. Komanduri, On multisensor approach to drill wear monitoring, Annals of the CIRP 42 (1) (1993) 71–74.

[24] A. Noori-Khajavi, R. Komanduri, Frequency and time domain analyses of sensor signals in drilling, part 1: correlation with drill wear, International Journal of Machine Tools and Manufacture 35 (6) (1995) 775–793.

[25] K. Ramamurthi, C.L. Hough Jr., Intelligent real-time predictive diagnostics for cutting tools and supervisory control of machining operations, Journal of Engineering for Industry, Transactions of the ASME 115 (1993) 268–277.

[26] M. Valikhani, S. Chandrashekhar, An experimental investigation into the comparison of the performance characteristics of TiN an ZrN coatings on split point drill using the static and stochastic models of the force system as a signature, The International Journal of Advanced Manufacturing Technology 2 (1) (1987) 75–106.

[27] E. Brinksmeier, Prediction of tool fracture in drilling, Annals of the CIRP 39 (10) (1990) 97–100.

[28] R.W. Barker, G.A. Kluthe, M.J. Hinich, Monitoring rotating tool wear using higher-order spectral features, Journal of Engineering for Industry, Transactions of the ASME 115 (1993) 23–29.

[29] S. Braun, E. Lenz, Machine Tool Wear Monitoring, Mechanical Signature Analysis, Theory and Applications, pp. 321–342, Academic Press Ltd, 1986.

[30] T.I. Liu, E.J. Ko, On-line recognition of drill wear via artificial neural networks, monitoring and control for manufacturing processes, PED, ASME 44 (1990) 101–110.

[31] T.I. Liu, S.M. Wu, On-line detection of drill wear, Journal of Engineering for Industry, Transactions of the ASME 112 (1990) 299–302.

[32] A.L. Quadro, J.R.T. Branco, Analysis of the acoustic emission during drilling test, Surface & Coating Technology 94-95 (1-3) (1997) 691–695.

[33] F. Erdélyi, C. Sántha, Monitoring tasks on boring and milling production cells, Computers in Industry 7 (1986) 65–71.

[34] X. Li, On-line detection of the breakage of small diameter drills using current signature wavelet transform, International Journal of Machine Tools & Manufacture 39 (1) (1999) 157–164.

[35] R.B. Randall, Application of B&K Equipment to Frequency Analysis, ISBN 87 87355 14 0, Brüel & Kjær, 1977.

[36] E. Jantunen, A solution for tool wear diagnosis, in: Proceedings

madem'99, Coxmoor Publishing Company, Oxford, UK, pp. 95–104.
tunen, H. Jokinen, Reduction of data needed in an expert ı for condition monitoring of FMS using regression analyhniques, in: Proceedings of Comadem'96, ISBN 1-85075-, University of Sheffield, Sheffield, 1996.
.tunen, H. Jokinen, R. Milne, Flexible expert system for automated on-line diagnosis of tool condition, in: Integrated Monitoring Diagnostics & Failure Prevention, Technology Showcase, 50th MFPT, Mobile, Alabama, 1996, pp. 259–268.
[39] A. Noori-Khajavi, R. Komanduri, Frequency and time domain analyses of sensor signals in drilling, part 2: investigation of the problems associated with sensor integration, International Journal of Machine Tools & Manufacture 35 (6) (1995) 795–815.

PUBLICATION II

Automated on-line diagnosis of cutting tool condition

In: International Journal of Flexible Automation and Integrated Manufacturing 1996. 4 (3&4), pp. 273–287.
Reprinted with permission from the publisher.

International Journal of Flexible Automation and Integrated Manufacturing, 4(3&4): 273–287 (1996)

Automated On-Line Diagnosis of Cutting Tool Condition

Erkki Jantunen* and Harri Jokinen

Technical Research Centre of Finland, Manufacturing Technology, P. O. Box 1702, FIN-02044 VTT, Finland

* To whom all correspondence should be addressed.

ABSTRACT

A machine tool utilization rate can be improved by an advanced condition-monitoring system using modern sensor- and signal-processing techniques. This paper gives a summary of a wide cutting test and analysis program for indirect tool-wear measurement. The test program covered both shank end and end mills with twist drills and tread taps. For monitoring tool wear, we tested vibration, acoustic emission, sound, spindle power, and axial force. The signals were analyzed in time domain using statistical methods such as root-mean-square value, standard deviation, maximum, and skewness. The signals were further analyzed using Fast Fourier Transform to find their frequency contents. The effectiveness of the best sensors and analysis methods has been verified with a program for predicting the remaining lifetime of a tool in use. The results show that vibration, sound, and acoustic emission measurements are more reliable for tool-wear monitoring than are the more commonly used power consumption, current, and force measurements.

KEYWORDS: condition monitoring, expert systems, flexible manufacturing systems, signal analysis, tool wear monitoring.

I. INTRODUCTION

An increasing number of flexible manufacturing systems (FMS) have been installed in Europe in the past few years. However, the availability of the installed FMS is not as high as was originally expected, and, in particular, unmanned use in three shifts has not been successful.[1] One of the most important reasons for this is that existing real-time tool-condition monitoring techniques do not cover the wide range of machining situations with different machining parameters that normally exist in practice.

For untended machining, process and tool-condition monitoring (tool identification, tool-wear monitoring, and tool-breakage detection) have great potential for increasing the capacity of the machine-tool systems. Automated on-line diagnosis refers both to machine-tool monitoring and to workpiece and tool-condition monitoring. Existing real-time condition- and tool-monitoring techniques for machine tools do not yet provide satisfactory information.

1064-6345/96 $5.00

Because of the poor correlation between the measured event, such as tool wear, and the sensor signal, there is no one measuring method that covers the wide range of situations with different machining parameters. (In addition, machining-process monitoring systems and maintenance modules have only limited commercial availability.) Signals must be grouped and synchronized to avoid poor correlation between a single signal source and the measured event. However, it is obvious that commercially available monitoring systems exploit few of the capabilities of modern sensor and analyzing techniques.[2]

The sensor validation in this survey is based on comprehensive laboratory tests. Several sensors were installed and tested before final validation. The aim was to monitor the condition of the tool at critical points in the machining and cutting process. Validation was performed according to the following criteria: sensitivity of the sensor to the measured event, correlation between signal and measured event, amount of deviation, and universality.

As part of the validation in the laboratory, measured signals from various sensors were analyzed using a number of methods in both time and frequency domains. The objective was to satisfy the economic and technical requirements of the industry. Suitable combinations of sensor and analysis methods are listed in Table 1. The information received from multiple sensors was analyzed simultaneously; thus, the correlation between different measuring signals can be used to localize a faulty component or identify wear. The relationships between the analyzed signals and wear form a basis for diagnostic rules that can be used in an AI system.

II. TEST ARRANGEMENT

A horizontal-type machining center was used in the cutting tests. The main specification of the machining center was as follows:

Machine tool: Niigata EN40B
Control unit: Fanuc 11 MA
Controlled axis: 4 axes (X, Y, Z, and B)
Table size: 400 × 400 mm
Spindle nose: NT No. 40 for BT
Number of tools: 30
Spindle speed: 15–6000 $^{1}/\text{min}$
Main motor power: 11/7.5 kW

II.1. Cutting Test Program

The tests concentrated on tool wear, tool breakage, and collision monitoring. Cutting tools of different types and shapes were used to learn machine tool behavior in the cutting process. Cutting tests were designed to cause situations in which a measurable event resulted from tool wear or failure. In this part of the cutting tests, different types of cutting tools were used to cover a wide range of cutting methods. The tools investigated in the tests were:

- *shank end mill,* diameter 6 and 10 mm, HSS
- *end mill,* diameter 50 mm with carbide inserts

TABLE 1. Description of Tools, Cutting Parameters and Monitoring Methods Used in the Cutting Tests*

	Shank end mill	End mill	Twist drill	Tread tap
Number of tools	18	8	26	27
Tool sizes	6 and 10 mm	50 mm	3.3, 5.0, 6.8, 8.5, 10.2 mm	M4, M6, M8, M12
Cutting speed, [m/min.]	78.5 and 85	250	22, 25, 29, 30, 32, 35, 37, 38, 40	6, 10, 12, 15, 18, 20
Table feed [mm/min.]	230, 270, 400	1300	140, 250, 272, 300, 350, 375, 448, 560, 620, 680	336, 478, 533, 640, 720, 864, 960
Depth of cut/Hole depth	3 and 5 mm	3 mm	27 and 30 mm (bottom hole) 35 mm (through hole)	10 and 20 mm (bottom hole) 40 mm (through hole)
Width of cut	5, 8, 9 mm	43 mm	-	-
Acoustic emission 200 kHz				
Acoustic emission 800 kHz				
Acoustic emission 100 - 1000 kHz				
Horizontal vibration				
Vertical vibration				
Sound				
Sound parabolic				
Niigata spindle power				
Spindle power				
Spindle current (P1, P2 & P3)				
Z-servo current (I1, I2 & I3)				
Force quideways				
Force dynamometer, Fz				
Torque dynamometer, Mz				
3-axis table dynamometer				
Tachometer				
X-servo power				
X-servo current				
X-servo voltage				

[a] Gray areas indicate combinations of tool-type and monitoring method used

- *twist drill,* diameter 3.3, 5.0, 6.8, 8.5, and 10.2 mm, HSS
- *thread tap* M4, M6, M8, and M12, HSS

Refer to Table 1 for a detailed description of the tools, cutting parameters, and monitoring methods.

II.2. Measuring Arrangement

The tool-wear monitoring methods investigated in the tests were mainly based on acoustic emission, vibration (acceleration), spindle and feed-drive power consumption, force, torque, and sound, using different measuring arrangements. Some of the measuring signals were kept the same and some were varied during the tests, because the instrument tape recorder only has 14 channels, of which 12 can normally be used for measuring data. Refer to Table 1 for information on the use of different measuring signals. Figure 1 illustrated the main configuration of the measuring arrangement.

Acoustic emission was measured by using one sensor with center frequency of 200 kHz (frequency range ± 20 kHz) for tool-wear monitoring, and another sensor with center frequency of 800 kHz (frequency range ± 50 kHz) for tool-breakage monitoring.[3] In addition, a broadband sensor with a frequency range of 100 kHz–1 MHz was used to find the best frequency range for further investigation. The 200 kHz sensor was used to measure the root-mean-square (RMS) values of acoustic emission, and the 800 kHz and 100 kHz–1 MHz broadband sensors were used to measure peak values of acoustic emission.

Vibration was measured in the frequency range of 0.2–5000 Hz. Two acceleration transducers were installed, one vertical and one horizontal, and were kept in the same measuring positions throughout the tests. Two microphones were installed—one normal and one with a parabola. The microphone with the parabola had a small direction angle intended to listen only to noise caused by the machining process.

Niigata's spindle-power consumption was measured through Niigata's own measuring system. Spindle-power consumption was also studied using separate measuring

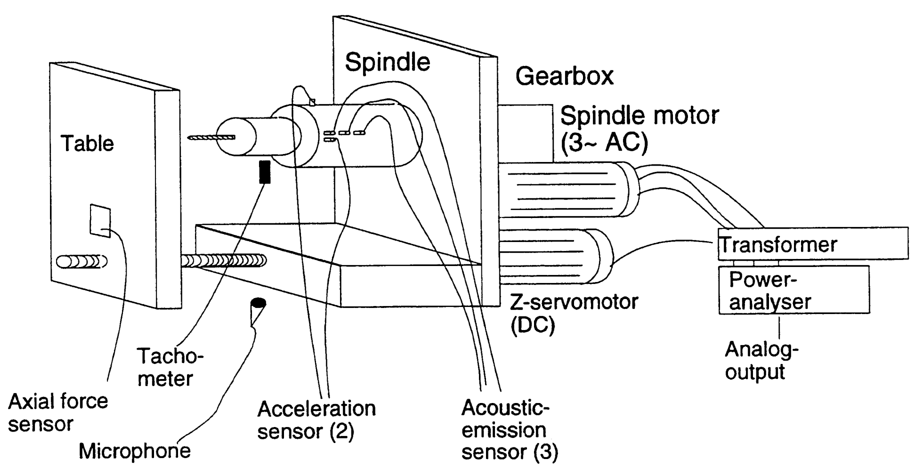


FIGURE 1
Measuring arrangement.

equipment, which measured power, current, and voltage. One channel was used to measure spindle-current consumption (RMS) and another to measure AC spindle current. Both X- and Z-servo power consumption were measured with separate measuring equipment. The measured values in these servo motors were power (P1), current (I1, I2), and voltage. Z-servo power/current consumption was mainly used to test drills and thread taps, X-servo power/current consumption to test shank-end mills and end mills.

The force sensor was installed in the guideways and measured the cutting force, mainly in the Z-axis direction. The 3-D dynamometer was used to measure the x, y, z forces in the tests, the 2-D dynamometer to measure axial force Fz and torque Mz in some of the drilling and tapping cases. These tests were intended to yield accurate information about cutting forces during the lifetime of the tool.

III. SIGNAL PROCESSING

For automatic analysis of the huge amount of data gathered, a user-friendly interface for a PC was created using the Visual Basic programming environment. The interface controlled functions of the instrument tape recorder through an IEEE-488 bus. Statistical analysis was based on data from a data acquisition board (maximum 16 A/D channels). Attached to the acquisition board, a sample and hold board was used to get synchronized dynamic signals from 4 channels simultaneously. The tacho pulse was used as a trigger for data collection.

The measured signals were analyzed with a number of methods in time and frequency domains. The analysis program enabled automatic analysis outside normal working hours, and the analysis results were stored in databases. A mathematical programming package, MatLab, was used in signal processing the dynamic signals, gathering only so-called "cursor values" to minimize the amount of information to be transferred and stored. These actions were also controlled by the interface.

III.1. Statistical Analysis

Depending on the measured events, the data to be analyzed were first cleaned of irrelevant signals that had not been recorded during the actual machining process, such as rapid movement during drilling. Cleaning was based on a master signal, according to which all the other signals were either accepted or omitted from the analysis. Data measured and recorded simultaneously with 12 sensors were studied by calculating a number of statistical parameters—arithmetic mean, RMS, mean deviation, standard deviation, skewness, kurtosis, maximum, and minimum. The two acceleration signals were analyzed using both low-pass filtering and no filtering. The results from the statistical and dynamic signal were saved in a separate database for every tool. Figure 2 shows an example of this statistical analysis.

III.2. FFT Analyses

For dynamic signals containing frequency information, Fast Fourier Transformation (FFT) techniques were employed.[4] Sample and hold functions of the data acquisition option board were used to get data simultaneously from 4 channels (vibration, force/torque, spindle-motor power, and sound), and the data was stored on disk. MatLab was used to read the data from the hard disk and to perform the necessary calculations. FFT was performed using both time and frequency domain averaging. About 20 kinds of analysis functions, such as spectrum, cepstrum, frequency response, and coherence, were calculated by FFT using the hanning window. The 20 highest maximums and/or minimums of each of these functions were tabulated and stored in a database with the corresponding frequencies and time values. The analyzed functions are shown in Table 2.

IV. RESULTS

The results from the statistical and FFT analyses were further analyzed using regression analysis techniques. Eight regression functions were tested to find the highest correlation between measured tool wear and analyzed measurement signals. Finally, the measuring sensors and analysis methods were tabulated and sorted according to the order of sensitivity of the sensor to the measured event, correlation between the signal and measured event, amount of deviation, and universality.

IV.1. Regression Analysis

The relationship between analyzed signals and tool wear was tested with 8 functions, which approximate the set of data points gathered from the cursor values of the spectrum analyses or from the statistical values of the measured signals. The sets of data points were then approximated as closely as possible with 4 smoothing functions—3 polynomials and 1 logarithmic function based on a simplified mathematical definition of wear[5]) using the least square principle.

$$Y = a + bT \tag{1}$$

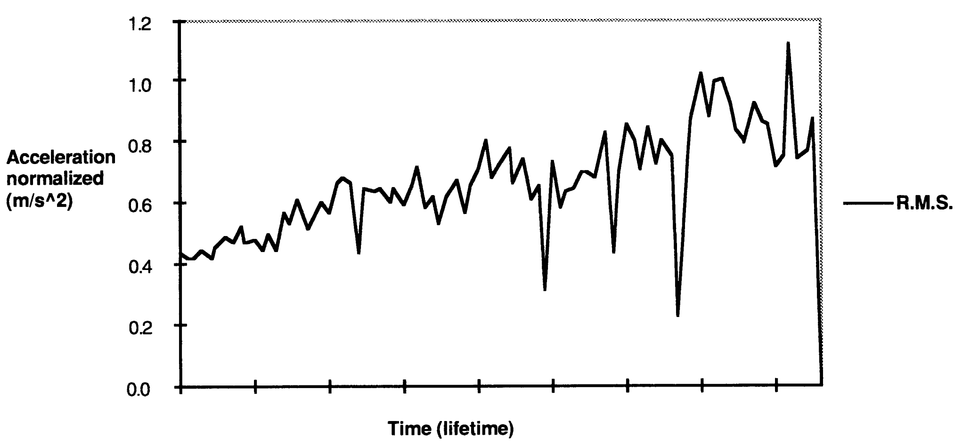


FIGURE 2
Root-mean-square (RMS) value of horizontal vibration acceleration for shank end mill.

$$Y = a + bT + cT^2 \quad (2)$$

$$Y = a + bT + cT^2 + dT^3 \quad (3)$$

$$Y = a + b \ln\left(\frac{T_{total}}{T_{total} - T}\right) \quad (4)$$

where $a, b, c, \& d$ = regression coefficients; T = time; $Ttotal$ = total tool lifetime in the test.

In the regression analysis, we also tested the weighting of each residual using a geometric series (1.0, 1.1, 1.21, 1.331 ...), which doubled the number of regression functions. To measure the goodness of the regression, the coefficient of determination (Ri) was calculated for each regression or curve fitting based on the sum of the squares of deviations about the regression (SSD) and the total sum of squares (SST) in the following way:

$$Ri = 1 - SSD / SST \quad (5)$$

TABLE 2. Functions Used in FFT Analyses

Functions used	Number of cursor values saved from each signal analyzed			
	Maximum	Minimum	Octaves	Harmonics
Spectrum	20			
Cross-spectrum	20			
Frequency response	20			
Coherence	20			
Coherent output power	20			
Autocorrelation	20	20		
Cross-correlation	20	20		
Signal-to-noise-ratio	20			
Cepstrum	20			
Liftered spectrum	20			
Scot	20			
1/3 Octave			36	
1/1 Octave			25	
Overall			1	
Six first harmonics and 1/2 and 1/4 harmonics				8
Total harmonic power				1
Multisignal frequency response 1. component	20			
Multisignal frequency response 2. component	20			
Multisignal frequency response 3. component	20			
Partial coherence	20			

The Ri and its product for each regression were stored in a database for further investigation. The regression was carried out for subsets of observations along the frequency axis or along the time axis starting from the origin. The frequency or time window was divided into 500 subsets. The frequency or time subset windows were accepted for regression analysis if there were more than 20 observations (refer to Table 2) in the window during the test of the specific tool. An example of an FFT analysis is shown in Figure 3.

Regression analyses were performed with all the information stored from the statistical and FFT analyses. For the statistical analyses, the results of the regression analyses were further analyzed by making so-called "ranking lists," in which the signals were sorted by descending coefficient of multiple determination of the third-order regression function (see Equation 3). The combined ranking list for all the tested tools and tool types is shown in Table 3, in which the statistical parameters are listed in order of their appearance in the ranking lists. (Note that those signals that were tested with only one tooltype are omitted from this list.)

Table 4 shows an example of the goodness of fit of the regression function (third order, Equation 3) of the FFT functions for different tool types. In this case, averaging was done in the frequency domain. In Table 4, each function is shown only once (the highest coefficient of determination), so it is a much shortened version of a complete list in which the total number of lines is in the order of 2000 for each tool and for the 2 averaging methods.

The degree of fit is much higher for cursor values of the FFT functions than for the statistical parameters. This result is logical, because the idea of the FFT analysis is to separate meaningful from meaningless information (noise). This result suggests that it would be good to use FFT analysis for tool monitoring. However, FFT analysis takes time, more if averaging is used to reduce the amount of noise, which makes it impossible to use for collision and tool-breakage monitoring but suitable for tool-wear monitoring.

Table 4 shows that the goodness of fit varies depending on the tool type. Drilling and shank-end milling are the easiest to monitor. It should be noted that such functions as cross-spectrum and coherence, which describe how much the signals are related to each other, show a rather high goodness of fit. One explanation for this is that the worn tool causes higher signals, and the signal-to-noise relation increases, so this type of function can be used for tool-wear monitoring. This suggests that it would be good practice to use

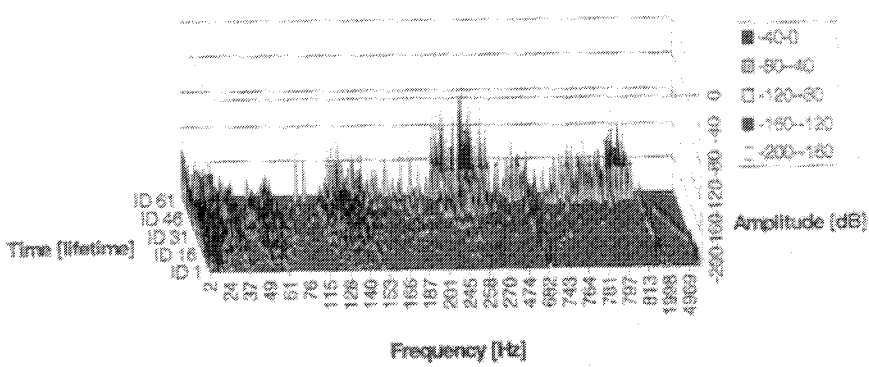


FIGURE 3
Cross spectrum of horizontal acceleration and spindle-motor current in drilling. (Note that the frequency axis is nonlinear.)

TABLE 3. Ranking (Ra) of Statistically Analyzed Signals for Different Tool Types[a]

Signal	Drill Ra	Drill Statistical parameters	Tread tap Ra	Tread tap Statistical parameters	Shank end mill Ra	Shank end mill Statistical parameters	End mill Ra	End mill Statistical parameters
Vibration horizontal (acceleration low-pass filtering)	1	3, 2, 6, 4, 7, 8, 5, 1			2	3, 2, 5, 6, 4, 8, 7, 1	16	8, 7, 3, 2, 6, 4, 5, 1
Axial force sensor (N)	2	3, 8, 6, 7, 4, 1, 5, 2	7	8, 2, 1, 4, 3, 7, 6, 5	-		-	
3-dimensional dynamometer X-axis, (N)	3	8, 7, 1, 2, 3, 4, 6, 5	4	7, 3, 2, 4, 8, 1, 5, 6	5	2, 8, 7, 1, 3, 4, 5, 6	-	
Vibration vertical (acceleration low- pass filtering)	4	3, 2, 6, 4, 8, 7, 5, 1			6	3, 5, 2, 6, 4, 7, 8, 1	15	7, 8, 6, 2, 3, 5, 4, 1
Acoustic emission 200 kHz/filter 180 (rms.)	5	1, 2, 8, 7, 3, 5, 4, 6	6	8, 1, 2, 7, 3, 4, 6, 5	23	7, 4, 8, 2, 6, 1, 3, 5	17	7, 2, 8, 1, 4, 3, 5, 6
Vibration horizontal (acceleration)	6	3, 2, 8, 4, 7, 6, 1, 5	13	8, 2, 7, 3, 4, 1, 6, 5	3	2, 3, 8, 7, 4, 5, 6, 1	19	8, 7, 3, 2, 4, 6, 1, 5
3-dimensional dynamometer Y-axis, (N)	7	8, 7, 1, 2, 4, 3, 6, 5	5	7, 4, 3, 1, 8, 2, 5, 6	4	7, 2, 1, 3, 8, 4, 6, 5	-	
Vibration vertical (acceleration)	8	3, 2, 7, 8, 4, 6, 1, 5	12	7, 8, 3, 2, 4, 1, 6, 5	7	2, 8, 3, 7, 6, 5, 4, 1	13	7, 8, 2, 3, 4, 1, 5, 6
Sound (normal microphone, AC)	10	3, 2, 7, 8, 4, 6, 1, 5	14	8, 1, 2, 7, 3, 4, 5, 6,	1	2, 3, 8, 7, 4, 6, 5, 1	-	
Z-servo current IT3 (AC, A)	11	1, 2, 8, 3, 7, 4, 6, 5	10	5, 1, 2, 8, 6, 7, 4, 3	-		-	
Force sensor (guideways, N)	12	7, 1, 8, 2, 4, 3, 6, 5	1	1, 2, 8, 7, 5, 6, 4, 3	16	3, 8, 4, 5, 2, 1, 7, 6	4	1, 7, 8, 2, 4, 6, 5, 3
3-dimensional dynamometer Z-axis, (N)	13	3, 1, 7, 2, 8, 5, 4, 6	3	8, 3, 4, 2, 1, 7, 6, 5	14	3, 5, 6, 7, 4, 8, 1, 2	-	
Spindle power P1 (rms, W)	14	2, 8, 7, 1, 3, 6, 4, 5	16	7, 3, 4, 2, 1, 5, 8, 6	24	6, 4, 2, 3, 7, 1, 8, 5	8	1, 2, 4, 7, 8, 6, 5, 3
Sound (parabolic microphone)	15	3, 2, 4, 8, 7, 6, 1, 5	9	3, 4, 2, 8, 7, 1, 5, 6	10	2, 3, 7, 8, 6, 4, 1, 5	-	
Acoustic emission 0 kHz-2 MHz/800 kHz peak	16	8, 1, 2, 7, 3, 4, 6, 5	2	1, 2, 8, 7, 3, 4, 5, 6	-		-	
Spindle current I1 (rms, A)	17	1, 8, 2, 6, 7, 4, 3, 5	11	7, 5, 3, 1, 8, 2, 4, 6	13	8, 7, 1, 6, 2, 4, 5, 3	6	1, 2, 6, 7, 4, 8, 5, 3
Z-servo current I3 (rms, A)	18	1, 2, 7, 8, 3, 4, 6, 5	18	8, 5, 7, 4, 3, 1, 2, 6	26	4, 6, 3, 1, 8, 5, 2, 7	-	
Niigata Spindle power (%)	19	8, 7, 2, 5, 1, 6, 4, 3	15	1, 2, 5, 7, 8, 3, 4, 6	17	7, 8, 1, 2, 3, 4, 6, 5	11	7, 4, 2, 1, 8, 6, 3, 5
Spindle current IT1 (AC, A)	20	2, 1, 8, 3, 7, 5, 6, 4	8	8, 7, 2, 4, 3, 1, 5, 6	15	1, 6, 8, 5, 3, 7, 2, 4	10	3, 7, 8, 4, 1, 5, 6, 2
Acoustic emission 400 kHz-2 MHz/400 kHz peak	21	8, 7, 3, 5, 4, 6, 2, 1	17	6, 8, 7, 5, 2, 1, 3, 4			-	
Spindle power P3 (rms, W)	-		-		19	2, 7, 4, 3, 8, 6, 1, 5	12	1, 2, 4, 6, 7, 8, 3, 5
Spindle current I2 (rms, A)	-		-		20	6, 8, 7, 1, 2, 5, 3, 4	7	8, 7, 1, 2, 4, 3, 6, 5
Spindle power P2 (rms, W)	-		-		25	2, 4, 8, 7, 6, 3, 1, 5	9	1, 2, 4, 7, 8, 6, 3, 5

[a]Statistical parameters: 1. Arithmetic mean 2. Root mean square 3. Mean deviation 4. Standard deviation 5. Skewness 6. Kurtosis 7. Maximum 8. Minimum

TABLE 4. Coefficients of Determination of Signals Analyzed with FFT Averaged in Frequency Domain for Different Tool-Types

Signal(s) and FFT function	Drill	Thread tap	Shank end mill(*	End mill(*
Spectrum of horizontal vibration (acc.)	0.99	0.85	0.98	0.83
Liftered spectrum of vertical vibration (acc.)	0.99	0.41	0.58	0.81
Spectrum of vertical vibration (acc.)	0.99	0.79	0.92	0.84
Cross-spectrum between horizontal vibration and vertical vibration (acc.)	0.99	0.92	0.95	0.83
Coherent output power between horizontal vibration and vertical vibration (acc.)	0.98	0.87	0.96	0.88
Autocorrelation of vertical vibration (acc.) (max.)	0.94	0.73	0.99	0.74
Cross-correlation between horizontal vibration and vertical vibration (acc.) (max.)	0.91	0.67	0.99	0.75
Coherence between horizontal vibration (acc.) and sound (normal microphone)	0.88	0.53	0.70	0.76
Spectrum of spindle current I1 (AC)	0.87	0.71	0.84	0.77
Autocorrelation of horizontal vibration (acc.) (max.)	0.85	0.86	0.99	0.88
Signal-to-noise-ratio of horizontal vibration (acc.) and sound (normal microphone)	0.84	0.87	0.94	0.60
Coherent output power between horizontal vibration (acc.) and sound (normal m.)	0.83	0.57	0.55	0.59
Spectrum of sound (normal microphone)	0.82	0.55	0.67	0.81
Cepstrum of vertical vibration (acc.)	0.82	0.57	0.53	0.90
1/1 Octave of spindle current I1 (AC)	0.80	0.50	0.43	0.88
1/3 Octave of spindle current I1 (AC)	0.80	0.62	0.77	0.88
Cepstrum of sound (normal microphone)	0.80	0.51	0.54	0.51
Cross-correlation between horizontal vibration and vertical vibration (acc.) (min.)	0.80	0.63	0.98	0.54
1/3 Octave of sound (normal microphone)	0.79	0.61	0.92	0.86
Cross-spectrum between horizontal vibration (acc.) and sound (normal microphone)	0.79	0.73	0.59	0.80
Autocorrelation of spindle current I1 (AC) (min.)	0.79	0.24	0.20	0.32
Autocorrelation of spindle current I1 (AC) (max.)	0.79	0.27	0.20	0.32
Overall of spindle current I1 (AC)	0.77	0.09	0.03	0.33
Cepstrum of horizontal vibration (acc.)	0.76	0.52	0.70	0.92
Cross-correlation between horizontal vibration (acc.) and sound (normal m.) (min.)	0.75	0.52	0.58	0.32
Autocorrelation of sound (normal microphone) (max.)	0.75	0.74	0.19	0.38
Autocorrelation of vertical vibration (acc.) (max.)	0.74	0.78	0.996	0.88

TABLE 4. (Continued) Coefficients of Determination of Signals Analyzed with FFT Averaged in Frequency Domain for Different Tool-Types

Signal(s) and FFT function	Drill	Thread tap	Shank end mill(*	End mill(*
Cross-correlation between horizontal vibration (acc.) and sound (normal m.) (max.)	0.74	0.06	0.74	0.35
Signal-to-noise-ratio of horizontal vibration (acc.) and vertical vibration (acc.)	0.73	0.95	0.80	0.45
Liftered spectrum of sound (normal microphone)	0.72	0.27	0.56	0.51
Coherence between horizontal vibration (acc.) and vertical vibration (acc.)	0.71	0.68	0.61	0.32
1/1 Octave of sound (normal microphone)	0.71	0.47	0.94	0.83
1/1 Octave of vertical vibration (acc.)	0.70	0.84	0.96	0.80
Autocorrelation of sound (normal microphone) (min.)	0.70	0.68	0.71	0.43
Cross-spectrum between horizontal vibration (acc.) and spindle current I1 (AC)	0.67	0.48	0.52	0.45
Cross-correlation between horizontal vibration and spindle current I1 (AC) (min.)	0.65	0.52	0.63	0.37
Autocorrelation of horizontal vibration (acc.) (min.)	0.62	0.78	0.89	0.88
Coherence between horizontal vibration (acc.) and spindle current I1 (AC)	0.61	0.59	0.70	0.79
Liftered spectrum of spindle current I1 (AC)	0.58	0.12	0.27	0.89
1/3 Octave of vertical vibration (acc.)	0.58	0.86	0.96	0.88
Coherent output power between horizontal vibration and spindle current I1 (AC)	0.56	0.73	0.99	0.44
Cross-correlation between horizontal vibration and spindle current I1 (AC) (max.)	0.55	0.54	0.56	0.28
1/3 Octave of horizontal vibration (acc.)	0.55	0.94	0.95	0.92
1/1 Octave of horizontal vibration (acc.)	0.55	0.94	0.96	0.91
Signal-to-noise-ratio of horizontal vibration (acc.) and spindle current I1 (AC)	0.53	0.48	0.62	0.47
Cepstrum of spindle current I1 (AC)	0.52	0.42	0.79	0.88
Overall of sound (normal microphone)	0.49	0.55	0.15	0.32
Liftered spectrum of horizontal vibration (acc.)	0.46	0.33	0.70	0.87
Overall of vertical vibration (acc.)	0.33	0.36	0.97	0.01
Overall of horizontal vibration (acc.)	0.25	0.76	0.82	0.71

at least two monitoring methods, for which it would be possible to adopt the FFT analysis. The use of at least two signals for tool-wear monitoring coincides with the findings of the statistical analysis, because there is variation between the tool types.

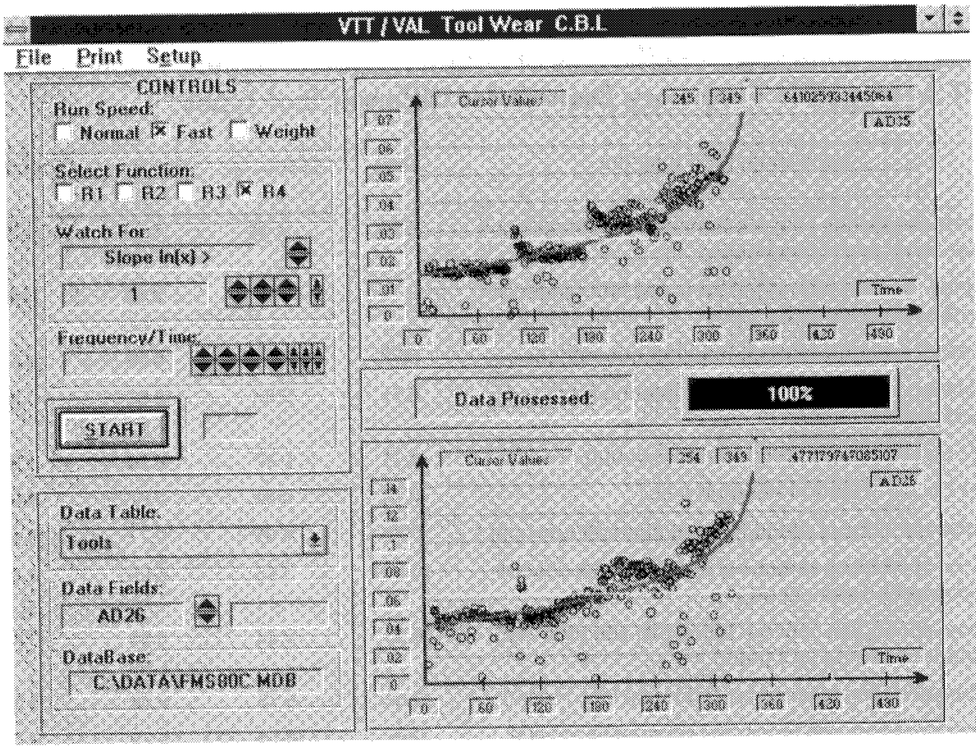


FIGURE 4
User-interface window of the simulation program, end milling: upper curve: horizontal vibration, mean deviation; lower curve: sound, root mean square.

IV.2. Simulation Program

One of the most critical tasks in maintenance (and one of the main goals of the project) is to detect faults and determine what actions need to be taken and when. This work is focused on the problems of automatically identifying the condition of the cutting process. The simulation program module is used to see how an expert system would react in different situations with different limit values for the chosen regression models. The module simulates the real machining process by using real data recorded and analyzed earlier. The program fits a selected regression curve to the existing data, the fitted curves being the same as those used in the regression analyses. Each time the curve fitting is done, the program checks whether the tool worn-out limit has been reached, and if so, a warning is displayed as a sign for when the machining process would have been terminated by the expert system.

In order to run the program, the user must define the following variables in the program window: name of the database, data table, and data fields; running speed for the program (normal = real time, fast = testing); which curve to fit into existing data; type of tool; worn-out limit and its value; which function to look for; and—for FFT analysis—the frequency range or time value.

Figure 4 shows an example of the simulation program user-interface window. In the near future, the simulation program, together with the results from the regression analyses, will be used to determine the most suitable and efficient analysis functions and the corresponding test/rule limits in an expert system.

All statistical signals were studied with the simulation program. After each signal was evaluated visually, the results indicated that when analyzing drills using acoustic emission and vibration, an expert system should be able to monitor tool wear. To improve accuracy and to compensate for dependency on tool size, the monitored signals should include at least one of the following signals: Z-servo current, microphone, or 3-D dynamometer.

It is clearly seen by visual evaluation that some of the calculated statistical parameters are not reliable. A total of 8 parameters were analyzed, and only 4 of them were found good for monitoring tool wear; disregarded from further examination were standard deviation, skewness, kurtosis, and minimum. The reason for higher-order parameters such as skewness and kurtosis being unreliable in a wide test series could be that they are too sensitive to noise, which means they work well sometimes and sometimes not at all.

The fully analyzed results indicate that the 8 signals shown in Table 5 with combinations of 4 statistical parameters are the best to monitor tool wear for drills. A significant dependence on the tool's size is noticed with the drills. With drills, the coefficients of determination of tested signals were approximately 0.25–0.84 (third-order regression).

For shank-end milling, all the signals were also studied with the simulation program. After visual evaluation of each signal, the results indicate that, when analyzing the shank-end mills using vibration, sound, and 3-D dynamometer, an expert system should be able to accurately monitor tool wear. Visual evaluation clearly shows that the conclusions regarding the number of useful statistical parameters that were drawn for drills also apply to the shank-end mills. The fully analyzed results indicate that the 7 signals shown in Table 6 with combinations of 4 statistical parameters are the best to monitor tool wear for shank-end mills. With shank-end mills, no relevant dependence on tool size was noticed, and the coefficients of determination of tested signals were approcximately 0.12–0.87 (third-order regression).

Based on the ranking list, Table 4, and visual evaluation, monitoring of tool wear in the case of end mills and thread taps is more difficult than in the case of drilling and

TABLE 5. Summary of the Best Signals and Statistical Parameters for Drilling[a]

	Signal	Statistical parameter	Suitable especially for size
1	Vibration horizontal (acceleration. low-pass filtering)	3, 2, 7	6.8 and 10.2
2	Vibration vertical (acceleration. low-pass filtering)	3, 2, 7	6.8 and 10.2
3	Acoustic emission 200 kHz/filter 180 (rms)	1, 2, 7, 3	all tested sizes
4	Vibration horizontal (acceleration)	3, 2, 7	all tested sizes
5	Vibration vertical (acceleration)	3, 2, 7	all tested sizes
6	Sound (normal microphone, AC)	3, 2, 7, 1	5, 6.8 and 10.2
7	3-dimensional dynamometer Z-axis, (N)	3, 1, 7, 2	3.3, 5 and 6.8
8	Z-servo current I3 (rms, A)	1, 2, 7, 3	3.3, 5, 6.8 and 10.2

[a] Statistical parameters: 1. Arithmetic mean 2. Root mean square 3. Mean deviation 7. Maximum. (Note: low pass filtered acceleration was only tested with 6.8 and 10.2 mm drills and 3-dimensional dynamometer was not tested with 10. 2 mm drill)

TABLE 6. Summary of the Best Signals and Statistical Parameters for Shank End Milling[a]

	Signal	Statistical. parameter.	Suitable especially for size
1	Sound (normal microphone, AC)	2, 3, 7	all tested sizes
2	Vibration horizontal (acceleration low-pass filtering)	3, 2, 7	all tested sizes
3	Vibration horizontal (acceleration)	2, 3, 7	all tested sizes
4	3-dimensional dynamometer Y-axis, (N)	7, 2, 1, 3	all tested sizes
5	3-dimensional dynamometer X-axis, (N)	2, 7, 1, 3	all tested sizes
6	Vibration vertical (acceleration low-pass filtering)	3, 2, 7	all tested sizes
7	Vibration vertical (acceleration)	2, 3, 7	all tested sizes

[a] Statistical parameters: 1. Arithmetic mean 2. Root mean square 3. Mean deviation 7. Maximum

shank-end milling. Although the coefficients of determination are reasonable, 0.36–0.74 in end milling and 0.21–0.86 for thread taps, the visual observations with the simulation program do not suggest that tool-wear monitoring could be done reliably with statistical parameters. Use of FFT improves reliability, but at this stage it has not been fully tested with the simulation program to determine whether the functions are reliable in practice.

V. CONCLUSION

There is great potential for improving machine-tool utilization rates with an advanced condition-monitoring system using modern sensor and signal-processing techniques. A comprehensive cutting-test procedure has been carried out with different tools and measurement arrangements. The recorded signal information has been processed in several ways, both in time and frequency domains. The effectiveness of the best sensors and analysis methods in the prediction of the remaining lifetime of a tool in use has been verified with a developed program module. The results from the statistical analysis show that vibration, sound, and acoustic emission measurements are more reliable for tool-wear measurement than the most common methods—power consumption, current, and force measurements—used in commercially available systems. Even better results are obtained with FFT methods, especially when at least 2 signals are available for frequency analysis. The relationships between the analyzed signals and tool wear form a basis for the diagnosis rules that are used in a diagnostic expert system.

ACKNOWLEDGMENTS

The results of the tool wear tests are used in development of an integrated condition and machining process monitoring system for flexible manufacturing systems and for stand-alone NC-machine tools (FMS Maint System, 1993-1996, EUREKA MAINE EU 744 project).

The work of the project team—Markku Karhu, Heikki Komulainen, Pasi Kontkanen, Antti Poikonen, Risto Vannela, and Kirsi Vähä-Pietilä—is gratefully acknowledged.

REFERENCES

1. **Lakso, T.,** The Influence of FMS technology on the efficiency of NC-controlled machine tools, *Tampere University of Technology Publications,* 50, 106, 1988.
2. **Kluft, W.,** Techniken zur Werkzeugüberwachung auf der Basis von Kraft-und Verformugsmessung, *Sonderdruck aus der Vortagsreihe des VDI-Seminars "Überwachung und Diagnose von Maschinen und Prozessen,"* Aachen, Germany, 1985.
3. **Kakino, Y.,** In-process detection of tool breakage by monitoring acoustic emission, *Cutting Tool Materials,* Ohio, 1980, in *ASM,* 5, 29-34, 1981.
4. **Randall, R. B.,** *Application of B&K Equipment to Frequency Analysis,* Brüel & Kjær, Nærum, Denmark, 239.
5. **Jantunen, E. and Poikonen, A.,** Dynamics in monitoring gear faults, *Proceedings of the 47th Meeting of the Mechanical Failures Prevention Group,* Virginia Beach, Virginia, April 13–15 1993, 279-288.

PUBLICATION III

The applicability of various indirect monitoring methods to tool condition monitoring in drilling

In: International Journal of Comadem 2004.
Vol. 7, No. 3, July, pp. 24–31.
(Also published in COMADEM 01,
September 4–6, Manchester, UK.)

The Applicability of Various Indirect Monitoring Methods to Tool Condition Monitoring in Drilling

Erkki Jantunen
Technical Research Centre of Finland, Manufacturing Technology
P.O.Box 1702, FIN-02044 VTT, Finland

Erkki Jantunen, (Lic.Tech.) is a senior research scientist who has been working in the field of vibration and structural analysis, development of condition monitoring methods and diagnostic expert systems for many years. At present he heads the machine operability research group at the Technical Research Centre of Finland.

Abstract

Condition monitoring of cutting tools is important for a number of reasons. The unmanned use of flexible manufacturing systems is not possible without a reliable system for tool condition monitoring. Tool wear affects the surface quality of processed workpieces. Tools cannot be optimally used based on tool change policy which relies on time and which easily leads to too frequent change of tools from which it follows that valuable production time is lost and the tool cost becomes high. There is great variation in how well different monitoring methods work in tool condition monitoring. It is well known and accepted that cutting forces increase as a function of tool wear and consequently thrust force and torque are often monitored in drilling. Feed drive and spindle current actually also measure the same thing as feed force and torque transducers although through a longer measuring chain. Tool wear also changes the dynamics of cutting processes and consequently drift forces, vibration and sound have been used for tool wear monitoring. Cutting dynamics change also at higher frequencies i.e. ultra sonic vibrations and acoustic emission are also used for tool wear and failure monitoring. In the paper some physical reasons behind the use of various indirect monitoring methods of tool condition in drilling are presented and the benefits and drawbacks of each method are discussed.

Keywords: Drill wear monitoring, drill failure monitoring, thrust force, torque, drift force, spindle power, vibration, sound, acoustic emission

1. INTRODUCTION

Many research projects have been carried out in the field of tool wear and failure monitoring. There are a number of reasons for this interest among the research society and industry. Probably the most important reason is that manufacturing technology has changed towards process industry in the sense that today production equipment are capable to work in production cells which can be fully automated. However, in order to automate the production a way of tool condition monitoring is needed because a worn or broken tool could cause a lot of damage either to the workpiece or workpieces or, in the worst case, to the machine tool itself. In case of drilling this is rather apparent because if the tool is broken there might not be a hole where the next tool e.g. thread tap is used. A less radical consequence is that with a worn tool the surface finish and dimensions are not as good as they should be. If tools are changed based on the time they have been used the economical life time of these tools cannot be benefited from because there is great variation in tool life. An other factor related to this conservative way of defining the tool change is that valuable production time is lost because of unnecessary tool changes.

Tool condition monitoring can be based either on direct or indirect methods. The monitoring methods are considered direct if they actually measure the amount of wear and correspondingly indirect if the change of the measured parameter is a consequence of wear i.e. such as the increase of cutting force or vibration. This paper covers only indirect methods mainly because the direct methods still seem

ISSN 1363 - 7681 © 2004 COMADEM International

to be much more expensive methods to use and in many cases they also cause restrictions to the manufacturing process as such.

2. DRILL WEAR MODEL

Based on a series of drilling tests following relations based on physical models in drilling cast iron have been observed [27]:

$$\text{Torque (M)} = a_1 H_B d^2 f + a_2 H_B d^2 r + a_3 H_B d^2 w \quad (1)$$

$$\text{Thrust (T)} = a_4 H_B d f + a_5 H_B d w + a_6 H_B d r + a_7 H_B d^2 \quad (2)$$

Where

H_B = Brinell hardness of work material

d = diameter of the drill

f = feed per revolution

w = average flank wear

r = radius at the cutting edge

$a_1 \ldots a_7$ = constants

The terms in Equation 1 are coming from three contact zones [27] namely: a) The rake face of the tool which contacts the chip and transmits most of the force necessary to perform the cutting action. b) The nonzero radius of the tool cutting edge (the transition surface between the rake and flank faces) which contacts the work material at the point where the chip and work separate. This edge radius causes an indenting force. (The nonzero intercept observed for zero feed on cutting force versus feed rate plots may be attributed to this effect.) c) An area on the flank face having 0 deg clearance, known as the flank wear land which rubs against the work surface. The shear stress between the flank and the workpiece has been determined to be approximately equal to the work material yield shear stress. The shear force caused by the flank wear is termed to be the third force component.

From the above given relationships it is apparent that there is a strong dependency of workpiece hardness which actually means that tool life varies remarkably as a function of this [27]. Consequently, cutting of a few random workpieces of large hardness may influence the drill life much more than a large number of workpieces of low hardness. Hence, in an industrial operation, drills may fail very early or after a long time, depending on the occurrence of these few workpieces of high hardness. This could explain the large variation in drill life observed in industrial conditions. The workpiece hardness also influences the thrust forces and torque occurring in a drilling operation. If the variation in thrust force, on account of changes in flank wear, is to be significant, the variation in workpiece hardness has to be held within 5 percent of the mean hardness value. This is very difficult to achieve in industrial castings. Hence, torque or thrust measurements for monitoring drill wear should be attempted only after a very close tolerance has been obtained in the workpiece hardness. Another observation in ref. [27] is the speed in which drill wear takes place at the end of the drill life.

2.1 Process Parameters

The approach developed in ref. [27] has been further studied in ref. [17] putting more emphasis on the cutting parameters when drilling copper alloys. In the test series thrust force and torque have been recorded at three different flank wear states, three cutting speeds, three feed rates and three drill diameters. It has been concluded that the relationships between the cutting force signals and drill wear as well as other cutting parameters including spindle rotational speed, feed rate and drill diameter were established. Tool wear can then be estimated using these relationships. It is also shown that the tool wear can be estimated knowing the thrust force signal, feed per revolution and drill diameter. Based on the studies conducted, the following conclusions are drawn [17]. 1) The effects of feed per revolution, depth of cut and tool wear on cutting force signals are significant, while the effect of cutting speed on the cutting force signals is relatively insignificant in the cases studied. 2) Both the thrust and the torque increase as the flank wear increases. 3) Thrust and torque can be well represented as functions of tool wear, drill diameter and feed per revolution. 4) Tool wear can be properly estimated knowing the thrust force and other cutting parameters, especially for larger tool wear.

2.2 Drill Geometry

Due to production variations, a drill is typically slightly asymmetric [6]. Accordingly, the two corners of the drill point wear gradually while maximum wear alternates from one cutting edge to the other [3,6]. This alternating process continues until both lips have zero clearance at the margin. The drill then adheres to the workpiece and breaks if the cutting process is not stopped in time. The described phenomena is the explanation why drift forces can also be used as an indicator of tool wear [14] together with feed force and torque.

3. MONITORING METHODS

Quite a number of indirect monitoring methods have been tested for drill wear and failure detection. The most popular methods reported in literature have been feed force, torque, drift forces, spindle motor and feed drive current, vibration, sound, ultrasonic vibration and acoustic emission. A summary of how popular each of these methods have been is shown in Figure 1 based on all references cited in this paper. Cutting speed and feed rate have also often been measured although they are not really used for tool wear monitoring. Since the other measured parameters are influenced by the cutting speed and feed rate they are also needed in a monitoring system or in adaptive control systems e.g. [11].

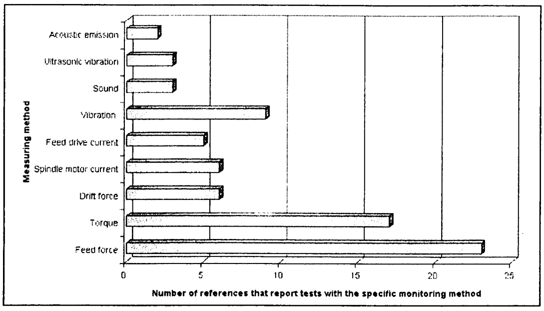


Figure 1: The popularity of measuring methods in drill condition monitoring.

In the subsequent paragraphs the monitoring methods are divided into two groups i.e. those that are related to measuring forces and those that are related to measuring vibration. The force measurements cover measurement of thrust (=feed) force, torque and drift forces together with measurement of spindle motor current and feed drive current. The spindle motor current actually corresponds to the measurement of torque although through a longer measuring chain and similarly feed drive current corresponds to the measurement of thrust force. The vibration related methods consist of mechanical vibration, sound, ultrasonic vibration and acoustic emission. Mechanical vibration is normally considered to take place from 1 Hz to about 10 kHz or 20 kHz. Airborne sound is often measured in the frequency range from 20 Hz to 20 kHz. Ultrasonic vibration starts from where mechanical vibration ends i.e. from about 10 kHz to about 80 kHz which then is the lower limit for acoustic emission which goes to as high frequencies as 1 MHz.

3.1 Drilling Forces

The equations presented in previous section drill wear model actually explain why drilling forces i.e. feed force and torque have so widely been tested for drill wear monitoring. From the equations it becomes clear that these forces increase with increasing tool wear. Drift forces also indicate tool wear because of the asymmetry of drills and also dynamics of the drilling process. Spindle motor current is an indicator of torque and feed drive current an indicator of feed force although through a longer measuring chain than the forces.

3.1.1 Feed Force

Feed force has been tested or is used as an indicator of tool wear and failure in references [2,3,5,8,10-15, 17-21,23-32] i.e. it is the most popular method for tool wear and failure monitoring in the cited literature. However, in many of the references the reported test material is very limited e.g. [2] and in some of the material the results are not encouraging i.e. the correlation between thrust force and wear has been found weak e.g. [3]. At the same time some other researchers have been successful in incorporating thrust force and torque into diagnostic approaches e.g. [15,17] and [18]. In references [19] and [20] the thrust force has been tested and used together with vibration with good success. The increase of dynamic variation of thrust force and torque as a sequence of drill wear has been verified to correlate with the surface quality of composite material in ref. [23]. An approach that can detect severe drill damage just before tool breakage occurs based on thrust force measurement has been proposed in ref. [28] and [29]. An other kind of approach for the same task i.e. detection of severe damage before the drill actually breaks, has also been developed in ref. [30]. In practise the measurement of feed force is rather a demanding task when it is done close to the tool i.e. it is difficult to find a way to measure force without causing some kind of problems. For example it is not very convenient to use an additional force transducer between the spindle and tool holder. Another place to position the force transducer is naturally between the workpiece and the table or between the table and the guide ways. In the tests reported in ref. [10] the feed force was ranked the second best measuring method after horizontal vibration when measured using a force transducer between the tool and spindle. The use of force sensor in the quideways was not as successful because in this kind of installation the force was actually influenced by the location of the hole in the workpiece, see Figure 2. However, it should be noted that at the end of drill life the indication of tool wear can be seen even with this rather poor signal.

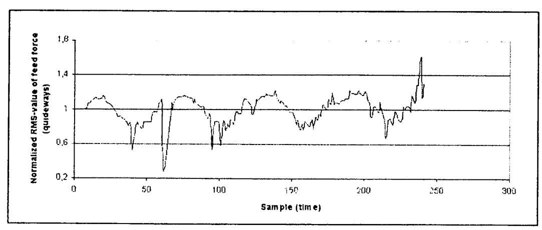


Figure 2:Normalised RMS-value of feed force measured from the guideways for a 10.2 mm twist drill.

All the results shown in this paper are based on measurements made with a horizontal machining centre. Four different kind of tools (shank end mill, end mill, twist drill and tread tap) were tested. The measurement system included force sensors, torque transducer, various equipment for the definition of voltage, current and electrical power consumption, accelerometers, acoustic emission sensors and microphones. The test and analysis program has been reported in more detail in ref. [10].

3.1.2 Torque

In most of the references where feed force has been tested also torque has been tested for the same purpose [3-5,8,10-12,14,15,17,18,21,23,25-27,31]. The conclusion of [3] does not give support to using torque as an indicator of drill wear. As could be expected on the basis of what has been presented about tool wear in the previous chapter the general experience with testing torque is more or less the same as with thrust force. One drawback of measuring cutting forces and torque has been the difficulty and, consequently, the expense of the measuring arrangement for practical applications. A new cheaper arrangement for measuring torque based on eddy current has been presented in ref. [4] with good experience.

3.1.3 Drift Forces

Drift force measurement is covered here in the same context as the other forces though it rather belongs to the same group as vibration measurements i.e. drift forces are a result of asymmetry in the drill and drilling process as explained earlier in this paper. Although the reported correlation with thrust force and torque with respect to tool wear was not good the results with drift force have been considered encouraging [3]. The result has been considered to give support to the theory of asymmetric wear of drills described in the drill geometry paragraph of this paper. However there seems to be a problem related to this type of measurement i.e. it has been indicated the parameters calculated from drift force seem to form a sort of an s-type trend index. At first the index increases when one side of the drill is wearing, and then it decreases and starts to increase again and so on. This makes it somewhat difficult to define when the drill is actually worn. Also the findings reported in ref. [14] support the theory of asymmetric wear of drills. It is suggested that tool life criterion could occur when the RMS of the drifting force achieves a minimum close to that of the sharp drill. In this paper all the figures that are shown and most of the findings in referenced papers are based on statistical parameters calculated from time domain signal. In most of these papers one or many of the tested methods have proved to be suitable for drill wear monitoring. In the studies reported in ref. [21] the conclusion for torque, feed and drift force measurements is simply that the signals in time domain do not show any correlation with drill wear. With more sophisticated signal analysis based on the use Fast Fourier Transform (FTT) the influence in the measured signals has been seen. The conclusion was that torque, feed force and drift force in x-direction showed good correlation with drill wear but the correlation of drift force in Y-direction was not as good. In ref. [31] after using torque, feed and drift forces in the test it has been suggested that the power spectrum of drift force could serve as an index to monitor the onset of tool failure.

3.1.4 Spindle Motor Current

Spindle motor current (or power depending what sort of measuring arrangement is used) is especially interesting as a measure of drill wear because it is so easy to monitor. In principle it can be expected that the same phenomena's as with torque should be possible to see in this signal. In ref. [16] spindle motor together with feed drive current has been tested for drill breakage detection. Also in ref. [24] good experience with diagnosis of drill wear based on spindle motor and feed drive current combined with diagnosis of failure with feed force is reported although no examples of actual signals are given in this reference. The experience reported in ref. [25] was not encouraging but the same also applies to feed force and torque measurements. In ref. [27] the reported spindle motor power and torque curves as a function of drill wear are very similar both indicating a very rapid increase in the signals at the end of drill life. In the tests reported in ref. [10] the analysis of spindle motor current or power was not successful as can be seen in Figure 3. It should be noted that there has been remarkable variation in the tests reported in ref. [10] but overall the results with electrical power and current measurements were below average level of the other signals. This could possibly be explained by the type of machine tool i.e. motors and also possibly into some extent to sensitivity of the measuring equipment used.

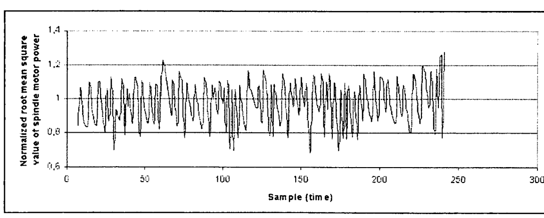


Figure 3:Normalised root mean square value of spindle motor power for a 10.2 mm twist drill.

3.1.5 Feed Drive Current

The measurement of feed drive current has been tested in many of the same references as spindle motor current or power [7,10,16,24,25]. It is somewhat surprising that not that many researchers have tried to use feed drive current since it should be able to give similar information as the feed drive force. Actually one could expect that the amount of noise in feed drive current would be less pronounced than is the case with spindle motor current. Quite similarly as with the spindle motor current drill breakage has also been detected with feed drive current in ref. [16]. An example of the root mean square value of feed drive servo motor current is given in Figure 4 based on the tests reported in ref. [10]. The indication of tool wear is not as evident as with some other measuring techniques although the level increases at the end of the tool life. Apparently some disturbance has taken place in the beginning of the test so that one measurement value clearly is not as it should be.

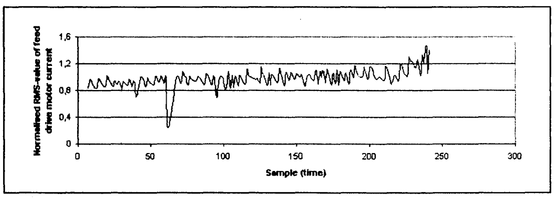


Figure 4: Normalised RMS-value of feed drive servo motor current for a 10.2 mm twist drill.

3.2 Vibration

It is very logical to measure vibration as an indicator of tool wear and failure. In principle, when the cutting forces increase due to wear also vibration and sound emitted by the structure in question increase. The increase of vibration at higher frequencies can also be expected when forces increase. Another logical reason for vibration monitoring to work is that when the tool becomes worn the cutting process tends to get somewhat more unstable i.e. the dynamic nature of the process becomes more apparent and vibration increases. This is also the reason why parameters that indicate the variation of the measuring signal (higher order terms like standard deviation and kurtosis) actually show some difference. It should be noted that vibration related measuring signals tend to be easier to use in practise than most of the force related methods since an accelerometer or microphone can be installed a bit further away from the tool and workpiece but in order to work really well force and torque should be measured between the motors that provide the forces and the tool which is somewhat more complicated or demanding.

3.2.1 Mechanical Vibration

Mechanical vibration has been studied by quite a number of researchers [1,2,6,7,9,10,13,19,20,26]. This is not surprising because of the previously described reasons and also the fact that vibration is the most widely used method in condition monitoring of machinery in general. The way some of the different types of drill wear i.e. chisel, outer corner, flank and margin is seen in the spectrum of vibration signal has been studied in ref. [6] with artificially produced wear. It has been concluded that monitoring vibration has been proved to be a useful method in predicting drill wear and failure. It should be noted that in the spectrums the most dominating increase of vibration due to wear has taken place at high frequencies (3 - 6 kHz) close to the natural modes of the tool and tool holder. The proposed analysis methods i.e. Kurtosis value together with cepstrum analysis, power spectrum and a statistical triggering parameter (ratio of absolute mean value) would seem to try to focus on the change of the dynamics of the signal. In tests reported in ref. [10] vibration (acceleration) was the best indicator of all of the tested methods including force, motor current, acoustic emission etc.. An example of vibration signals is given in Figure 5. Thrust force and vibration have been tested in references [19] and [20] simultaneously. It has been concluded that either of these could be used for on-line classification of drill wear. However, integrating both signals yields better results.

3.2.2 Sound

The use of sound measurements has only been tested in a few references [2,3] and [10]. In [3] it has been considered striking that the curves of two completely different physical values i.e. drift force and sound have been very similar. However, it could be debated whether the result could actually be anticipated since airborne sound actually is a result of the mechanical vibration of the parts of the machine tool, tool and work piece and the vibration of these is a function of the dynamic forces present in the drilling process. Actually it could be claimed that all the same information that is available in mechanical vibration signal should be available in the sound signal. However the problem with sound signal is that it is very diffused i.e. it is reflected from various surfaces in various directions. The real benefits in using sound measurements for tool wear and failure detection are in the ease of installation of the transducer since a microphone can very easily be installed rather close to the tool and the price of a microphone compared to an accelerometer is really low. An other factor which one could think of that would have encouraged the use of sound measurements in drill wear monitoring is the fact that the machine tool operators often rely on their hearing when they define whether the tool is worn i.e. the sound the tool produces changes with tool wear. The results reported in ref. [10] can be considered promising. In Figure 6 a sound signal curve is shown for the same drill for which also vibration curve has been shown. Although the standard deviation value of sound signal is not quite as clear an indicator as the low pass filtered RMS-value of vibration signal it tells the same story.

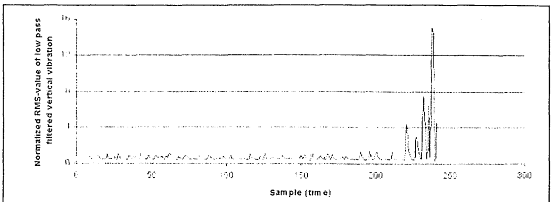


Figure 5: Normalised RMS-value of low pass filtered vertical vibration for a 10.2 mm twist drill.

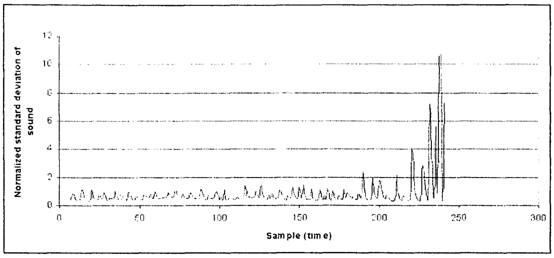


Figure 6: Normalized standard deviation of sound signal for a 10.2 mm twist drill.

3.2.3 Ultrasonic Vibration

Some researcher have tested ultrasonic vibration [9,13,26] for drill wear monitoring and breakage detection. However, in some cases vibration measurement below 80 kHz has also been called acoustic emission because it does not seem to be commonly accepted how vibration at these higher frequencies should be called. In ref. [9] vibration in frequency range from 20 kHz to 80 kHz is defined as ultrasonic vibration and the same definition is used in this paper. The use of ultrasonic vibration as an indicator of tool wear is explained and compared to other methods in the following way in ref. [9]. Acoustic emission is considered to suffer from severe attenuation and multi-path distortion caused by bolted joints commonly found in machine tool structures. It has especially been noted that ultrasonic vibration does not suffer as much because it takes place at lower frequencies and consequently the transducer can be placed fairly far from the chip forming zone. When compared to lower frequency vibration, ultrasonic vibration is considered better in the sense that the structural modes of vibration do not affect it because the structural modes in this range are so closely spaced that they form a pseudo-continuum. In ref. [26] ultrasonic vibration has been tested together with torque, feed and drift force measurements. In tests ultrasonic vibration has been the most effective method both for wear and failure monitoring especially when the signal analysis of ultrasonic vibration is based on band-pass filtering (10 kHz bands with 10 kHz steps in frequency range from 10 to 70 kHz.

3.2.4 Acoustic Emission

Acoustic emission takes place when a small surface displacement of material surface is produced [22]. It is considered that acoustic emission can be used to monitor crack growth, sudden impacts and rubbing of material against another which all cause vibration of the structure at very high frequencies (from 80 kHz to 1 MHz). Monitoring of acoustic emission has been rather popular in turning but surprisingly it does not seem to be that popular in drill wear and failure monitoring. One possible explanation to this is that in turning the AE transducer can be positioned closer to the tool than is the case in drilling. One of the benefits of acoustic emission is that since it takes place at very high frequencies it does not travel very far i.e. noise from other sources such as electrical motors do not travel to the tool in turning due to damping. The same feature is actually very easily a drawback in drilling since in practise the transducer has to be positioned rather far away from the tool and there might be a number of joints on the way where the AE needs to travel from one part to another which is very disadvantageous to the signal. The above explanation is probably the reason why in the tests reported in ref. [10] acoustic emission was not found to be one of the best methods for tool wear monitoring in drilling. The normalized root mean square value of acoustic emission (200 kHz centre frequency) is shown in Figure 7. In the example some disturbances are seen in the early part of the signal. Unfortunately it is rather typical that something, which destroys the measuring signal, happens during the machining process. There are a number of possibilities that could be the cause of this type of jump in the analysed signal e.g. something has hit the transducer or cable or some outside source has caused high vibration noise. It is possible to try to avoid wrong conclusion if a number of signals from various transducers are used in the diagnosis as the basis for defining whether the tool is worn or not. In ref. [22] acoustic emission has been tested with drill with and without TiN (Titanium Nitride) coatings. It has been suggested that AE works best when certain phases of drilling process are analysed because there are peaks in the signal in the beginning and end of the process of drilling a hole.

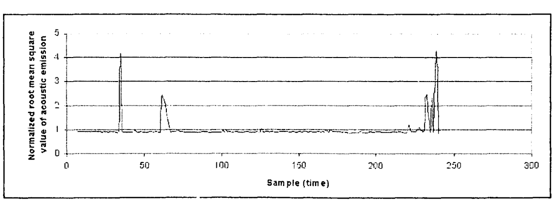


Figure 7: Normalized root mean square value of acoustic emission for a 10.2 twist drill.

4. CONCLUSION

Condition monitoring of cutting tools is important for a number of reasons. The unmanned use of flexible manufacturing systems is not possible without a reliable system for tool condition monitoring. Tool wear affects the surface quality of processed workpieces. Tools cannot be optimally used based on tool change policy which relies on time and which easily leads to too frequent change of tools from which it follows that valuable production time is lost and the tool cost becomes high. Based on the reports of a number of researchers it can be claimed that there is great variation in how well different monitoring methods work in tool condition monitoring. It is well known and accepted and also mathematical models have been developed which show that cutting forces increase as a function of tool wear and consequently thrust force and torque are often monitored in drilling. Feed drive and spindle current actually also measure the same thing as feed force and torque transducers although through a longer measuring chain and therefore also they can be used for tool wear and failure monitoring. Tool wear also changes the dynamics of cutting processes and consequently drift forces, vibration and sound have been used for tool wear monitoring. Cutting dynamics change also at higher frequencies i.e. ultra sonic vibrations and acoustic emission are also used for tool wear and failure monitoring. Based on the reported test material in the literature and the tests reported in this paper it would seem that thrust force, torque, drift forces, mechanical vibration, sound and ultrasonic vibration are all potential monitoring methods for drill wear monitoring although not all of the experience gained is as good. The result is not surprising since all of these methods are linked and those methods that have not been as successful simply suffer from a longer measuring chain that dampens the signals and introduces noise.

5. REFERENCES

1. Barker R.W., Kluthe G.A. and Hinich M.J., Monitoring Rotating Tool Wear Using Higher-Order Spectral Features, Journal of Engineering for Industry, Vol. 115, Transactions of the ASME, pp23-29, (1993).

2. Braun S. and Lenz E., Machine Tool Wear Monitoring, Mechanical Signature Analysis, Theory and Applications, Academic Press Ltd, pp321-342, (1986).

3. Braun S., Lenz E. and Wu C.L., Signature Analysis Applied to Drilling, Journal Mechanical Design, Vol. 104, Transactions of the ASME, pp268-276, (1982).

4. Brinksmeier E., Prediction of Tool Fracture in Drilling, Annals of the CIRP, Vol. 39, No. 1, pp97-100, (1990).

5. Christoffel K. and Jung W., Überwachungseinheit für die Bohrbearbeitung, Industrie Anzeigner, Vol. 103 (62), pp198-199, (1981).

6. El-Wardany T.I., Gao D. and Elbestawi M.A., Tool Conditon Monitoring in Drilling Using Vibration Signature Analysis, International Journal of Machine Tools & Manufacture, Vol. 36, 6, Pergamon Press Inc., pp687-711, (1996).

7. Erdélyi F. and Sántha C., Monitoring Tasks on Boring and Milling Production Cells, Computers in Industry, Vol. 7, Elsevier Science B.V., pp65-71, (1986).

8. Govekar E. and Grabec I., Self-Organizing Neural Network Application to Drill Wear Classification, Journal of Engineering for Industry, Vol. 116, No. 3, Transactions of the ASME, pp233-238, (1994).

9. Hayashi S.R., Thomas C.E. and Wildes D.G., Tool Break Detection by Monitoring Ultrasonic Vibrations, Annals of the CIRP, Vol. 37, No. 1, pp61-64, (1988).

10. Jantunen E. and Jokinen H., Automated On-line Diagnosis of Cutting Tool Condition (Second version), International Journal of Flexible Automation and Integrated Manufacturing, 4 (3&4), Begell House Inc., pp273-287, (1996).

11. Kavaratzis Y. and Maiden J.D., System for Real Time Process Monitoring and Adaptive Control during CNC Deep Hole Drilling, Proceedings of Comadem, '89, Kogan Page, London, UK, pp148-152, (1989).

12. König W. and Christoffel K., Sensoren für die Bohrbearbeitung, Industrie Anzeigner, Vol. 103 (100), pp29-33, (1980).

13. Kutzner K. and Schehl U., Werkzeugüberwachung von Bohrern kleinen Durchmessers mit Körperschallsensoren, Industrie Anzeigner, Vol. 110 (82), pp32-33, (1988).

14. Lenz E., Mayer J.E. and Lee D.G., Investigation in Drilling, Annals of the CIRP, Vol. 27, No. 1, pp49-53, (1978).

15. Li P.G. and Wu S.M., Monitoring Drilling Wear States by a Fuzzy Pattern Recognition Technique, Journal of Engineering for Industry, Vol. 110, No. 2, Transactions of the ASME, pp297-300, (1988).

16. Li X., On-Line Detection of the Breakage of Small Diameter Drills Using Current Signature Wavelet Transform, International Journal of Machine Tools & Manufacture, Vol. 39, Issue 1, pp157-164, (1999).

17. Lin S.C. and Ting C.J., Tool Wear Monitoring in Drill-

ing Using Force Signals, Wear, 180 1-2, Elsevier Science S.A., pp53-60, (1995).

18. Liu T.I. and Anantharaman K.S., Intelligent Classification and Measurement of Drill Wear, Journal of Engineering for Industry, Vol. 116, Transactions of the ASME, pp392-397, (1994).

19. Liu T.I. and Ko E.J., On-Line Recognition of Drill Wear via Artificial Neural Networks, Monitoring and Control for Manufacturing Processes, PED. Vol. 44, ASME, Dallas, TX, pp101-110, (1990).

20. Liu T.I. and Wu S.M., On-line Detection of Drill Wear, Journal of Engineering for Industry, Vol. 112, Transactions of the ASME, pp299-302, (1990).

21. Noori-Khajavi A., Frequency and Time Domain Analyses of Sensor Signals in a Drilling Process and Their Correlation with Drill Wear, Ph.D. Thesis, Oklahoma State University, Stillwater, OK, (1992).

22. Quadro A.L. and Branco J.R.T., Analysis of the Acoustic Emission during Drilling Test, Surface & Coating Technology, Vol. 94-95, No.1-3, Elsevier Science S.A., pp691-695, (1997).

23. Radhakrishnan T. and Wu S.M., On-Line Hole Quality Evaluation for Drilling Composite Material Using Dynamic Data, Journal of Engineering for Industry, Vol. 103, Transactions of the ASME, pp119-125, (1981).

24. Ramamurthi K. and Hough C.L. Jr., Intelligent Real-Time Predictive Diagnostics for Cutting Tools and Supervisory Control of Machining Operations, Journal of Engineering for Industry, Vol. 115, Transactions of the ASME, pp268-277, (1993).

25. Routio M. and Säynätjoki M., Tool Wear and Failure in the Drilling of Stainless Steel, Journal of Materials Processing Technology, Volume 52, Issue 1, Elsevier Science B.V., pp35-43, (1995).

26. Schehl U., Werkzeugüberwachung mit Acoustic-Emission beim Drehen, Fräsen und Bohren, Dissertation, RWTH Aachen, (1991).

27. Subramanian K. and Cook N.H., Sensing of Drill Wear and Prediction of Drill Life (I), Journal of Engineering for Industry, Vol. 99, Transactions of the ASME, pp295-301, (1977).

28. Tansel I.N., Mekdeci C., Rodriguez O. and Uragun B., Monitoring Drill Conditions with Wavelet Based Encoding and Neural Network, International Journal of Machine Tools & Manufacture, Vol. 33, No. 4, Pergamon Press Ltd, pp559-575, (1993).

29. Tansel I.N., Rodriguez O. and Mekdeci C., Detection of Tool Breakage in Microdrilling Operation with RCE Neural Networks, PED, Vol. 47, No. 1, ASME, pp83-88, (1992).

30. Thangaraj A. and Wright P.K., Computer-asisted Prediction of Drill-failure Using In-process Measurements of Thrust Force, Journal of Engineering for Industry, Vol. 110, Transactions of the ASME, pp192-200, (1988).

31. Valikhani M. and Chandrashekhar S., An Experimental Investigation into the Comparison of the Performance Characteristics of TiN an ZrN Coatings on Split Point Drill Using the Static and Stochastic Models of the Force System as a Signature, The International Journal of Advanced Manufacturing Technology, 2 (1), pp75-106, (1987).

32. Von Nedeß C. und Himburg T., Automatisierte Überwachung des Bohrens, VDI-Z, Bd. 128, Nr. 17, pp651-657, (1986).

The article is reprinted from Condition Monitoring and Diagnostic Engineering Management, (Comadem 2001), Proceedings of the 14th International Congress 4-6 September 2001, Ed. A.G. Starr and Raj B.K.N Rao, Manchester, UK, ISBN 0 08 0440363, Copyright 2001, Pages No. 781-791, with permission from Elsevier Science.

PUBLICATION IV

Dynamic effects influencing drill wear monitoring

In: Proceedings of the MFPT 58th Meeting.
Ed. by H.C. Pusey, S.C. Pusey & W.R. Hobbs.
Virginia Beach, USA, April 25–30, 2004. Pp. 51–60.

DYNAMIC EFFECTS INFLUENCING DRILL WEAR MONITORING

Erkki Jantunen
VTT Technical Research Centre of Finland
P.O.Box 1702, FIN-02044 VTT, Finland
Email: Erkki.Jantunen@vtt.fi

Abstract: Tool wear monitoring is economically very important but technically a rather demanding task. In this paper an attempt is made to get further understanding of the dynamics that influence the drilling process and especially what happens when a drill is worn. A very simplified approach is tested in the development of the cutting forces and modeling the influence of wear in these forces. The developed horizontal forces are used for excitation of a simplified one degree of freedom model of the drill. The dynamic model is used for producing vibration velocity signal as a function of drill wear. Using this signal the most typical and widely used signal analysis techniques are tested. The signal analysis data produced with the developed simplified model shows similar trends as data measured in laboratory tests and can be considered useful in the development of an automatic diagnosis program for drill wear monitoring.

Key Words: Tool wear; Drilling; Monitoring methods; Signal analysis; Dynamic model; Cutting force

Introduction: Tool wear monitoring is important due to a number of reasons such us: Unmanned production is only possible when machine tools are equipped with a reliable tool wear monitoring system. Tool wear influences the quality of surface finish of the products produced and thus, if unnoticed, can cause high costs. The economical tool life can not be benefited from without tool wear monitoring. Unfortunately tool wear monitoring is a very difficult task. There are methods available that monitor the tool directly i.e. measure the tool wear but these methods are not practical enough to be used outside laboratories. Indirect monitoring methods such as measurement of cutting forces or vibration are technically demanding to be used. Reference [1] gives a summary of the indirect monitoring methods that have been used for drill wear monitoring. Feed force and torque measurements have been widely used in laboratory tests but it could be claimed that they are not methods practical enough for everyday use, especially if they are measured between the tool and the spindle. If measured from the table the measuring points are located further away from the point where the forces are initiated, and consequently these measuring points do not give as reliable results. Vibration monitoring is one of the most widely used methods which according to the literature survey [1] and reported tests [2] has proved to work well in practice. In this paper a simplified simulation model is developed in order to gain further understanding how horizontal forces and vibration could be used for drill wear monitoring. It is also hoped that the simplified model could serve as a testing and training tool when automated diagnostic tools such as fuzzy logic, neural networks or rule based expert systems are developed.

Cutting force model: In theory drilling does not induce horizontal forces i.e. forces that are perpendicular to the drill axis, if the drill has two cutting lips because these two lips cancel the influence of each other. In practice horizontal forces exist and they can be measured, and also due to these forces horizontal vibration occurs. There are a number of reasons for these forces: The drill geometry is not perfect i.e. the cutting lips do not have exactly similar geometry and consequently forces are induced, the work piece material is never exactly homogeneous causing some horizontal force components, the drilling process does not take place exactly perpendicular to the surface of the work piece. When the drill is worn the two cutting lips do not wear exactly to the same extent causing some unbalance of forces which can vary from side to side depending on which cutting lip has worn more [3]. When the drill starts to vibrate because of the reasons described above, and also due to forces that are induced to the drill from the spindle, the vibration causes horizontal movement resulting in unbalance in the horizontal forces and further vibration. The proposed model tries to take into account all the above named factors. However, the model does not try to predict the exact drill forces nor the unbalance in horizontal direction, but instead it merely tries to show the influence of various factors so that the force predicted and the vibration velocity calculated with the model would have the characteristics of those forces and vibrations measured in laboratory tests.

The first component in the drilling force model is a factor that takes into account the discontinuous nature of the drilling process i.e. always when a new hole is drilled the forces start from zero see e.g. reference [4]. This process can be described mathematically in a simplified form with the following feed force function:

$$F_{dh}(t) = (t - i \cdot t_d)/(t_d/b_l) \qquad \text{if} \qquad i \cdot t_d \leq t < i \cdot t_d + t_d/b_l \tag{1}$$

$$F_{dh}(t) = 1 \qquad \text{if} \qquad i \cdot t_d + t_d/b_l \leq t \leq i \cdot t_d + t_d \tag{2}$$

Where t is time, i is a counter for the hole number, t_d is the time it takes to drill one hole and b_l is a coefficient which defines the relation of the increasing part and the stable part of the thrust force. Figure 1 shows the simulated feed force when t_d is 4 seconds and altogether 15 holes are drilled. It should be noted that all the forces i.e. torque and horizontal forces can be expected to perform similarly.

The second step in the development of a simulation model is to introduce the actual drilling force models. Based on a series of drilling tests the following relations in drilling cast iron have been observed [5]:

$$\text{torque } (M) = a_1 \cdot H_B \cdot d^2 \cdot f + a_2 \cdot H_B \cdot d^2 \cdot r + a_3 \cdot H_B \cdot d^2 \cdot w \tag{3}$$

$$\text{thrust } (T) = a_4 \cdot H_B \cdot d \cdot f + a_5 \cdot H_B \cdot d \cdot w + a_6 \cdot H_B \cdot d \cdot r + a_7 \cdot H_B \cdot d^2 \tag{4}$$

where H_B is Brinell hardness of work material, d is diameter of the drill, f is feed per revolution, w is average flank wear, r is radius at the cutting edge and a_1 ... a_7 are constants. The use of the above formulas enables the scaling of force defined in formulas 1 and 2. In the above relationships there is a strong dependency on work piece hardness which actually means that tool life varies remarkably as a function of this [5]. Consequently, cutting of a few random work pieces of high hardness may influence the drill life much more than a large number of work pieces of low hardness. Hence, in an industrial operation, drills may fail very early or after a long time, depending on the occurrence of these few work pieces of high hardness. This could explain the large variation in drill life observed in industrial conditions. Since the purpose in the development of a simulation model really is to be able to see the influence of wear in the measured signals, it can be concluded from the above formulas that it is logical to develop an approach where part of the forces is a function of wear, and part is not, and that both parts strongly depend on the drill diameter and hardness of the material. Another possible way to calculate the drilling forces would be the kind of approach developed by Chandrasekharan [4] and used by Yang et. al. [6] where drilling forces are calculated based on the geometry of the drill and results from turning tests which define the necessary parameters for the approach. In this kind of approach the cutting lips are divided into a number of sections where the forces are calculated and then integrated. However, since the purpose of this study is not to define the exact forces that could be measured, it is easier and much faster to use a statistical approach for scaling the forces so that the effect of work piece material and cutting conditions can be taken into account.

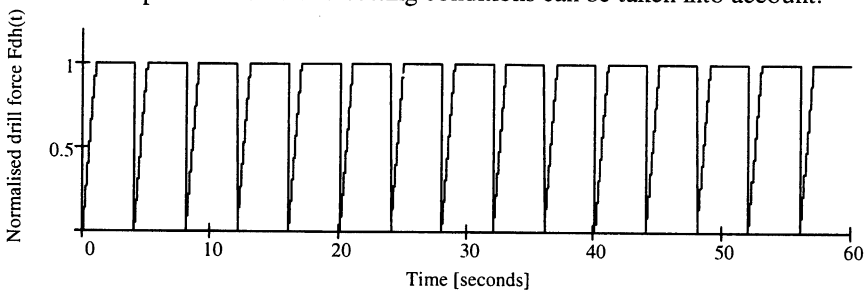


Figure 1. Simulated feed force in drilling.

In the developed approach it is assumed that there are a number of reasons for horizontal forces to appear in drilling. In theory these forces do not exist because typically there are two cutting lips in drills. However, according to the tests and various references [1] horizontal forces can rather well be used for drill wear monitoring. Possible reasons that can cause unbalance in these forces are e.g. geometrical differences between the two cutting lips and differences in the wear of the two cutting lips. In fact, it can be assumed that if there are differences in the beginning due to geometrical errors, there will be differences in wear of the two cutting lips since the forces, which are the cause of wear, are different. Following formulas have been developed in order to take into account the difference of the forces of the two cutting lips:

$$F_{rpm1}(t) := F_{dp}(t)\cdot\left(c_1 - c_2\cdot\ln\left(1-\frac{t}{t_c}\right)\right)\cdot\cos\left(2\cdot\pi\cdot\omega\cdot t + \phi_{ge} + \phi_{wd}\cdot\sin\left(\omega\cdot\frac{t}{c_3}\right)\right) \quad (5)$$

$$F_{rpm2}(t) := F_{dp}(t)\cdot\left(c_1 - c_2\cdot\ln\left(1-\frac{t}{t_c}\right)\right)\cdot\cos\left[2\cdot\pi\cdot\omega\cdot t + \pi\cdot\left(1 + \phi_{wd}\cdot\sin\left(\omega\cdot\frac{t}{c_4}\right)\right)\right] \quad (6)$$

where $c_1 \ldots c_4$ are constants, t_c is the total lifetime of the drill, ω is the angular speed of rotation, ϕ_{ge} is the angular geometrical error due to the tolerance in manufacturing the drills, ϕ_{wd} is the difference in wear of the two cutting lips of the drill and F_{dp} drilling process force that scales the size of the forces. The random variation of wear from one cutting lip to another is taken care of by varying the phase between the cutting forces i.e. the effect of the difference between the two constants c_3 and c_4. The use of a logarithmic wear function is based on the use of very simplified wear model which tries to describe progressive wear [7]. In this kind of case wear rate increases as the forces increase and since the forces are initiated by the wear this is an accelerating process. The drilling process force F_{dp} can e.g. be calculated based on equations 3 and 4 so that it would take into account the change of drilling parameters i.e. feed and also the hardness of the drilled material. Since the simulation model that is developed here is used for the purposes of development of tool wear monitoring and it is not supposed to predict the horizontal forces physically correctly, the effect of the tool diameter and radius at the cutting edge can be neglected as these are not variables for a specific drill that is monitored. Based on the above the following equation is used for drilling process force:

$$F_{dp} = c_5 \cdot H_B \cdot f \cdot F_{dh} \tag{7}$$

where c_5 is a constant, H_B is the Brinell hardness of the work piece material, f is feed per revolution and F_{dh} is calculated according to equations 1 and 2.

In the vibration velocity signal of most rotating machines, vibration amplitudes at the harmonics of rotating speed can be seen. There are a number of reasons for this i.e. if the vibration is distorted in the sense that it is not sinusoidal, Fast Fourier Transform (FFT) produces these harmonics and also there are quite a number of possible excitations at these frequencies such as bearing frequencies and those excited by the driving engine which most often is an electrical motor. In the developed approach a number of excitation forces at the harmonic frequencies of the rotating speed are assumed to exist. These are defined by the following summary function which defines harmonic components starting from the 3rd and reaching to the 11th harmonic force component:

$$F_{nrpm}(t) := \sum_{n=3}^{11} \left[F_{dp}(t) \cdot \left(\frac{c_6}{n} - \frac{c_7}{n} \cdot \ln\left(1 - \frac{t}{t_c}\right)\right) \cdot \cos(n \cdot 2 \cdot \pi \cdot \omega \cdot t) \right] \tag{8}$$

where c6 and c7 are constants, n defines the order of the harmonic component, $F_{dp}(t)$, ω and t_c as defined above.

Another typical factor that is always present in vibration measurements is the noise of the signal i.e. random fluctuation of the measured signal. Also in the case of noise there are a number of reasons for it, some of which originate from the measured machinery due to random excitation which could be caused by many sorts of reasons such as movement of the machinery or some other machine. In the cutting process the cutting fluid is one source, and also chip flow causes random vibration. The electrical measuring equipment is also a source of random fluctuation in the measured signal. In order to make the simulation model to provide more natural signals, a random component is also introduced to the calculation of the excitation force. The random force is defined by following formula:

$$F_{rnd}(t) = rnd(c_8) - c_8/2 \tag{9}$$

where c_8 is a constant and rnd denotes the MathCad program function [9] that produces an equally distributed random number between 0 and c_8.

When the drill starts to vibrate during the drilling process one consequence from this is that the cutting lips do not cut a round hole and as a result of that the horizontal forces are not in equilibrium [6]. Because the drill together with the tool holder basically vibrates like a beam that is only supported from one end it can be expected that vibration at the first natural frequency of that structure is rather pronounced. Following from this it is logical to introduce a horizontal force into the dynamic model that gives excitation to the model at the natural frequency:

$$F_o(t) := \cos\left(2\cdot\pi\cdot f_o\cdot t\right)\cdot F_{dp}(t)\cdot\left(c_9 - c_{10}\cdot\ln\left(1 - \frac{t}{t_c}\right)\right) \tag{10}$$

where c9 and c10 are constants, t_c total tool life time, drilling force F_{dp} as defined above and f_o is the first natural frequency of the drill and tool holder calculated with the following formula [8]:

$$f_o := \frac{1}{2\cdot\pi}\cdot\sqrt{\frac{k}{m}} \tag{11}$$

where m is the mass of the drill and tool holder, and k is the stiffness of the structure. It should be noted that although the above formula is very simple it is not easy to define the natural frequency exactly without measuring it. In the following analyses the mass m = 1.4 kg and the stiffness k = 395 N/mm have been chosen to be the same as used in reference [6] for a d = 15.9 mm drill.

It could also be assumed that due the inhomogeneous nature of the work piece horizontal forces would be seen. These are not modeled separately, instead it is assumed that a static force acting for some tenths of seconds into one direction would mainly induce vibration at the natural frequency of the drill and tool holder. This kind of source is taken care of by equation 10 and also partly by the random excitation defined by equation 9. In fact in reference [6] inhomogeneous work piece material is considered the main source of initial excitation and is induced to the model as a randomly acting force. The assumptions as described above apply also to the effect that is caused by the fact that drilling does not in practice start exactly perpendicular against the work piece surface i.e. it is assumed that equations 9 and 10 take care of this effect, too. Naturally it could be even argued that this geometrical error might in many cases be very small and consequently also the forces would be very small.

The final step in the development of the excitation force in the simulation model is to add together all the five components which have been introduced above. This can be done simply by calculating the sum of all the five components:

$$F_x(t) = F_{rpm1}(t) + F_{rpm2}(t) + F_{nrpm}(t) + F_{rnd}(t) + F_o(t) \tag{12}$$

Since the developed model is not physically exact i.e. it is assumed that the force components that have been presented above do exist in reality but it would be very difficult to calculate the exact size of each force, and instead of exact solution the model tries to bring out the features that can be seen in drill wear monitoring. Therefore the features of the calculated force sum function fully depend on the chosen parameters. It should be noted that in theory it would also be possible to try to approach the problem in a more precise way i.e. trying to look for the actual physical solution. In such a case one possible approach would be similar to the one that has been developed in reference [6]. The approach chosen by Yang et. al. follows the principles developed by Chandrasekharan [4]. In the approach the cutting lips of a drill have been studied in a number of sections, typically 50 sections and for each one of these the different force components have been calculated based on tests in oblique cutting. In principle it would be possible to introduce wear into such a model by looking at each individual section and by saving the history of the cutting process in each section so that when the forces get higher the probability of wear would get higher, and using a random function the material loss would be described. Naturally this kind of a solution would not really have an equivalent case in reality but statically this kind of a model could be adjusted to correspond to measured values in laboratory. The other force components that have been introduced above i.e. harmonic components and a random component could be with some accuracy defined based on laboratory tests. The influence of vibration could then be defined using a similar approach as has been used by Yang et. al. [6] where the actual cutting path influences the drilling forces. The approach suggested in reference [6] could be further developed if instead of two degrees of freedom a higher number would be used e.g. using finite element method (FEM). However, the purpose of this study is to show that even with a relatively simple approach with proper choice of parameters the typical features of vibration velocity signal can be seen. The sum force function calculated using equation 12 with the following values of constants c1 = 20, c2 = 400, c3 = 2, c4 = 1.7, c6 = 0.04, c7 = 0.08, c8 = 0.5, c9 = 0.02 and c10 = 0.04 and also assuming F_{dp}= 0.102 (equation 7) is shown in Figure 2 for the first hole and in Figure 3 for the last hole. In the example the total life time of the drill is defined to be 15 holes i.e. total tool life is 60 seconds when it takes 4 seconds to drill one hole. The signals shown in Figures 2 and 3 are supposed to show the development of the horizontal drilling forces in the sense that in the beginning the time signal of the force seems to be rather noisy and no one frequency component stands out of the others. Towards the end of the life of the drill the forces get bigger and the influence of the defined frequencies such as speed of rotation can be seen.

Dynamic model: The development of the dynamic model follows the principles used by Yang et al. in reference [6]. It is assumed that the drill and the tool holder can be modeled like a beam that is rigidly supported at one end and the excitation force influences at the other end. In the above mentioned reference [6] two degrees of freedom have been studied basically because of the development of a dynamic model for the drilling force based on the influence of vibration to the shape of the hole which becomes distorted if compared to the theoretically round shape. In this study only one degree of freedom is studied since the excitation force is supposed to take into account the above described phenomenon. The simplified dynamic model can then be described with the following differential equation [8]:

$$mx'' + cx' + kx = F_X(t) \tag{13}$$

where m is the mass of the vibrating tool and tool holder, c is damping, k the stiffness and Fx(t) the dynamic horizontal drilling force defined in the previous chapter. The forced vibration differential equation can be solved using Runge-Kutta method [8]. In the analysis MathCad program package has been used for calculation of the vibration response [9]. Figure 4 shows the excitation force and vibration displacement for the modeled tool life time of 60 seconds. In order to make Figure 4 clearer vibration displacement curve has been moved from above the force by adding five to the response and subtracting five from the force values so that the curves do not coincide each other.

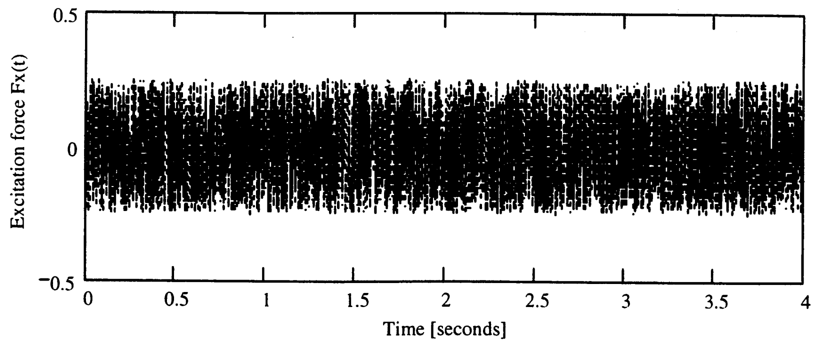


Figure 2. Excitation force in the beginning of the simulation.

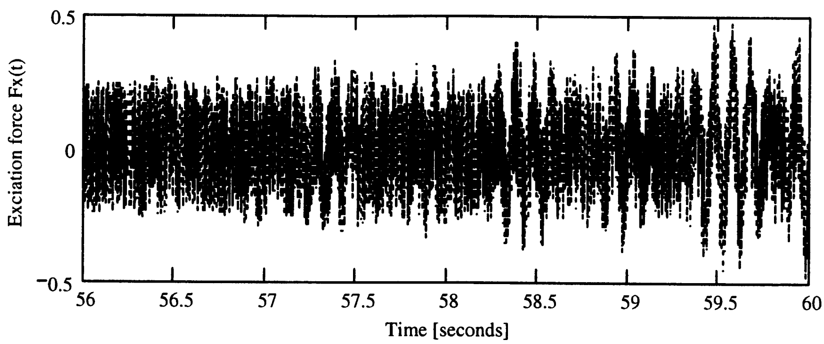


Figure 3. Excitation force in the end of the simulation.

Signal analysis: Reference [1] gives a summary of the signal analysis methods that have been used for drill wear monitoring. Most of the references that have been reviewed use statistical parameters such as root mean square (rms) value in analyzing the time domain signal. Figure 5 shows the development of such statistical parameters as rms and maximum value of simulated horizontal vibration velocity as a function of time for the total tool life time. (It should be noted that for this kind of a simulated signal, the rms value and the standard deviation value is actually the same.) The parameters have been calculated using a time constant of 0.05 seconds. The constants and parameters values have been the same as in the previous chapters for the development of the excitation force. In reference [2] it is reported that such statistical parameters as rms, mean deviation and maximum where the best statistical parameters in the analysis of the best measuring signal i.e. horizontal vibration. The reported findings correspond very well with the trends seen in Figure 5 with simulated data.

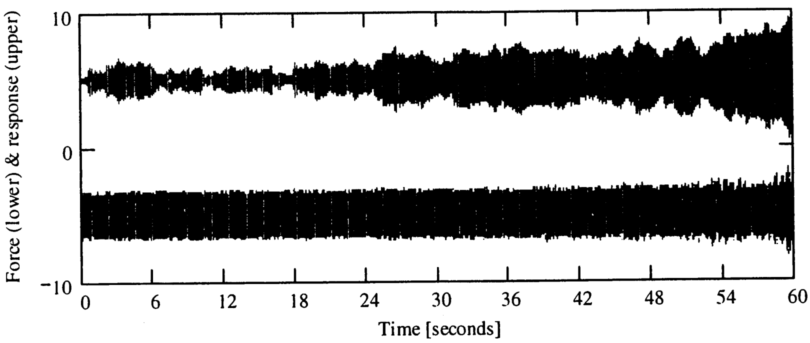


Figure 4. Excitation force from equation 12 (lower curve) and response (upper curve).

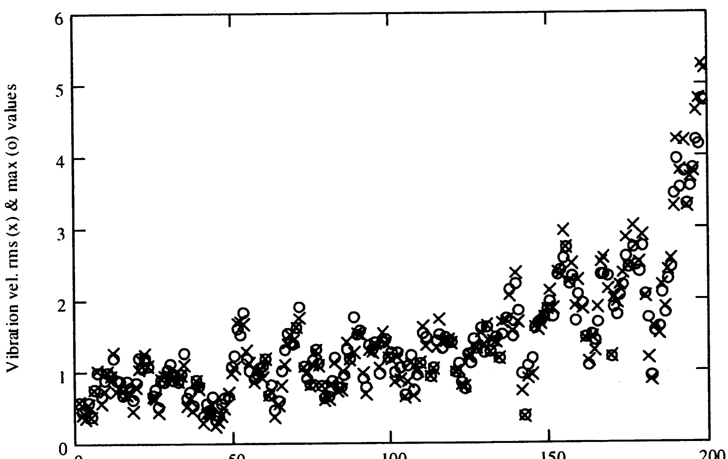


Figure 5. Simulated statistical vibration parameters as a function of drill wear (time).

Figure 6 shows the corresponding statistical parameters from laboratory tests reported in reference [2]. Although there are remarkable differences between laboratory tests and simulation the trend is very similar. The biggest difference is that in simulation there is not much difference whether normalized rms or maximum value is used but in laboratory tests there is more variation in the maximum value i.e. the process is not as stable as has been defined in simulation. However, it should be remembered that there is much more variation when the laboratory test results of individual drills are compared with each other. The life time of drills varies a lot and also the increase of the normalized statistical parameters during the lifetime of the tool varies remarkably. Based on the above it can be suggested that the simulation model can be used e.g. in the development and testing of expert systems for drill wear monitoring.

An other signal analysis method that has been widely used is the Fast Fourier Transform (FFT). Figure 7 shows the waterfall presentation of a simulated vibration velocity spectrum in the frequency range from 0 – 150 Hz. The FFT analysis has been done using the same constant and parameter values as in the case of statistical signal analysis. In spectrum analysis hanning window has been used and the number of points has been 2000 and in the analysis the shown logarithmic spectrums represent the average of three spectrums calculated with 50% overlap. In Figure 7 the drill wear can be seen rather clearly. This result again corresponds to the reported result in reference [2] i.e. more sophisticated analysis functions show the development of tool wear more clearly than just statistical parameters. However, it should be noted that with such analysis functions like FFT it is important to know at which frequencies the amplitudes should be followed.

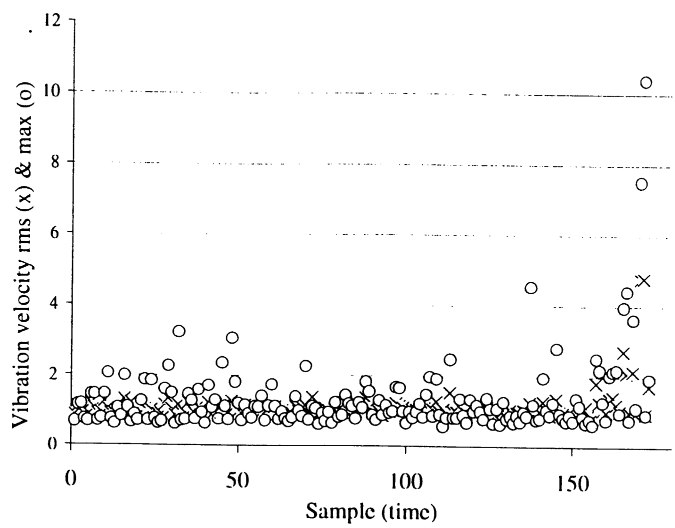


Figure 6. Statistical vibration parameters from laboratory tests.

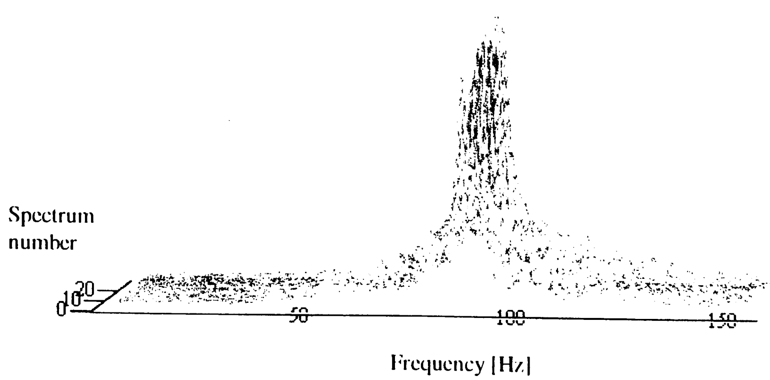


Figure 7. Waterfall presentation of frequency spectrum showing the influence of wear.

In this simulation the indication of drill wear can be seen at the excitation frequencies i.e. multiples of rotational speed and the first natural vibration of the drill and tool holder. In order for a tool wear monitoring system to work it should have the capability of calculating these frequencies and following the amplitude trend at these specific frequencies.

Conclusion: Tool wear monitoring is economically very important but technically a rather demanding task. In this paper an attempt has been made in order to reach further understanding of the dynamics that influence the drilling process and especially what happens when a drill is worn. A very simplified approach has been tested in the development of the cutting forces and modeling the influence of wear in these forces. Such factors as geometrical difference of the cutting lips, different kind of wear history of the lips, vibration at first natural frequency and excitation at harmonics of the speed of rotation have been taken into account in the development of the excitation force. The developed forces have been used for excitation of a simplified one degree of freedom model of the drill. The dynamic model has been used for producing vibration velocity signal as a function of drill wear and with this signal the most typical and widely used signal analysis techniques i.e. statistical time domain parameters and spectrum analysis have been tested. The signal analysis data produced with the developed simplified model shows similar trends as data measured in laboratory tests and can be considered useful in the development of an automatic diagnosis program for drill wear monitoring.

References:

[1.] Jantunen, E., A Summary of Methods Applied to Tool Condition Monitoring in Drilling, International Journal of Machine Tool & Manufacture 42, 2002, pp. 997-1010

[2.] Jantunen, E. & Jokinen H., Automated On-Line Diagnosis of Cutting Tool Condition, International Journal of Flexible Automation and Integrated Manufacturing, 4(3&4), 1996, pp. 273-287

[3.] El-Wardany, T.I., Gao, D. & Elbestawi M.A., Tool Condition Monitoring in Drilling Using Vibration Signature Analysis, International Journal of Machine Tool & Manufacture 36 (6), 1996, pp. 687-711

[4.] Chandrasekharan, V., A Model to Predict the Three-Dimensional Cutting Force System for Drilling with Arbitrary Point Geometry, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1996

[5.] Subramanian, K. and Cook, N.H., Sensing of Drill Wear and Prediction of Drill Life, Transactions of the ASME, Journal Of Engineering for Industry, 99, 1977, pp. 295-301

[6.] Yang, J.A., Jaganathan V. and Du, R., A New Dynamic Model for Drilling and Reaming Process, International Journal of Machine Tools & Manufacture 42, 2002, pp. 299-311

[7.] Jantunen, E., Prognosis of Wear Progress Based on Regression Analysis of Condition Monitoring Parameters, Proceedings of the Comadem 2003 conference. August 27-29, Växjö, Sweden.

[8.] Thomson, W.T., Theory of Vibration with Applications, 4th edition, Prentice Hall, New Jersey, ISBN:0-13-915323-3

[9.] Mathsoft, Mathcad 11 User's Guide, Mathsoft Engineering & Education Inc., 2002, ISBN:1-57682-297-4

PUBLICATION V

Flexible expert system for automated on-line diagnosis of tool condition

In: Proceedings of a Joint Conference, Technology Showcase, Integrated Monitoring, Diagnostics & Failure Prevention, MFPT 50th Meeting. Joint Oil Analysis Program Technical Support Center, University of Wales. Ed. by H.C. Pusey & S.C. Pusey. Mobile, Alabama, USA, April 22–26, 1996. Pp. 259–268.

FLEXIBLE EXPERT SYSTEM FOR AUTOMATED ON-LINE DIAGNOSIS OF TOOL CONDITION

Erkki Jantunen
Technical Research Centre of Finland, Manufacturing Technology,
P.O.Box 1702, FIN-02044 VTT, Finland

Harri Jokinen
Technical Research Centre of Finland, Manufacturing Technology,
P.O.Box 1702, FIN-02044 VTT, Finland

Robert Milne
Intelligent Applications Limited,
1 Michaelson Square, Livingston, W. Lothian, Scotland EH54 7DP

Abstract: An increasing number of Flexible Manufacturing Systems (FMS) have been installed in Europe during the past few years. The general experience is that the availability of the FMS is not as high as was originally expected and especially their unmanned use during three shifts has not been successful. One of the major problems is the deterioration or failure of the tools. To develop a system to address this, a wide cutting test and analysis program for tool wear was performed. The test program covered both shank end and end mills together with twist drills and tread taps. For monitoring the tool wear a number of monitoring methods such as vibration, acoustic emission, sound, spindle power and current, axial force, torque were tested. The relations between the analysed signals and tool wear form a basis for the diagnosis rules that are used in an diagnostic expert system module. An expert system for automated on-line diagnosis of tool wear of different types of tools was built using a new approach. In this approach the faults are described in a fault tree database and the corresponding features of condition monitoring signals together with the machine status information are described in a symptom tree database. Using a rule synthesiser program the information gathered in the databases is automatically converted to expert system code.

Key Words: Artificial intelligence, Condition monitoring, Expert systems, Flexible manufacturing systems, Tool wear monitoring

INTRODUCTION. An increasing number of Flexible Manufacturing Systems (FMS) have been installed in Europe during the past few years. A general experience is that the availability of the installed FMS is not as high as was originally expected, and especially the unmanned use has not been successful [1]. A major problem is the condition of the tool. One of the most important reasons for this is today's existing real-time tool condition monitoring techniques do not cover the wide range of different machining situations and machining parameters that normally take place in practice.

There is a need to group and synchronize sensor signals together to avoid poor correlation between a single signal source and the measured event. However, it is obvious that the commercially available monitoring systems are exploiting just a limited number of the capabilities of modern sensor and analyzing techniques [2]. In this survey several sensors were installed and comprehensive laboratory tests done before the concluding sensor validation [3]. The aim was to accomplish the requirements of condition monitoring at critical points of the machine tool and in the cutting processes. The validation was performed according to the following criteria: sensitivity of the sensor to the measured event, correlation between the signal and measured event, the amount of deviation and universality. The information received from multiple sensors was analyzed with many different methods. The relations between the analyzed signals and wear form a basis for the diagnosis rules that can be used in an AI system.

The diagnosis of the condition monitoring signals has to be done using an expert system since the goal is to be able to use FM-systems unmanned in three shifts. One big problem in reaching this goal is how much the FM-systems differ from each other. From this it follows that an expert system should be easy to configure for each task. Unfortunately, the behavior of measuring signals is a function of the machine tool and tool type. In this study a lot of emphasis has been put on creating a generic solution, and a new approach to configuring an expert system. The basic idea is to use databases for defining the varying information, and to use specific software to write the expert system code automatically. The user only defines the necessary data with the aid of modern tools for database handling and the computer translates that information into a working expert system code.

TEST ARRANGEMENT. A small horizontal-type machining center with an 11 kW main motor power was used in the cutting tests for tool and machine tool condition monitoring. The tests concentrated on tool wear, tool breakage and collision monitoring. Cutting tests were performed in order to create situations where a measurable event was present due to tool wear or failure. In this part of the cutting tests, different types of cutting tools were used to cover a wide range of different cutting methods. The tools investigated in the tests were shank end mill (diameter 6 and 10 mm, HSS), end mill (diameter 50 mm with carbide inserts), twist drill (diameter 3.3, 5.0, 6.8, 8.5 and 10.2 mm, HSS) and thread tap (M4, M6, M8 and M12, HSS). The tool-monitoring tests were carried out by using different kinds of measuring arrangements. The main configuration of the measuring arrangement can be seen in Figure 1. A more detailed description of the measuring arrangement is in reference [3].

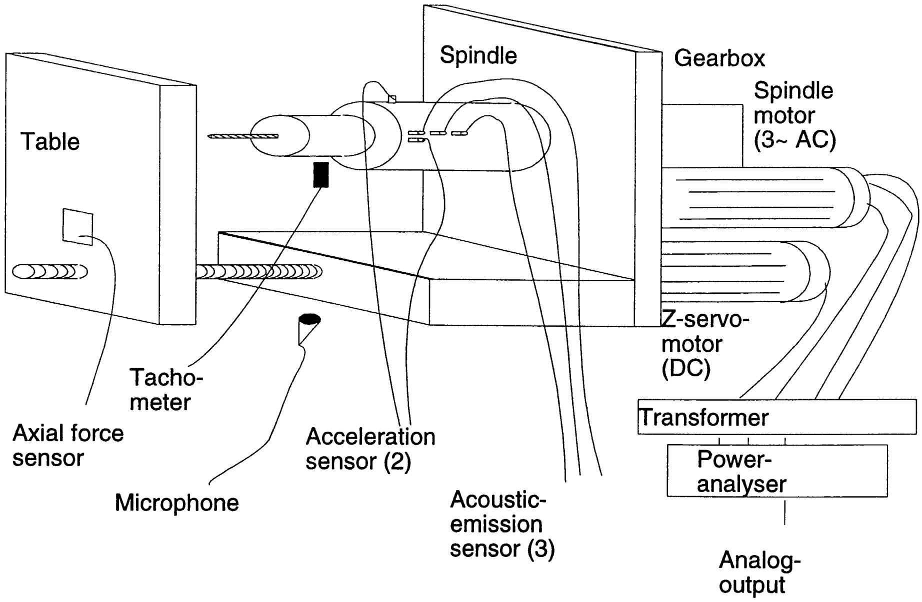


Figure 1. Measuring arrangement

EXPERT SYSTEM. With the increasing complexity of NC machine tools and flexible manufacturing systems, and with the growth in diversity of systems, it is increasingly difficult for maintenance personnel to assess the problems rapidly. Long delays in identifying the precise problem greatly increases downtime and causes even bigger secondary problems in plants. This is especially true for FM systems where the material flow from a machine tool to another is performed quickly with a minimum number of buffers.

To automatically identify the condition of the machine tools and cutting process two types of diagnostics are needed. A reactive diagnosis is needed to identify the cause of a current problem and a predictive diagnosis is needed where ever it is possible to anticipate the need for a maintenance action based on the condition monitoring of the machine tool or cutting process.

The development of a diagnostic expert system is based on diagnostic rules which are derived from the results of the condition monitoring tests. From the results of the analysis it became apparent that rules would become rather complex if the system were completely generic, i.e. suitable for a number of different types of FMS environments, and also that there would be a need to divide the tasks within the system so that responses would be fast enough to perform the diagnosis in the case of collision and tool breakage also.

Principles of the chosen approach. In order to make an expert system flexible and suitable for a wide range of FM systems, a new approach for defining and modifying the rules was developed.

The basic idea is to use the fault tree database definition program for defining the faults, and describe corresponding condition monitoring tools (symptoms) using the symptom tree database definition program. After that, the user starts a rule synthesiser program that translates the contents of the fault and the symptom databases into expert system rule code for the computer performing the monitoring task. In principle the translation program simply takes one page from the symptom tree at a time and writes a module to the expert system code from that. The procedure is shown in Figure 2.

It is considered that there are several advantages to this approach: It is not necessary to write enormous amounts of expert system code manually. It is very easy to make changes or add more information and, especially, it is possible to configure the program for a specific FM-system. Apparently, there are also certain disadvantages to this approach: The amount of code in the final expert system will be considerable since it is not possible to use the sophisticated features of an expert programming package. Instead of a sophisticated system the expert system programming module always writes rather simple modules for each condition defined in the symptom tree.

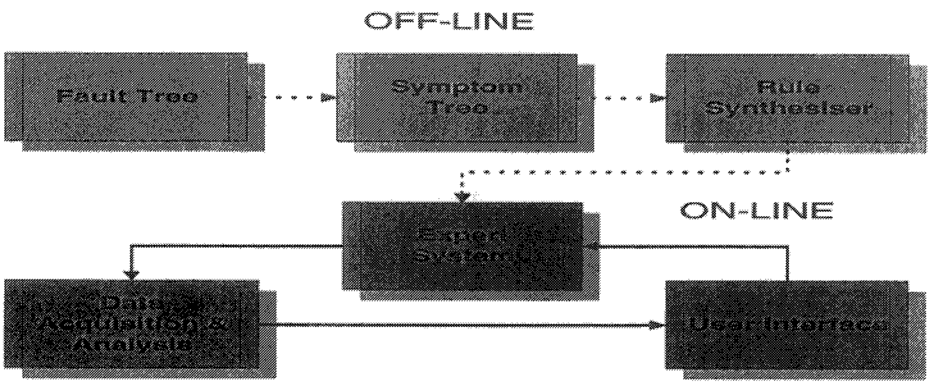


Figure 2. Principles of the new approach to expert system rule generation.

Fault tree. In the fault tree database the machine tool is defined with the chains of subcomponents. The number of subcomponent levels is limited to five. For the lowest level of subcomponents where a principal fault can take place, all the possible faults are described. The fault tree database program has all the typical search and editing functions of a normal database program. The structure/window of the fault tree is shown in an example in Figure **3**.

Symptom tree. Following the definition of all the relevant faults with the fault tree database, the next step in building an expert system with this new approach is to define all the symptoms related to the faults together with information about the machining process. As shown in Figure 4 the following input is defined in the symptom database: fault tree component chain identification, fault to be processed, tool identification, the status information of the machine tool, machining information and condition monitoring method information (symptom). The symptom is defined with signal, general, analysis and limit value information (Figure 4).

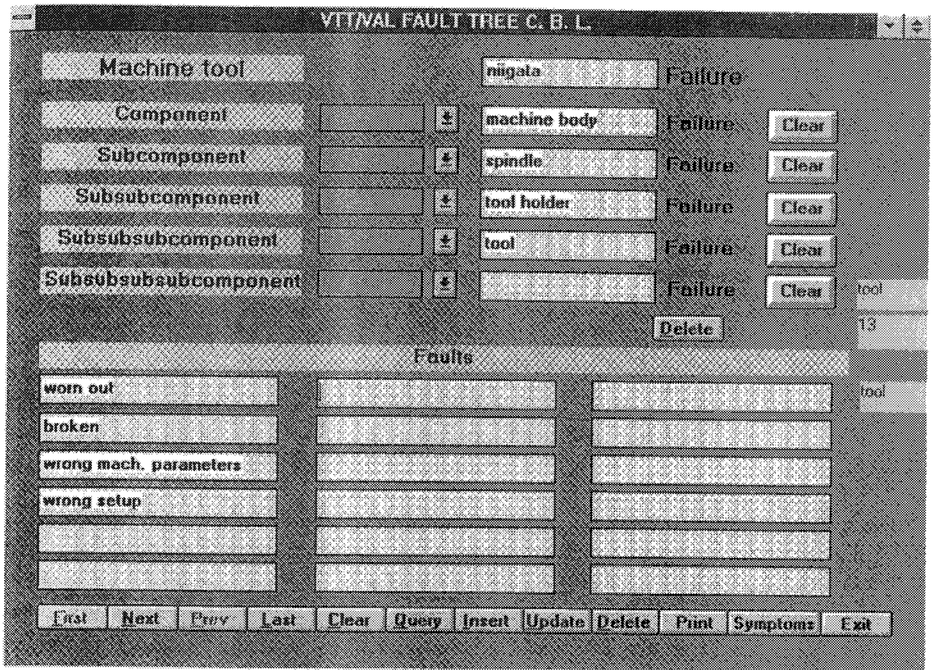


Figure 3. Fault tree database program window

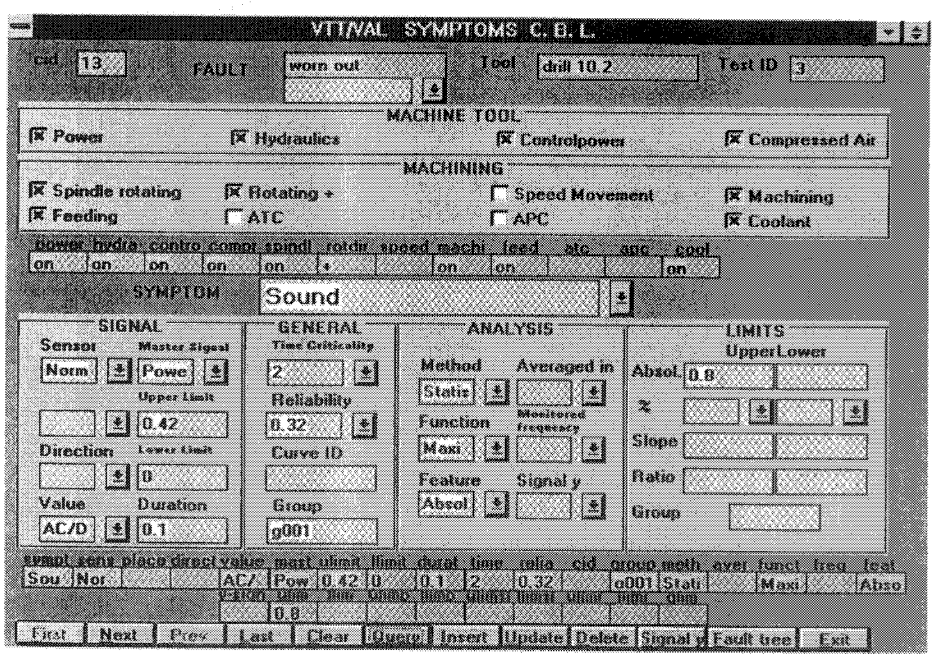


Figure 4. Symptom tree database program window

SIGNAL PROCESSING. For the automatic analysis of the huge amount of data to be gathered an interface was created using the Visual Basic programming environment for Windows. The system gathers the data with a data acquisition board (16 A/D channels, Keithley) and the necessary calculation procedures are defined using a collection of subroutines (VTX) for this AD card family.

Data acquisition. The statistical analysis is based on the acquired data. Attached to the acquisition board, a sample and hold board is used to synchronize dynamic signals. The measured signals are analyzed with a number of different methods in the time domain and in the frequency domain. In the case of dynamic signal analysis, only so-called cursor values are gathered to minimize the amount of information to be stored in the databases. These actions are also controlled by the interface. The flowchart of data acquisition is shown in Figure 5.

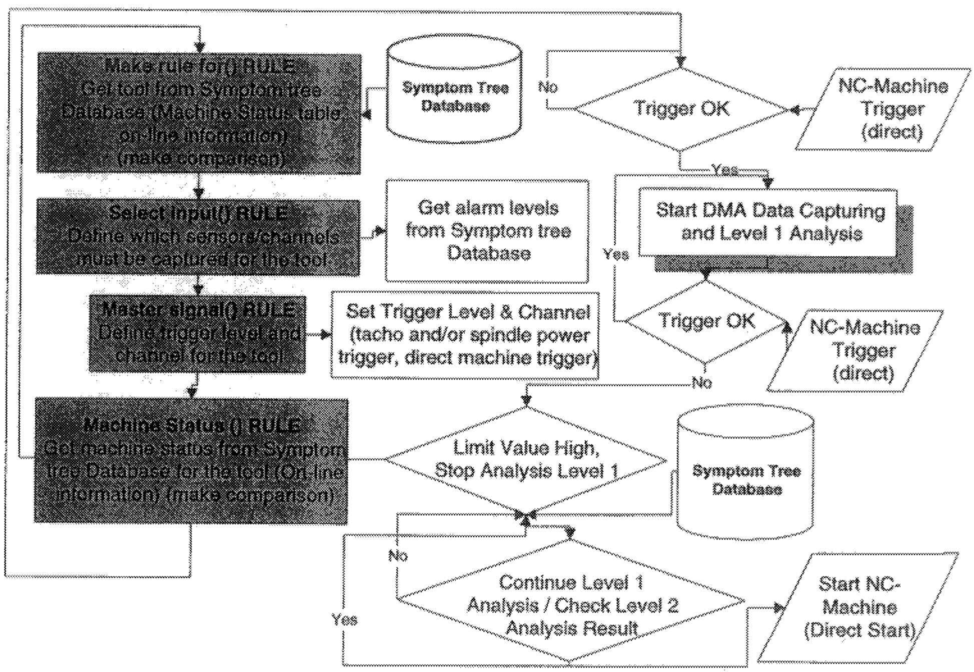


Figure 5. Level 1 data acquisition

Signal analysis. Depending on the measured events, the data to be analyzed is first cleaned of irrelevant signals, e.g., rapid movement during drilling, that has not been recorded during the actual machining process. Data measured and recorded simultaneously from the sensors is studied by calculating a number of statistical parameters: arithmetic mean, root mean square (RMS), mean deviation, standard deviation, skewness, kurtosis, maximum and minimum.

In the case of dynamic signals containing frequency information, Fast Fourier Transformation (FFT) techniques are employed [4]. The sample and hold function of the data acquisition board is used to get data from four channels simultaneously. It is possible to perform a FFT with both time and frequency domain averaging. Different kinds of analysis functions such as spectrum, cross-spectrum, frequency response, coherence, coherent output power, autocorrelation, crosscorrelation, cepstrum, liftered spectrum, 1/3 octave spectrum, 1/1 octave are available.

Regression analysis. The results of the statistical analysis and FFT analyses are further analyzed using regression analysis techniques. Different regression functions were tested to find the highest correlation between measured tool wear and analyzed measurement signals. The sets of data points are approximated as closely as possible with the four smoothing functions [3]: first, second and third order polynomials and one logarithmic function based on a simplified mathematical definition of wear [5] using the least square principle.

The degree of the fit is much higher for cursor values of the FFT functions than for the statistical parameters. This result is logical, since the idea of the FFT analysis is to separate meaningful information from noise. However, it takes time to carry out the FFT analysis, which makes it impossible to use the FFT for collision and tool breakage monitoring but enables tool wear monitoring. The goodness of fit varies with the tool-type. Drilling and shank end milling are the easiest to monitor. Functions describing how much two signals are related to each other, show a rather high goodness of fit (3). The use of at least two signals for tool wear monitoring coincides with the findings of the statistical analysis, since the best methods for monitoring purposes vary between the tool-types.

Simulation module. This work is focused on problems of automatically identifying the condition of the cutting process. The simulation program module is used as a tool to see how the expert system reacts in different kinds of situations with different kinds of limit values for the chosen regression models. The program fits a selected regression curve to the existing data. The fitted curves are the same as those used in the regression analyses. After each time the curve fitting has been done, the program checks to see whether the tool worn-out limit has been reached, giving a warning when the machining process has to be terminated by the expert system. The advantage of using the simulation module together with the regression functions is that with this procedure the amount of measuring data to be stored in the databases is greatly reduced and the method is not too sensitive for sudden changes in the measuring signals i.e. it is possible to avoid most false alarms.

Rule Synthesiser. Knowing what signals to process, how to process them and what features or thresholds to look for after processing, requires considerable knowledge about how tools deteriorate and the signal processing capabilities that are available. This 'expert' knowledge was gained during the initial experimentation of the project. A key facet of the knowledge is the processing to perform changes from tool type to tool type. Like traditional expert systems, a mechanism is needed to allow the expert to naturally specify his knowledge of how to process the signals for a given tool. However, the system cannot use one fixed knowledge base, since the best way to analyze the signals will vary from tool type to tool type. Hence a more flexible approach to building the expert system is needed, since the 'expert system element' will also change from tool type to tool type.

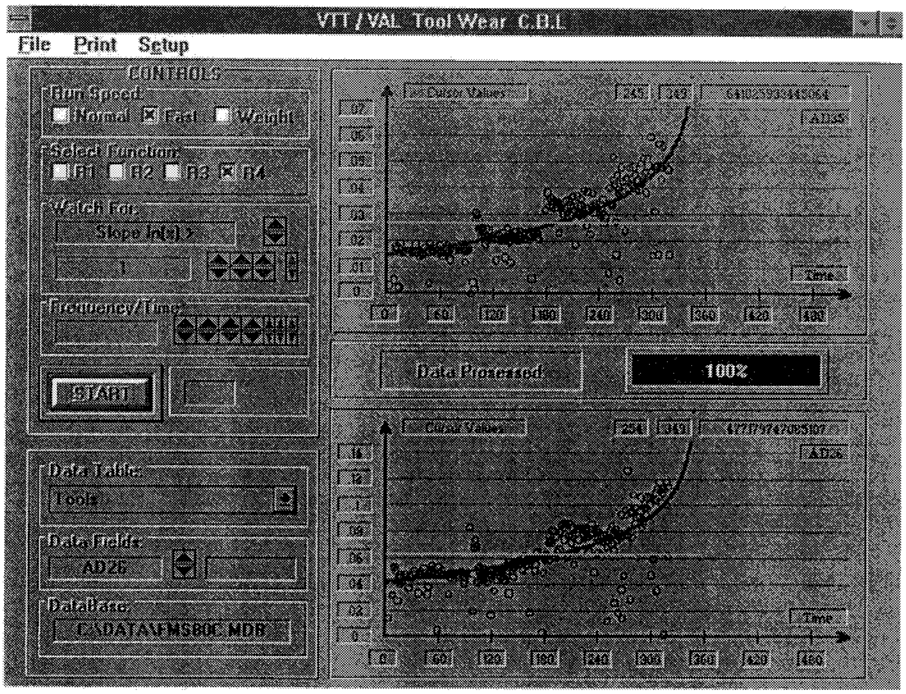


Figure 6. User-interface window of the simulation program, end milling, upper curve: horizontal vibration, mean deviation, lower curve: sound, root mean square

The information in the symptom tree database can be viewed as a specification for the knowledge based rules. Each entry specifies what a rule should look like; what signals to examine, what processing to perform, which features to extract and what thresholds to use. Using a traditional expert system environment would be too complex for experts in this domain to specify all this information. Also, hidden from the expert is the fact that each 'rule' has several parts, corresponding to the above steps. The database front end provides an user interface that is natural for the expert to use and hides this underlying complexity. Our system uses a 'rule synthesiser'. This takes the specification of a rule that is contained in the symptom tree database and automatically generates the computer programs needed to implement these 'rules'. Figure 7 shows the steps and database interactions that result from each rule. Contrasting this with the simple layout of Figure 4, illustrates the gain that results from the use of the rule synthesiser.

The rule synthesiser works by processing each rule specification in the symptom tree database, breaking each rule into several function calls. It also builds the links between these function calls so that the data can progress through the steps of acquisition, signal processing, feature extraction and testing against the specified limits. In addition, it automatically combines rules into groups, for example, all the rules to detect worn out 10.2 mm drills. This process hence implements one of the key ideas of expert systems - let the expert specify the knowledge in a natural way and have the system do all the hard work.

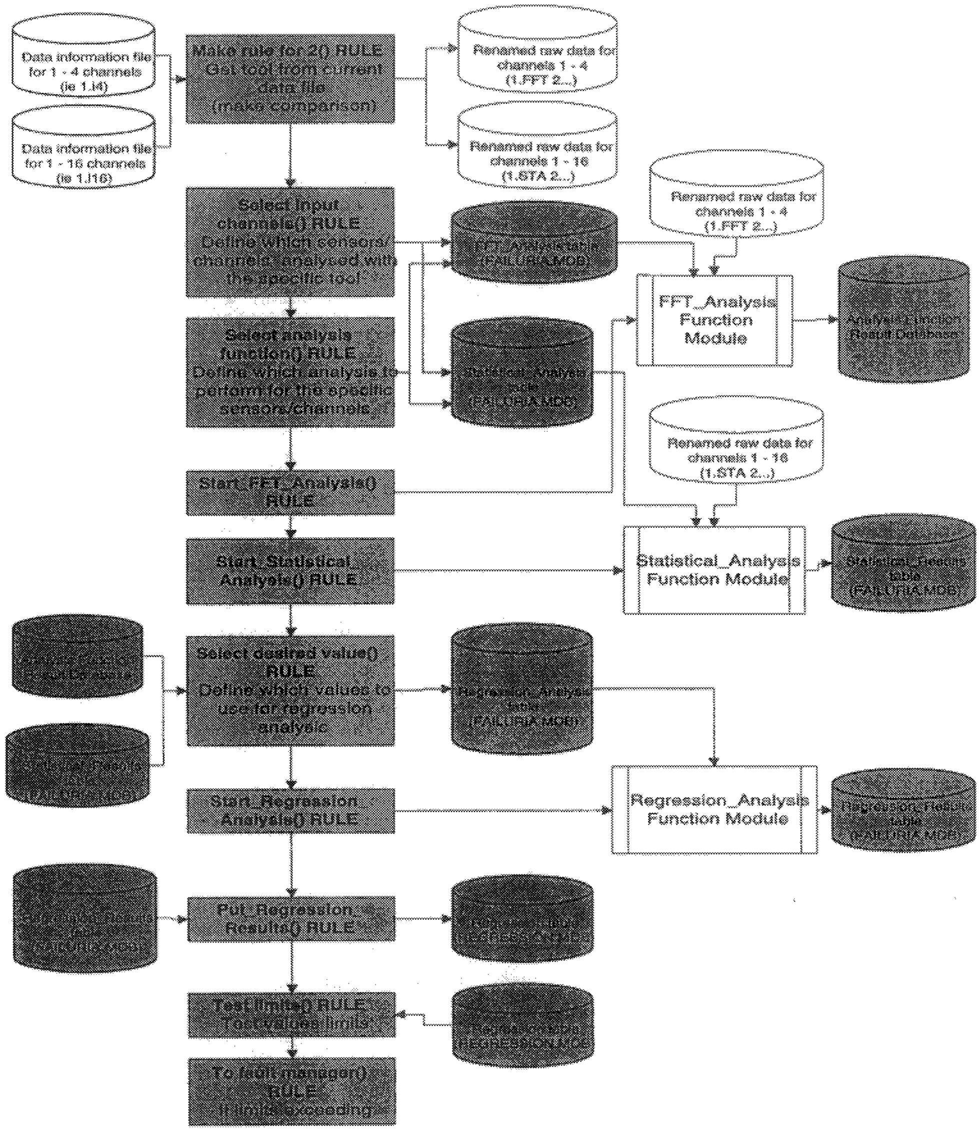


Figure 7. Level 2 rules and data model

CONCLUSION. It has been shown that there exists a great potential to improve the utilization rate of a machine tool by an advanced condition monitoring system using modern sensor and signal-processing techniques. A comprehensive cutting test procedure has been carried out with different tools and measurement arrangements. The recorded signal information has been processed in several ways, both in the time and the frequency domain. The effectiveness of the best sensors and analysis methods have been verified in the prediction of the remaining lifetime of a tool in use. The relations between the analyzed signals and tool wear form a basis for the diagnosis rules that are used in a diagnostic expert system. A new approach to development of a rule-based expert system is reported. The approach makes it easy to configure the expert system for different types of FM-systems with different types of tools. The solution is based on the adoption of fault and symptom tree databases with sophisticated user interfaces for the definition of the relevant fault types together with the corresponding monitoring methods. A rule synthesiser is used to take the specification of a rule and produce the detailed expansion to control the various sub systems of the condition monitoring system. The hides from the expert the underlying complexities of the task and lets him specify knowledge in a way that is natural to him. Without this capability, the system would be too complex to be used by the people who have the appropriate knowledge, and hence the value of the whole system would be greatly reduced. This work has combined extensive laboratory testing with the implementation of a state of the art signal processing environment and an easy to use way to specify the knowledge about how to interpret the data that is collected. Such a system is not only practical, but is an essential part of the how automation can be used to increase the utilization of flexible manufacturing systems.

ACKNOWLEDGMENTS. The results of the tool wear tests are used in the development of an integrated condition and machining process monitoring system for flexible manufacturing systems and for stand-alone NC-machine tools (FMS Maint System, 1993-1996, EUREKA MAINE EU 744 project). The work of the project team, D. Deasy, J. Hinkkanen, M. Karhu, H. Komulainen, P. Kontkanen, A. Poikonen, D. Smith, R. Vannela and K. Vähä-Pietilä is gratefully acknowledged.

REFERENCES.

1. Lakso, T. The Influence of FMS-techonology on the Efficiency of NC-controlled Machine Tools. Tampere University of Technology. Publications 50. Tampere 1988. 106 p.
2. Kluft, W. Techniken zur Werkzeugüberwachung auf der Basis von Kraft- und Verformugsmessung. Sonderdruck aus der Vortagsreihe des VDI-Seminars "Überwachung und Diagnose von Maschinen und Prozessen " von 3 -4-Juni 1985 in Aachen, Germany. 14 p.
3. Jantunen, E. & Jokinen, H. Automated on-line diagnosis of cutting tool condition. 5th International Conference of Flexible Automation & Intelligent Manufacturing, FAIM 95, June 28-30, 1995, Stuttgart, Germany
4. Randall, R.B. Application of B&K equipment to frequency analysis. In Brüel & Kjær. Nærum, Denmark. 239 p.
5. Jantunen, E. & Poikonen, A. Dynamics in monitoring gear faults. Proceedings of the 47th meeting of the mechanical failures prevention group. Virginia Beach, Virginia, USA, April 13-15 1993. p. 279-288.

PUBLICATION VI

Prognosis of wear progress based on regression analysis of condition monitoring parameters

In: Finnish Journal of Tribology 2003.
Vol. 22, 4, pp. 3–15.
(Also published in COMADEM 03,
August 27–29, Växjö, Sweden.)

PROGNOSIS OF WEAR PROGRESS BASED ON REGRESSION ANALYSIS OF CONDITION MONITORING PARAMETERS

Erkki Jantunen

Technical Research Centre of Finland, Industrial Systems,
P.O.Box 1702, FIN-02044 VTT, Finland, Email:
Erkki.Jantunen@vtt.fi,
Phone +358 9 4565319, Fax +358 9 460627

Reprinted from Proceedings of the 16th International Congress COMADEM 2003, 27-29 August, Växjö, Sweden, editors Shrivastav O.P., Al-Najjar B. & Rao R.B.K.N., Växjö University Press, ISBN: 91-7636-376-7 pp. 481-491, with permission from Växjö University

ABSTRACT

For the maintenance personnel the key questions in every day life are: Is everything working properly and if not should we do something? It is especially important to know when some action should be taken i.e. will the machine in question hold until the next scheduled maintenance or does it not? Considering this from the condition monitoring point of view it is important to get a reliable indication of an upcoming failure so early that the necessary maintenance tasks can be planned well in advance and really perform them when the production machinery is stopped for scheduled maintenance. It is not an easy task to predict from measured parameters how quickly the fault will progress. The paper discusses some possible models for the progress of condition monitoring parameters i.e. how the condition monitoring parameters indicate the development of wear as a function of time. The prediction of the development/increase of these parameters is based on regression analysis techniques. The choice of these models is discussed keeping in mind that for practical purposes they should be simple and fast to use. The models are tested with some very common components which suffer from a type of wear which tends to progress with increasing speed towards the end of the life of the component. The first example is from tool wear monitoring where the life of the tool is very short and the measured values usually follow a certain trend and the second example is from a bearing test where the trend of the measured parameter is not that obvious. In both cases the suggested regression analysis technique works very well and can give prognosis of the further development of the monitored parameter.

KEYWORDS

Rotating machinery, wear progress, bearing fault, tool wear, condition monitoring, monitoring parameters, regression analysis, diagnosis, prognosis

INTRODUCTION

In the industry the maintenance personnel need to know when to take action i.e. when it is necessary to carry out maintenance.

Usually, due to economical reasons, the maintenance actions should be performed during a specific period of time when normal production is not disrupted i.e. they should be postponed until the next scheduled maintenance. The big question in this planning of maintenance is: How do the maintenance personnel know how long the machinery keep on running if an indication of some developing fault has been seen e.g. in the condition monitoring signals? Normally the decision whether or not to stop the machinery immediately or whether production may continue is based on the experience of the maintenance personnel. Skilled personnel who have many years of experience might have seen a similar case and can therefore say with some kind of reasonable probability whether the component will hold or not. Unfortunately this kind of diagnosis is not always correct i.e. every now and then the diagnosis is wrong and the production equipment has to be stopped which in turn causes unscheduled maintenance action with very high costs. The problem could be avoided if good methods of prognosis existed, that could well in advance predict how the fault will develop based on condition monitoring data. Unfortunately this is not the case, this kind of models are available but only for a rather limited number of cases. Especially, this kind of wear models are not available for rotating machinery. The wear progress models are not well known and also it is not known how the monitored parameters indicate the wear rate. Another factor that makes the situation even more challenging is the fact that often the start of the wear progress is some odd situation which has possibly only lasted for a very limited amount of time, e.g. the loads have for some time increased to such a high level that wear has started or something has momentarily gone wrong with lubrication so that tribological surfaces have suffered.

Vibration monitoring

The judgement of condition monitoring parameters is typically based on amplitude levels, i.e. if the amplitude of a certain parameter e.g. root mean square (rms) value of vibration velocity in a specified frequency range exceeds a predefined value a fault condition is diagnosed. The diagnosis can be based on broadband analysis i.e. the signal is not filtered [1]. Normally an unfiltered broadband or overall measurement that provides the total vibration energy between 10 and 10000 Hz is used for this type of analysis. The overall analysis does not provide any innovation pertaining to the actual machine problem or failure mode. Changes in both the speed and load of machinery will have a direct effect on the overall vibration levels of the machine, which makes it very problematic in practise to diagnose whether a fault is developing. Narrowband trending, like broadband, monitors the total energy for a specific bandwidth of vibration frequencies [1]. The technique uses vibration frequencies representing specific machine components or failure modes. This method provides the means to quickly monitor the mechanical condition of critical machine components, not just the overall machine condition. The technique provides the ability to monitor the condition of gear sets, bearing and other machine components without manual analysis of vibration signatures. As in the case of broadband trending, changes in speed, load and other process parameters will have a direct, often dramatic, impact on the vibration energy produced by each machine component or narrowband. To be meaningful, narrowband values must be adjusted to the actual production parameters. Unlike the two trending techniques above, signature analysis (frequency analysis) provides a representation of each frequency component generated by a machine [1]. Vibration signatures can be used to determine the specific maintenance required by plant machinery. Most vibration-based condition monitoring programmes use some form of signature analysis in their programme. In this kind of monitoring some kind of warning limits are used. There can actually be a number of limits so that, if the amplitude of vibration at some frequency is below a certain limit, the situation is considered good, and if it then gets higher, it is considered as a warning. The latter case could result in that the interval between measurements is decreased and then if the amplitude exceeds a certain value, it is considered that a fault is present which should be taken care of very quickly. Even more limits could be used, i.e. if the amplitude gets higher than the previous limit, the machine has to be stopped immediately. The vibration standards also recognise some kind of prognosis [2] i.e. if the trend (linear regression) drawn from three last measurements indicate that an alarm limit would be reached before the next scheduled measurement, the situation is considered as a warning. The term alarming rate of change has been used to describe this kind of situation. Naturally, the maintenance personnel are expected to take measures in this kind of situation, e.g. at least additional measurements should be made prior to the next scheduled measurement.

Wear models

Wear of rotating machinery is a very complicated phenomenon since normally there are two surfaces in interaction though they are separated with a lubrication fluid. Basically two types of wear progression can be distinguished i.e. progressive and cumulative [3]. An example of the progressive type of wear process is the wear volume of a plain journal bearing, operating with some metal-to-metal contact. After running-in, there might be a stable period with a constant wear rate, until the bearing clearance is high enough to change the dynamic behaviour of the shaft, causing an accelerating wear process. A ball bearing gives an example of the cumulative type of wear process. After some minor running-in wear, the wear rate is almost zero for a long period of time. During this period, surface fatigue damage accumulates. Fatigue cracks are initiated, and after some time the first metal flakes start to loosen from the surface of a bearing race. In addition to the above, quite often the development of a fault starts when something abnormal takes place either in relation to lubrication or load. When an initial fault has occurred wear usually progresses with an increasing if not exponentially increasing rate. Based on a number of studies, Onsoyen [3] has summarized a simple model for the wear depth shown in Eqn. 1.

$$h(t) = h_o + h't \qquad (1)$$

where h(t) is the wear depth, t is the time, h_o is the contribution from running-in and h' is the wear rate (the increase in wear depth per unit of time). The time to failure is the time t_c until h(t) reaches a critical wear depth h_c. In [4] it was assumed that the wear progression during the tests had been of a progressive type [3] so that the wear behaviour at the beginning was described as mild wear and at the end as severe wear [5]. To fulfil this assumption, a simplified numerical expression for the wear rate was chosen see Eqn 2 [4].

$$h'(t) = A^* t_c/(t_c - t) \qquad (2)$$

where A is a coefficient which does not vary as a function of time t. For simplicity, running-in wear is not accounted for in the above expression. By integrating the above formula, a numerical expression for the wear depth has been developed as shown in Eqn. 3 [4].

$$h(t) = - A^*t_c^*\ln(1 - t/t_c) \qquad (3)$$

Assuming that vibration at some frequency is a function of the physical irregularity of the contact surface, i.e. the fault and initial vibration which is caused by unbalance, loads in the motor etc., vibration follows the format of wear depth shown in Eqn. 3. This kind of development of vibration level is schematically shown in Figure 1.

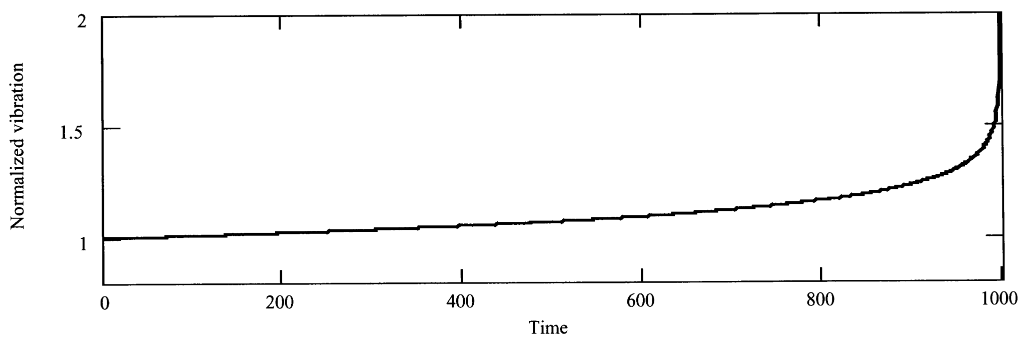


Figure 1. Normalized vibration showing schematically the influence of progressive wear.

It should be noted that, even though the formula in Eqn. 2 chosen for wear rate development is very simple, the curve shown in Figure 1 very well follows the type of wear development curves shown in literature. It also fits well with recorded vibration data, i.e. very often vibration in rotating machinery starts to increase exponentially when a failure has occurred and is developing in size. In wear prognosis the purpose is to be able to diagnose the current state of wear and to predict the development of wear based on possible wear models. Naturally it would be even more tempting to try to develop a method that could be used for many types of components suffering from different types of wear.

POSSIBLE FUNCTIONS FOR WEAR MONITORING AND PROGNOSIS

Most of the diagnosis tools that are used today are based on their capability to recognise or classify the changes in the parameters that they follow. For example autoregressive models can well be used to recognise the change in measured parameters especially if the process is stable, i.e. when the models have been defined based on recorded data they can distinguish when a change has taken place. Assuming that it would also be possible to model the behaviour of condition monitoring parameters in faulty situations as a function of the fault, it would be possible to predict with these models how much time there is left until complete failure of the machinery. Unfortunately this type of approach is not practical because of the amount of modelling involved. It can be claimed that another very popular approach today i.e. Artificial Neural Networks (ANNs) suffer from similar kind of restriction. ANNs are typically good in classification tasks i.e. for diagnosis of changes in the situation. For example if a net has been trained with measured parameters in a good condition of a machine they are capable of recognising a change in the parameters and thus diagnose a possibly faulty situation. In [6] Nandi gives a good comparison of the classification success rate of various ANNs and also Support Vector Machines (SVMs) approaches together with simple thresholding, and methods like Principal

Components Analysis (PCA). It is very interesting to note the trend that if the diagnosis method is very simple but relies on more sophisticated features (bispectrum) the results are rather good, and if the diagnosis method is more sophisticated, more reliable classification results can be reached with more basic features (statistical and spectrum). But then again the opposite seems also to be true, i.e. with sophisticated approaches based on sophisticated features the results are not as good and similarly using simplistic approaches with simple features does not give that reliable results. However, it is very time consuming to try to teach ANNsto recognise different phases of fault development. It is by no means an impossible task but for the method it possibly is not the best way of using them. Neural nets have been successfully used for prediction or prognosis when the approach is based on a number of input parameters and the development of only one or very few output parameters is predicted. It can be claimed that in condition monitoring the goal is quite different. In condition monitoring the purpose is to be able to do prognosis of the development of the health of the machinery in question based on a minimum number of inputs. When building diagnostic systems rule based approaches offer the possibility of effectively programming the rules of thumb used in condition monitoring standards, e.g. if the measured vibration parameter has become twice that it originally was in the beginning of the trending, then a fault is developing and the machine will probably only last one month etc. Now if it were possible to model the wear development as shown in the previous chapter of this paper it would be easier to predict the development of the condition of the machinery. Unfortunately, in the case of rotating machinery it is very seldom the case that the exact wear development of the machinery could be modelled. However, the idea with the use of regression analysis is to be able to adopt the previous development of the monitored parameter and then based on this and the knowledge of typical development of similar cases to be able to predict the future i.e. do prognosis of the future development of the condition monitoring parameter in question and in this manner do prognosis of the development of wear and predict when the machine part will collapse so that the machine will not be able to work properly. In practice condition monitoring should be easy to perform and the number of transducers that are needed should be very limited. Also because of economical reasons the human involvement should be minimal and even the computers used for recording and diagnosing the data should be as cheap as possible so that the monitoring system could in practise be widely used. Based on the above, regression analysis techniques offer a number of advantages that are listed below:

- Smoothens the variation of the data between individual measurements.
- Makes it possible to remember the history of the measured parameter with a limited number of terms stored in a database or a file i.e. it is sufficient to store only the summary terms.
- Makes it easier to notice trends in the data.
- Makes it possible to predict the future, i.e. how much time there is before the component will be totally destroyed.
- Enables percentage prognosis, e.g. it is possible to predict 3 percent in the future in stead of using an arbitrary value like one day that would be the case if the prognosis would be based on prognosis of some amplitude values measured earlier (which would have been measured at constant intervals).

- Makes it possible for the prognosis to be based on the trend of the parameter in question i.e. the prognosis can be based on how the parameter is developing as a function of time, in stead of a single amplitude value which can vary a lot from case to case due to varying loading conditions, structural differences etc.
- The results of the regression analysis can be used as input to different kind of models e.g. ANN and SVM and if the regression models are used for the prognosis i.e. the values of the parameters are predicted then the diagnostic models can be used for prognosis.

Ln Function

The logarithmic wear curve and consequently vibration parameter model would at first sight look rather a promising basis for regression analysis and has actually been tested in [7] but as a function it is very problematic in the sense that parameter t_c has to be known in beforehand or otherwise the mathematics become very laborious and the actual solution of the regression curve would be based on iteration resulting in that the whole idea of using a simple approach with a very limited number of summary terms would be ruined.

3rd Order Polynomial

In [7] the third order polynomial of the type shown in Eqn. 4 proved to be very promising for tool wear monitoring.

$$y(t) = at^3 + bt^2 + ct + d \qquad (4)$$

Where y(t) is the monitored parameter as a function of time, a, b, c and d are regression coefficients and t is time. Similarly, good results with the third order function are reported in [8] for monitoring the development of a bearing fault. However, the third order polynomial regression curve does seem to have some drawbacks. If it is tested against the type of vibration parameter curve shown in Figure 1 it would not be flexible enough to adopt the exponential shape in the end of the life of the component. Another very important factor is that if the regression function has been used for a long time it tends to get very stable, i.e. if in practise, it would have been used for five years for monitoring a bearing in the industry it would take very long time for high parameter values to change the indication of the regression curve. Due to the drawbacks of the 3^{rd} order polynomial another basically as simple regression function has been developed i.e. higher order polynomial that emphasizes current data with a limited number of terms.

Polynomial Model of Higher Degree with Limited Number Of Terms

The idea with higher order polynomial regression function which has a limited number of terms and emphasizes current data, is really to be able to adopt the trend that can be seen in many of the condition monitoring parameters, i.e. exponential growth of the parameter towards the end of the life of the component when wear is taking place with increasing speed, see Figure 1. Higher order polynomial can mimic the ln-function to a certain extent. Figure 2 shows the end of the life-part of the same simulated ln-function as in Figure 1. In Figure 2 the regression analysis is based on data that is supposed to be available when only about three percent of the lifetime remains. Together with the ln-function is shown the prediction made with a higher degree polynomial, the third degree polynomial regression function, and also the first order regression as suggested by the standard [2]. The prognosis made with the higher degree function (e=9, f=6, g=3, constant=1, k=0.95, see Eqn. 5) gives the best estimate what is going to happen even though it does not give a clear view how rapidly the monitored parameter is changing if wear would be taking place as fast as indicated in Figure 1. The reason for third degree polynomial for not to work at all is simply the fact that in this function all the data is equally weighted, i.e. current data is not emphasized and consequently the function reacts very slowly to the change. Based on the above it is suggested that the benefits of polynomial model of higher degree regression analysis that emphasizes current data, are:

- Higher order function reacts sufficiently quickly to the changes for the maintenance personnel to react, even if the fault in the end of the life of the component is increasing in size and severity very rapidly.

- Emphasizing current data is another means to make the analysis quick enough to adapt the current changes. (In fact in the approach given in [2] emphasis is given to only the three last measurements which actually tends to make the method in some cases rather too sensitive, even to the extent that it might be difficult to say how reliable the prognosis is when at one time it shows descending trend and then after the next measurement the situation seems to be critical.)

- Higher order function is especially suitable for rotating machines where the fault, when initiated, often develops with an exponentially increasing rate caused by the fact that when the fault gets bigger the loads get bigger which in turn increases the rate of wear etc.

- Emphasizing current data makes it possible to use the approach also in case of varying loading conditions assuming that the consequences in the amplitudes of the parameters that are used in the analysis are limited, or information of the change of loading condition can somehow be passed to the diagnosis model/system.

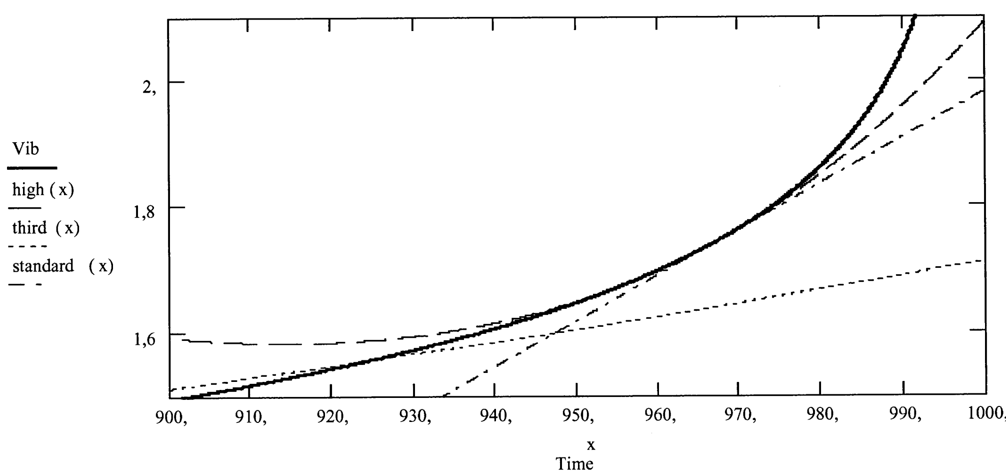


Figure 2. Prognosis of the development of a monitoring parameter based on regression analysis, Vib=data also shown in figure 1, high(x)=higher degree polynomial current data emphasised, third(x) = 3rd degree polynomial, standard(x)=regression as suggested in standard.

It is possible in practise to fine-tune the way the regression analysis emphasises past and current data. If the load varies (indication as a parameter or as a change of relation parameters of which one is more sensitive to the load than the other) it is possible to make the analysis more sensitive to current data so that the functions adapt to the current status more quickly, and then let the emphasizing move to the direction of putting more weight to the history, and consequently make the regression curve more stable. The development of the function given in Eqn. 5 follows the principles given for example in [9].

$$y(t) = at^e + bt^f + ct^g + \text{constant} \qquad (5)$$

Where y(t) is the monitored parameter as a function of time, a, b and c are regression coefficients and t is time. Parameters e, f and g define the degree of the function and there is also a constant in the function. The solution given in [9] is based on the idea of minimizing the sum of the squares of residuals where the residual means the difference between the observed and the estimated response. The minimization of the sum of the squares of residuals is done by finding the partial derivatives for a, b and c. These derivatives are then set equal to 0 to form a system of normal equations. In case of Eqn. 5 there are three unknown terms i.e. a, b and c and three equations. In this solution, in the end there are nine summary terms that need to be calculated and saved for the definition of the regression function. It is often good practice to normalize the parameters that are fed to the regression curve and, as a consequence of that, it is practical to use a constant in the equation that has a value of one/unity. Equation 5 actually becomes a second order polynomial regression function if e is set to 2, f to 1 and g and the constant are set to 0. Similarly the function corresponds to third order polynomial regression function if e is set to 3, f to 2 and g to 1 and the constant is set to 1. However this kind of function is not really the complete third degree polynomial since the constant is given and not calculated which of course could be done if the number of unknowns in the linear set of equation would be increased to four. In order to make regression function more sensitive or aggressive the degree of terms can be increased e.g. e can be set to 9, f to 6 and g to 3. Naturally this kind of function does not have all the features of a complete higher degree function but it is as easy to calculate as the second degree function and still behaves especially towards the end of life of a component very sensitively assuming that the later measurements are emphasised at the cost of the values in the beginning. Introducing a term shown in Eqn. 6 does this.

$$p = k^{(n-i)} \qquad (6)$$

Where n is the current total number of samples, i the index in the calculation summary terms and k is the constant that defines how much weight the early terms get when all the terms in the calculation of summary terms are multiplied with p. Typically k can have a value such as 0.99 if the process is stable with frequent measurements where as a value such as 0.6 would mean that the last measurements are very much emphasised just like the case is with the standard [2]. Actually the method given in the standard [2] corresponds to the use of the first derivative of the regression curve as a means to predict the future assuming that regression curve is behaving in a similar manner as the final three data points suggest. It is suggested that it is practical to use a general form higher order regression polynomial with a constant term of one/unity when the process is stable and the analysis is based on a normalised parameter, i.e. it starts from one, and this way it is possible to monitor the development. However, if the process varies, e.g. because of varying load conditions, it is not practical to use the coefficient in the prognosis because, it in a way stabilises the starting point which is not true. Instead the power of the third term could be zero which in practise means that actually the value of the constant/coefficient, when calculated in this way can have different values as the loading condition varies (which also means that the regression function actually has to be made rather easily adaptive to the current state, i.e. it should not stick to the old value very strongly, i.e. k has a low value.)

TESTS

Two examples of the use of the proposed approach are given. The first example deals with tool wear monitoring which is a very similar problem as that of condition monitoring of rotating machinery. The big difference actually is that tool wear takes place in a very short time scale compared to the wear of machine components. However, the signals being monitored and their behaviour are very similar, and therefore tool wear monitoring is very suitable for testing purposes. In fact, because of the short timescale, monitoring is of even greater importance in the case of tool wear than with condition monitoring of components of rotating machinery. The other example deals with bearing failure, which is one of the most, if not the most common part of rotating machinery that is monitored. The chosen example is somewhat more complicated than an ordinary case even though it is from laboratory tests with constant loading, but it is really showing the potential of the chosen approach to deal with more complicated shapes of signal history.

Tests with drills

Figure 3 shows the results of tests with twist drills (diameter10.2 mm, cutting speed 22 m/min, feed 140 mm/min). The measured parameter is standard deviation of vibration velocity. Figure 3 shows the situation when the twist drill is still in quite good condition and it can still be used. The regression curves of a higher degree polynomial (e = 9, f = 4, g = 0, constant = 0 and k = 0.99) and the third degree polynomial predict that the level of the signal will stay more or less at the same level as has been recorded earlier. Since standard deviation is a very sensitive condition monitoring parameter and varies quite a lot, in this kind of a test the use of linear prognosis based on three last measurements, as suggested in [2], is not practical. The data shown in Figure 4 is from the same test as in Figure 3, only from a later stage of the life of the tool in question. From Figure 4 it is possible to see that both the higher degree polynomial and the linear approach according to reference [2] indicate the end of the life of the tool very well but the third order function seems to react a bit too slow. The main difference between the higher order and third order function is the fact that in the higher order regression function current data is emphasised compared to the preceding data (k=0.99, see Egn. 6). From this example of tool wear monitoring it can be concluded that higher order regression function that emphasises current data with a limited number of terms seems to work very well with this data.

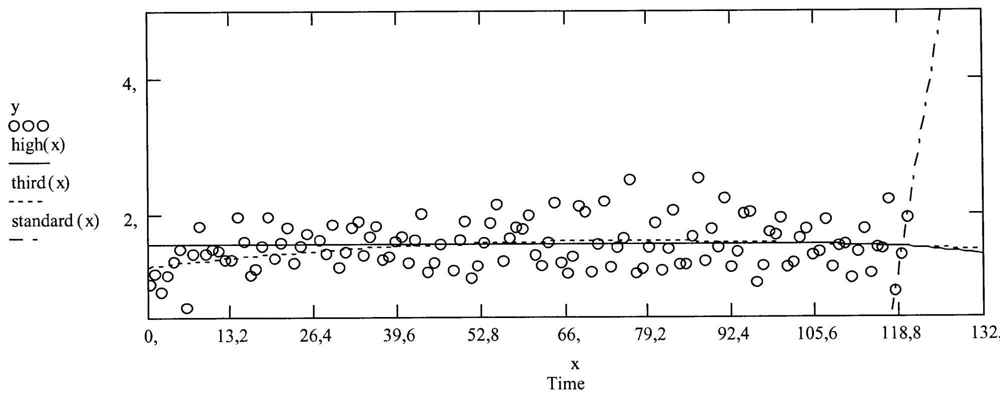


Figure 3. Standard deviation of horizontal vibration velocity of a twist drill, about 1/3 of life time remaining, y=normalised measured values, high(x)=higher degree regression function, third(x)=third order regression function, standard(x)=three point regression according to standard.

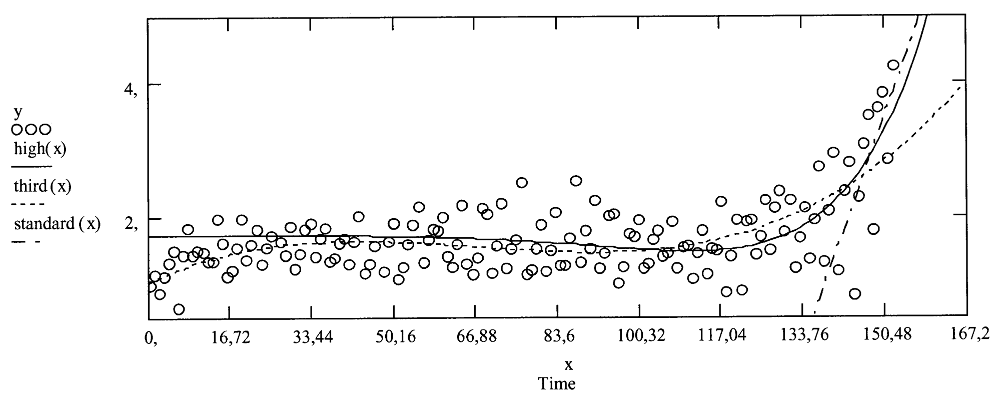


Figure 4. Standard deviation of horizontal vibration velocity of a twist drill, almost complete life time shown, y=normalised measured values, high(x)=higher degree regression function, third(x)=third order regression function, standard(x)=three point regression according to standard.

Tests with bearing data

Figure 5 shows the results of a bearing test in laboratory with a small bearing. The measured parameter is the normalised rms-value of vibration velocity. Together with the measured parameter also the higher degree polynomial regression function (e=9, f=4, s=0, constant=0, k=0.99) and the third degree polynomial regression function as well as linear regression as suggested in [2] are shown. All of the three regression techniques seem to indicate that an immediate increase of the measured parameter could be expected, i.e. the prognosis is that the bearing will suffer from a failure within near future. However, it should be noted that the rate of the increase of the measured parameter at this moment is possibly not that strong that it would mean that the component should not be used anymore.

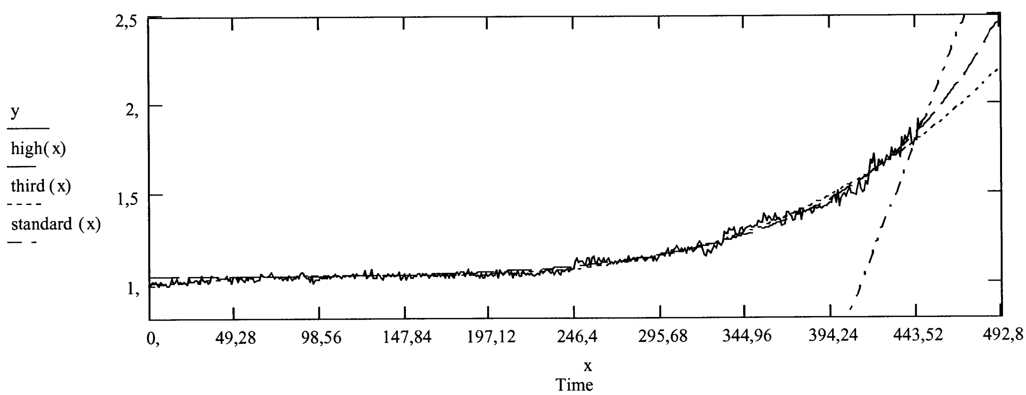


Figure 5. Averaged vibration velocity rms-value from a bearing test, about half the life time of the bearing, y=rms value, high(x)=higher degree polynomial regression function, third(x)=third degree polynomial regression, standard(x)=linear regression based on the last three measured values.

At the specific moment which is studied in Figure 5, linear regression based on the last three measured values gives the highest estimate, higher degree polynomial gives the next highest and third degree polynomial gives the lowest estimate for the following measurement values. This result is very natural since the third degree function gives more emphasis to the past history than higher degree polynomial function and the linear estimation is really based on only the latest data.

Figure 6 shows the results from the same bearing test in laboratory as shown in Figure 5, but now from a much later stage of the test. Together with the vibration velocity rms-value the same regression functions are as shown in Figure 5 are also shown in Figure 6. At this moment in time of the test, linear regression based on the last three measured values gives an indication of rapidly decreasing trend but it should be noted that this type of regression varies a lot if the signal has not been averaged so that the data points represent a long period of time.

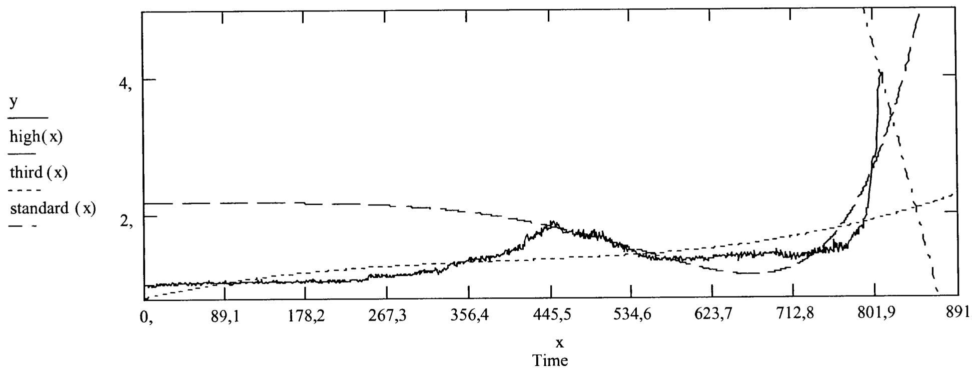


Figure 6. Averaged vibration velocity rms-value from a bearing test, total life time of the bearing, y=rms value, high(x)=higher degree polynomial regression function, third(x)=third degree polynomial regression, standard(x)=linear regression based on the last three measured values.

The third degree polynomial does not really follow the measured values because it gives so much weight to the earlier part of the measured rms-value. The higher degree polynomial, which emphasises the current data more than the data from the beginning of the test, gives a relatively close estimate of the measured signal. In this test the bearing actually suffered from a complete failure at the moment shown in Figure 6. Based on the tested data it could be suggested that the higher order polynomial regression function which emphasises current data and which is calculated with limited number of terms seems to follow the condition monitoring parameters even in a rather complex case so that it can give reasonable estimates of the very near future, i.e. predict the trend of the measured parameter or, in other words, it can be used for the prognosis of the development of condition monitoring signals. It should be noted that when a function is used for the prognosis of the development of a monitoring signal it is not of great importance how closely that signal actually shows what has happened in the past. Another finding is that if a method that is extremely sensitive to current data is used it is important to use averaged data so as to get rid of the extreme variation of the regression function. However there is a problem related to this, i.e. how to define how many points are used in the averaging process so that the function is not made too slow-moving to react to the changes of the measured signal.

CONCLUSION

For the maintenance personnel relying on condition based maintenance it is of great importance to know when they should perform the maintenance actions. Is it possible to carry on with production until the next scheduled maintenance or should the production be stopped immediately? A similar problem exists with machine tools and especially with the cutting tools that are used and changed very frequently. The question is when should the tool be changed, since a worn tool can cause a lot of damage but also the changing of tools too frequently causes excessive downtime and higher tool costs. Due to the complicated nature of wear it is not easy to predict the future, especially since in rotating machinery wear tends to progress exponentially towards the end of the life of the component in question. In this paper some possible models for the development of condition monitoring parameters are given i.e. how the condition monitoring parameters may indicate the development of wear as a function of time. The prediction of the progressive change of these parameters is based on regression analysis techniques. The models have been developed keeping in mind that, for practical purposes, they need to be simple and fast to use. In practise, a high degree polynomial regression function that has a limited number of terms and that emphasises current data seems to work very well with a simulated exponentially developing wear. The developed function also works well in the case of monitoring drill wear and bearing failure. The real benefit of a regression function is that it can into some extent predict the future i.e. give a prognosis of wear development based on condition monitoring parameters.

REFERENCES

[1] Mobley R.K., Condition Based Maintenance. in Handbook of Condition Monitoring, ed. A. Davies, Chapman & Hall, 1998, ISBN 0 412 61320 40

[2] Condition Monitoring, Vibration Measurement, Planning of Condition Monitoring, PSK 5705 standard in PSK-käsikirja 3,Vibration Measurement, Condition Monitoring, 4th ed, 2000, ISBN 951-97098-4-3

[3] Onsoyen E., Accelerated Testing of Components Exposed to Wear, in Operational Reliability and Systematic Maintenance, ed. K. Holmberg and A. Folkeson, Elsevier Applied Science, 1991, ISBN 1-85166-612-5

[4] Jantunen E. and Poikonen A., Dynamics in Monitoring Gear Faults, Proceedings of the 47 th Meeting of Mechanical Failures Prevention Group, 1993, Virginia Beach, Virginia, USA

[5] Holmberg K., Tribological Bases for Accelerated Testing, ed. K. Holmberg and A. Folkeson, Elsevier Applied Science, 1991, ISBN 1-85166-612-5

[6] Nandi A.K., Vibration Based Fault Detection - Features, Classifiers and Novelty Detection, Proceeding of the Condition Monitoring and Diagnostic Engineering Management (COMADEM) conference, ed. Raj B.K.N. Rao & A K. Nandi, 2002 Comadem International

[7] Jantunen E. and Jokinen H., Automated On-Line Diagnosis of Cutting Tool Condition, International Journal of Flexible Automation and Integrated Manufacturing, 4 (3&4), Begell House Inc., 1996, pp. 273-287

[8] Vidqvist V., Development of Vibration Based Condition Monitoring in Rolling Bearing of Reel, Master's Thesis (in Finnish), Lappeenranta University of Technology, 2001, 69 p.

[9] Milton J.S. and Arnold J.C., Introduction to Probability and Statistics, Principles and Applications for Engineering and the Computing Sciences, McGraw-Hill, 3rd ed., 1995, ISBN 0-07-042623-6

PUBLICATION VII

Diagnosis of tool wear based on regression analysis and fuzzy logic

In: IMA Journal of Management Mathematics 2006.
Vol. 17, No. 1, January, pp. 47–60.
Reprinted with permission from the publisher.

IMA Journal of Management Mathematics (2006) **17**, 47–60
doi:10.1093/imaman/dpi027
Advance Access publication on May 20, 2005

Diagnosis of tool wear based on regression analysis and fuzzy logic

ERKKI JANTUNEN†
Technical Research Centre of Finland, Industrial Systems,
P.O. Box 1702, FIN-02044 VTT, Finland

[Received on 29 February 2004; accepted on 24 February 2005]

Tool wear monitoring is important for a number of reasons. Automatic diagnosis of tool wear enables the unmanned use of flexible manufacturing systems and machine tools. Besides, a worn tool if unnoticed could cause a lot of damage, i.e. the machined products could be damaged and unfit for their planned use. As such the machining process is very challenging to monitor due to various reasons. Tool type and cutting parameters may vary resulting in variation of the monitored parameters. Also, there can be a lot of noise in the measured signals. The paper deals with the use of regression analysis techniques together with fuzzy logic in order to overcome the challenges in tool wear monitoring. Regression analysis, based on a higher order polynomial function that emphasizes the most recent measured data and has a limited number of terms, can very well follow and give prognosis of the development of the monitored parameters from such signals as vibration, sound and acoustic emission. The use of fuzzy logic makes it possible to automatically define limits for the monitored parameters and to combine the information from a number of signals. The proposed approach is tested with data from drilling tests.

Keywords: tool wear; drilling; tool condition monitoring; regression analysis; fuzzy clarification; diagnosis.

1. Introduction

Tool wear and failure monitoring is very important due to a number of reasons. Unmanned production is not possible without tool wear and breakage detection. The quality of surface finishing and dimensions of the product can only be guaranteed with proper tool wear monitoring methods. Full benefit from the economical tool life cannot be realized without tool wear monitoring. In principle, there are two approaches in tool wear monitoring, i.e. direct and indirect.

The direct methods include methods that can monitor or measure the tool wear as such. In spite of many attempts, direct methods such as visual inspection or computer vision etc. have not yet proven to be very attractive economically or technically. The challenges with direct measuring methods are related to the demanding environment in machining. The flow of cutting fluid and the flow of removed material (chips) together with mechanical vibration do not encourage the use of sensitive measuring equipment. Naturally, it is possible either to protect the measuring equipment during machining or to do the measurement outside the machine tool, but in both cases the measurement is not taking place on-line and it has an adverse influence on the effective cutting process since machining cannot take place when the tool is being measured, taken away from the machine tool or put back in place again.

Indirect methods include methods like vibration, sound, acoustic emission, force, torque and electrical power measurements. These measuring signals are influenced by the wear of a cutting tool and can

†Email: Erkki.Jantunen@vtt.fi

thus be used for monitoring tool wear. Jantunen (2002) gives a summary of the indirect tool wear monitoring methods in drilling. The summary also covers the used signal analysis and diagnosis methods.

Tool wear monitoring with indirect monitoring methods is a very challenging and demanding task. The cutting parameters may vary, causing variations in the monitored parameters. The work piece material can also vary, causing different kinds of response in the measured signals. In the cutting process, there is a lot of noise in the measured signals due to various reasons, e.g. the flow of cutting fluid influences the measured signals as does the flow of cutting chips. One way to solve the problem, i.e. to reduce noise and the influence of outside factors, is to use sophisticated signal analysis methods. Unfortunately, this kind of approach does not seem very beneficial because tool wear at the end of tool life often increases very rapidly as can be seen in the results of this paper or e.g. in the tests carried out by Subramanian & Cook (1977) and by Pan *et al.* (1993). This tendency of rapid development of wear at the end of the tool life means that when an indication of wear is seen in the signals, there is not necessarily much time left until the complete failure of the tool takes place. In addition, because the number of tools that can be used in one flexible manufacturing system can be really large, i.e. hundreds of tools, the analysis and data handling have to be effective.

The suggested approach in this paper relies on the use of a number of indirect tool wear monitoring methods so as to make the process more robust in the varying cutting conditions. The use of higher order polynomial regression functions is briefly described by Jantunen (2003). The use of regression analysis can solve the problem of saving data for a number of tools and it can also make the signal analysis more stable. To some extent, it also enables prognosis of the development of monitored signals. It is also suggested here that the automatic use of fuzzy classification could be utilized after the regression analysis in making the decision whether the tool is worn or not.

2. Tool wear

Tool wear, and especially drill wear, is a rather complicated phenomenon. In their review of tool condition monitoring in milling, turning and drilling, Rehorn *et al.* (2004) explained that drilling operations differ significantly from turning and face milling for several reasons. The major difference is the fact that drilling is a complex 3D material removal operation, unlike the relatively simple cases of orthogonal and oblique cutting. Drills also have vastly different geometries than turning and face milling tools. They are usually much longer than a turning cutter and have far less cross-sectional area than a face milling cutter. Drilling operations are also different in that they require the full immersion of the tool, rather than operating on the periphery or surface as is the case in face milling and end milling.

Thangaraj & Wright (1988) explained that in principle, drill wear is an accelerating process which takes place at the outer margin of the flutes of the drill due to the intimate contact and elevated temperatures at the tool work piece contact. They also point out how there is a period of initial wear, then a period of moderate wear and in the third phase a period of excessive wear. El-Wardany *et al.* (1996) point out that due to production variations, a drill is typically slightly asymmetric. Accordingly, the two corners of the drill point wear gradually while maximum wear alternates from one cutting edge to the other. This alternating process continues until both lips have zero clearance at the margin. The drill then adheres to the work piece and breaks if the cutting process is not stopped in time. In addition, chip flow creates significant friction between the cutter and the work piece inside the drill hole. Rehorn *et al.* (2004) point out that these frictional forces can significantly change the dynamics of the system and they can cause the cutter to break. Drills, like other cutters, can fail either from breakage or excessive wear. Thangaraj & Wright (1988) determined that drills of a diameter less than 3 mm tend to fail by fracture, while larger tools will fail by excessive wear.

It is generally known and accepted that cutting forces increase as tool wear increases, see e.g. Subramanian & Cook (1977), Pan *et al.* (1993) or Lin & Ting (1995). The same references also show the logical dependency between cutting forces and wear. The increase of cutting forces due to increased wear is then the cause of the increase of wear rate although the whole process is more complicated if also the dynamics and the material properties of the tool were taken into account. From the above, it follows that drill wear is accelerating, i.e. the wear rate increases as the wear progresses. Therefore, it is possible to use a similar simplified theoretical approach as has been used in the case of rotating machinery when defining how wear depth could develop as a function of time, i.e. the shape of this development.

Based on a number of studies, Onsøyen (1991) has summarized a simple formula for the wear depth shown in (2.1).

$$h(t) = h_0 + h't, \tag{2.1}$$

where $h(t)$ is the wear depth, t is the time, h_0 is the contribution from running-in and h' is the wear rate (the increase in wear depth per unit of time). The time to failure is the time t_c until $h(t)$ reaches a critical wear depth h_c. It should be noted that (2.1) does not have any physical parameters in it which means that it is very basic and simply defines the relationship of wear and wear rate. When this formula is used for tool wear then, instead of using the term running-in wear depth, it might be more relevant to assume that there are originally differences in the dimensions of the unused tools, i.e. they are not absolutely symmetrical and their dimensions do not absolutely fulfil the defined geometry. Jantunen & Poikonen (1993) assumed that the wear progression (of rotating machinery, gear) during the tests had been of accelerating type, see e.g. Onsøyen (1991) for definition, so that the wear behaviour at the beginning was described as mild wear and at the end as severe wear, see e.g. Holmberg (1991) for definition. To fulfil this assumption, Jantunen & Poikonen (1993) chose a simplified numerical expression for the wear rate; see (2.2).

$$h'(t) = At_c/(t_c - t), \tag{2.2}$$

where A is a coefficient which does not vary as a function of time t. For simplicity, running-in/(geometrical tolerance) wear is not accounted for in the above expression. By integrating the above formula, Jantunen & Poikonen (1993) developed a numerical expression for the wear depth shown in (2.3).

$$h(t) = -At_c \ln(1 - t/t_c). \tag{2.3}$$

Again, it should be noted that (2.2) and (2.3) are not physically explaining tool wear but are numerical expressions which as a function of time try to mimic the trend seen in tests with cutting tools as a function of wear. The reason for using this kind of numerical approach for drill wear is simply the fact that on the basis of literature studies there does not exist a published wear model for twist drills.

Assuming that vibration at some frequency (e.g. natural frequencies of the drill can be seen) is a function of the physical irregularity of the contact surface, i.e. the fault/wear (influencing the cutting and chip formation process), and initial vibration which is caused by unbalance, loads in the motor etc., vibration follows the format of wear depth shown in (2.3). Jantunen (2003) defined this kind of development of vibration level as shown schematically in Fig. 1.

Even though the formula in (2.2) chosen for wear rate is very simple and does not physically describe what is taking place when a drill is getting worn, it should be noted that Fig. 1 describes quite well in principle the development of vibration or any other similar monitoring signal in tool wear reported in the literature, see e.g. the test results of Subramanian & Cook (1977) and Pan *et al.* (1993). In the case of vibration and forces that are measured perpendicular to the axis of the drill, it is well worth remembering that in theory an ideal drill that has two cutting lips is balanced in this direction. Therefore,

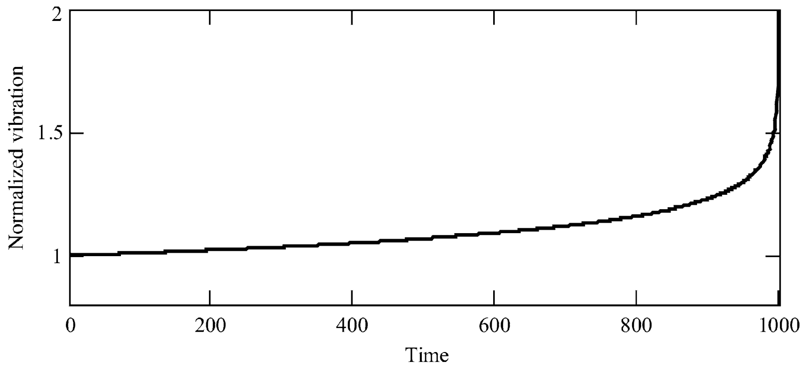


FIG. 1. Normalized vibration showing schematically the influence of progressive wear.

the forces and vibration should not actually exist but due to the dynamics and irregularities the type of development shown in Fig. 1 can be seen in practice. It should also be noted that when the purpose is to give prognosis of the future development of a monitoring signal, and consequently the wear of the component in question, it is essential to know the shape of the wear function as e.g. shown in Fig. 1.

3. Tool wear monitoring

There are so many influencing factors due to random noise, variation of cutting parameters and cutting tool and work piece material that tool wear and failure monitoring is a very demanding task. As Jantunen (2002) showed in the summary based on 31 references in which tool condition monitoring in drilling have been studied, quite a number of different methods have been tested and suggested. The most widely tested method is the measurement of the feed force but it could be argued whether feed force measurement can be done so effectively in practice that it could be in daily use. Vibration, sound and acoustic emission are easier to measure when the sensor type and its positioning is considered. In their study based on about 25 drill monitoring related references, Rehorn *et al.* (2004) came to the same conclusion about the popularity of monitoring methods. The similarity of conclusion is very natural since both reviews use mostly the same references which have been available in this field of research.

Jantunen & Jokinen (1996) tested a number of measuring methods together with a number of signal analysis methods. The conclusion was that vibration was the most effective method in monitoring tool wear, and that more complicated signal analysis methods such as fast Fourier transform (FFT) gave a more reliable indication of tool wear than the simpler statistical parameters. However, the problem with more sophisticated analysis methods is that the analysis takes time and in that sense the more sophisticated approaches might sometimes be too slow to react to the rapid degradation of the tool. Also, the amount of data that needs to be saved during the monitoring might be excessive to be handled in case a large number of tools are used. Figure 2 gives an example of statistical measuring data. Four drills have been tested and the statistical parameter shown is the normalized vibration velocity root mean square value (rms-value) on a logarithmic scale. It is evident that there is a lot of variation from one tool to another. The tool life varies as does the measured parameter value. Although some of the variation is due to the difference of cutting parameters, there is also remarkable variation from test to test shown in Fig. 2 and also reported by Jantunen & Jokinen (1996) which means that it is not realistic to expect that

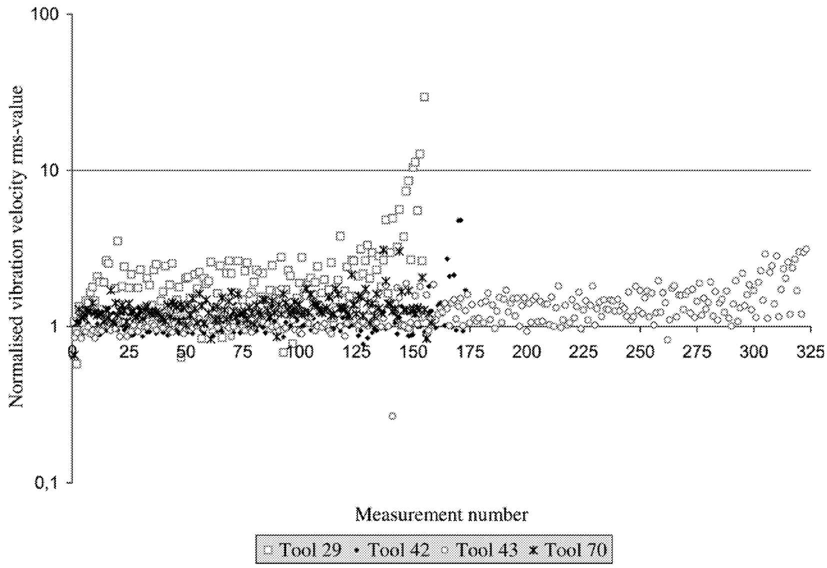


FIG. 2. Vibration velocity rms-value from four drilling tests.

some absolute level of such a simple parameter could give an indication that the tool should be changed due to wear.

El-Wardany *et al.* (1996) described similar difficulties with vibration monitoring in the following way: 1) Materials such as cast iron are not homogeneous and will affect the amplitude of the vibration measured, and this may cause false alarms. 2) Tool damage in drilling produces a high level of transient vibrations (spikes) which are largely attenuated by the averaging procedure typically used in spectrum calculation, and this makes it difficult to extract a discriminating feature to distinguish the change in the tool conditions. 3) Non-uniform hardness of the work piece material, built-up edges and micro-cracks can also cause false alarms by increasing the vibration amplitude. In order to minimize the effects of these difficulties when monitoring the change in drill condition, El-Wardany *et al.* (1996) have used averaging, varying the number of averages from 6 to 10 depending on the phase of the drill life.

4. Polynomial regression model of higher degree with limited number of terms

One possible way to handle the problem described above of saving a lot of data is to use regression analysis techniques which results in that only the summary terms need to be saved for each tool and each analysed parameter. A higher order regression function can actually quite well mimic the wear development or the development of a monitoring parameter shown in Fig. 1. Jantunen (2003) suggested the use of polynomial regression models of higher degree that emphasize the most recent data, with a limited number of terms.

The development of the function described by Jantunen (2003) follows the principles described e.g. by Milton & Arnold (1995). Jantunen (2003) suggested the use of the following type of regression function.

$$y(t) = at^e + bt^f + ct^g + d, \tag{4.1}$$

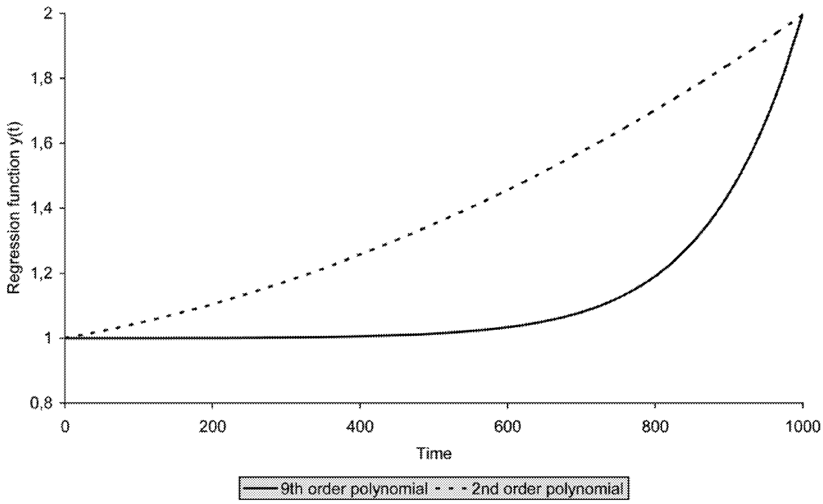


FIG. 3. Comparison between second-order and ninth-order polynomial regression functions with limited number of terms.

where $y(t)$ is the monitored parameter, as a function of time. The parameter can be either a statistical time domain parameter such as rms-value or an amplitude value at specific frequency if FFT has been used. In this equation a, b and c are regression coefficients and t is time. Exponents e, f and g define the degree of the function and there is also a constant d in the function. Jantunen (2003) suggested that relatively high values, such as $e = 9$, $f = 6$ and $g = 3$ give good results.

Figure 3 shows a comparison of a second-order polynomial ($e = 2$, $f = 1$, $g = 0$ and $d = 0$) and a ninth-order polynomial ($e = 9$, $f = 6$, $g = 3$ and $d = 1$) regression function with limited number of terms. As can be seen from Fig. 3, the higher order function has a rather similar shape as the wear function in Fig. 1 and it can be expected to be fast enough to react to the rapid changes close to the end of the life of the monitored tool. The optimal order of the polynomial regression function has not been defined, but based on trial and error with limited number of test data, polynomial functions of sixth or ninth order have given good results while second- and third-order functions react too slowly. It should be noted here that the original logarithmic wear function shown in (2.3) is not suitable for regression analysis because it would be very complicated to be handled since it has the time to failure term t_c inside the logarithm.

It could be claimed that a simple exponential function could also be used so that the exponent would be calculated as one of the regression coefficients. This kind of a function really reacts quickly enough and also has a similar shape as the normalized vibration in Fig. 1. However, the drawback of the exponent function could actually be that it is too sensitive and, consequently, with noisy data it could react too quickly and thus give unreliable indications. Certainly, a number of other possible functions exist that could be tested for the purpose of mimicking the development of wear and those parameters used for monitoring it but it is not in the scope of this paper to try cover the whole range of possibilities. Instead, the purpose is to suggest a working solution which is much closer to reality than e.g. linear regression with the last three measurements which is a rather widely used approximation of the development of the monitoring parameters.

For emphasizing the most recent data, Jantunen (2003) introduced a factor that is used when the summary terms in regression analysis are calculated.

$$p_i = q^{(n-i)}, \tag{4.2}$$

where n is the current total number of samples, i the index in the calculation of the summary terms and q is a constant that defines how much weight the earlier terms are given when all the terms in the calculation of the summary terms are multiplied with p. The most important reason for the introduction of the factor q is that regression analysis functions tend to get very stable, i.e. they do not react to current data very rapidly if they have been used for some time with similar data. This lack of response is naturally very contradictory to what has been presented in Section 2 about the rapid development of wear towards the end of tool life, and hence the introduction of this factor q is needed.

Typically q can have a value such as 0.99 if the process is stable with frequent measurements whereas a value such as 0.6 would mean that the last measurements are very much emphasized. Again, the suggested values are based on limited testing with a trial and error approach. However, the use of weighting function worked also well with bearing data that Jantunen (2003) has tested. In order to keep the summary terms from getting too big for practical reasons, they can be scaled down by dividing them all with a suitable number. Such a limit could e.g. be two, i.e. whenever t gets higher than two, it is divided by two and in this manner kept small and consequently the summary terms will not grow too big.

The benefits of the suggested approach are:

- A higher order function reacts sufficiently quickly so that a worn tool can be noticed in time. The data shown by Jantunen (2003) compare linear, second-order, third-order and higher order polynomial regression functions and the results are promising as are the results shown later in this paper.
- Emphasizing the most recent data is another means to make the analysis quick enough to adapt to the current changes. This means that it is possible to handle the change of the cutting parameters and the change of work piece in this way, assuming that it is accepted that the diagnosis can start some time after the change has taken place and also assuming that information of the change is passed to the measuring system. It should also be remembered that a regression function tends to become slow to react to changes if a lot of data in a constant situation have been gathered without any changes in the monitored parameter and this kind of phenomenon can also be avoided with the introduction of the weighting function shown in (4.2).
- A higher order function is especially suitable for tool wear which, towards the end of the tool life, develops with an exponentially increasing rate, due to the increasing wear leading to increasing loads which in turn increases the rate of wear and so on as explained in Section 2 of this paper.
- In practice, the suggested higher order function is very easy and fast to calculate and only nine summary terms need to be saved. In order to keep the summary terms from getting too big, for practical reasons they can be scaled down by dividing them all by a suitable number. Such a limit could e.g. be two, i.e. whenever t gets higher than two, it is divided by two and in this manner kept small and consequently the summary terms will not grow too big.

When considering the possible drawbacks of using a higher order polynomial regression function with limited number of terms, it could be argued that higher order polynomial regression functions tend to be unstable especially with noisy data. This is actually a drawback but the use of a limited number of terms improves the situation at the same time as it makes the calculation more simplified. Another possible weakness is the problem that higher polynomial dependencies increase the relative error. Again this is a drawback but unfortunately there is no way avoiding this because of the nature of the problem.

The development of the monitored parameters tend to be very progressive and whichever regression function was used, the same problem would exist if the regression analysis was able to follow the trend of the monitored parameter because here we have a phenomenon where at first there are very small changes taking place and then after a long period of time the changes are taking place with an increasing speed.

The use of the weighting function does also help because with that the influence of data gathered long before the time of monitoring is not so remarkable. In fact, it can be claimed that in the developed approach the regression analysis function actually acts as a filter that removes some of the unwanted variation of the measured parameters and then gives prognosis of the trend in the measurement. This trend can be seen in the analysed examples shown later in this paper. Naturally, if only time-series data smoothing would have been the target of the data manipulation, a much more simplified function would have been available such as described by Williams *et al.* (1994) when they give examples of the use of moving average or exponential smoothing in condition monitoring. The biggest difference between the suggested approach and those very simple methods is that they do not give prognosis of the forthcoming trend of the monitored parameter. With this restriction, simple smoothing techniques do not react as quickly to the changes of the monitored parameter. On the other hand, the suggested approach in this paper is so simple that there are no practical problems with it in relation to the time it takes for a cheap and low speed processor to calculate the regression analysis of higher order polynomial functions with limited number of terms.

5. Fuzzy classification

In order to automatically classify the results of the higher order polynomial regression function, a very simple fuzzy classification approach following the example shown by Rao & Rao (1993) has been used. In principle, the idea is that in the beginning, when data are measured from a tool that is in good condition, some of the early data are used for the definition of fuzzy classification limits for the analysed parameters. In the approach, the number of classes is limited to eight with the assumption that class 2 means that the tool is in good condition. Higher classes then mean that the monitored component is getting into a worse condition. Class 1 is reserved for lower values of the monitored parameter, meaning possibly that the cutting conditions are not similar as in the beginning when the limits were defined in the first place.

In the definition of the classes, it is logical to use the mean and the standard deviation of the parameter in question for the definition since there is always variation in the monitored parameters and it can be expected that this variation can be used as a basis for the future change of the parameter as a function of wear. Naturally, this kind of choice has to be tested with the mechanical component, monitoring the signal and parameter in question so that proper choice of sensitivity can be made. It is probably worth noticing that it is typical to use mean value and variation or standard deviation of the monitored parameter when so-called health indices are defined, e.g. the health index shown by Williams *et al.* (1994) uses these parameters. Similarly, these parameters are used for the definition limits in vibration data trending in international and national condition monitoring standards such as PSK 5705 Standard (2004).

In the developed approach, the classes are defined using the following definitions. The mean value of each class (class index $i = 1, \ldots, 8$) is defined according to the following formula:

$$\text{ClassMean}_i = (i - 2)j\sigma + \mu, \tag{5.1}$$

where j is a coefficient defining the size of the classes, k is a coefficient that defines the shape of the classes and μ is the mean value and σ the standard deviation of the first measured parameters.

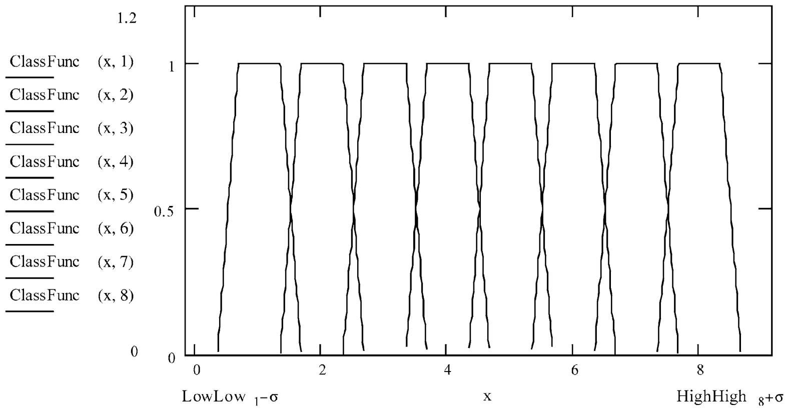


FIG. 4. An example of classification limits.

The upper and lower limits of the classes are defined as follows:

$$\text{LowLow}_i = \text{ClassMean}_i - j(1+k)(\sigma/2). \tag{5.2}$$

$$\text{LowHigh}_i = \text{ClassMean}_i - j(1-k)(\sigma/2). \tag{5.3}$$

$$\text{HighLow}_i = \text{ClassMean}_i + j(1-k)(\sigma/2). \tag{5.4}$$

$$\text{HighHigh}_i = \text{ClassMean}_i + j(1+k)(\sigma/2). \tag{5.5}$$

Figure 4 shows an example of the classification limits ($j = 2$, $k = 0.33$, $\mu = 2$ and $\sigma = 0.5$). Assuming that monitoring is based on vibration measurements and the monitored parameters are statistical values calculated from the time domain data, it can well be assumed that limits can be automatically defined based on e.g. the 20 first measured values. In practice, the most important parameter to define in this automatic approach is the value of coefficient j which defines how sensitive the approach really is in diagnosing wear. In this paper different values of coefficient j are tested with data from drilling tests. It should be noted that for practical purposes the first class can be defined to start from minus infinity, and corresponding to that, the highest class can be defined to go to infinity. Also, only half of the limits need to be saved in the memory since the HighLow value of a lower class is the same as the LowLow of a higher class etc. as can easily be seen in Fig. 4.

It could be argued that the introduction of fuzzy logic does not bring a lot of advantages, if any, in such a simple case as the classification of the drill monitoring data. In fact, if $k = 0$ was used, this would actually mean that the limits would become crisp in the above equations (5.2)–(5.5) which is the same as no fuzziness at all. However, it can be claimed that in real life the limits actually are fuzzy and we cannot define exact limits for this type of classes. Also, using fuzzy limits brings some benefits in the following steps of the approach. In the automatic approach that has been developed for tool wear monitoring, the step after the fuzzy classification is the use of neural networks for distinguishing between various cutting conditions and at that stage the introduction of fuzzy classes has a beneficial influence in making the neural net model more robust. The use of fuzzy logic in pre-processing the input data follows the principles shown by Rao & Rao (1993).

As discussed in Section 3, the measured data in tool wear monitoring are noisy in practice due to the nature of the machining process. One possible way to handle this and to make the monitoring more

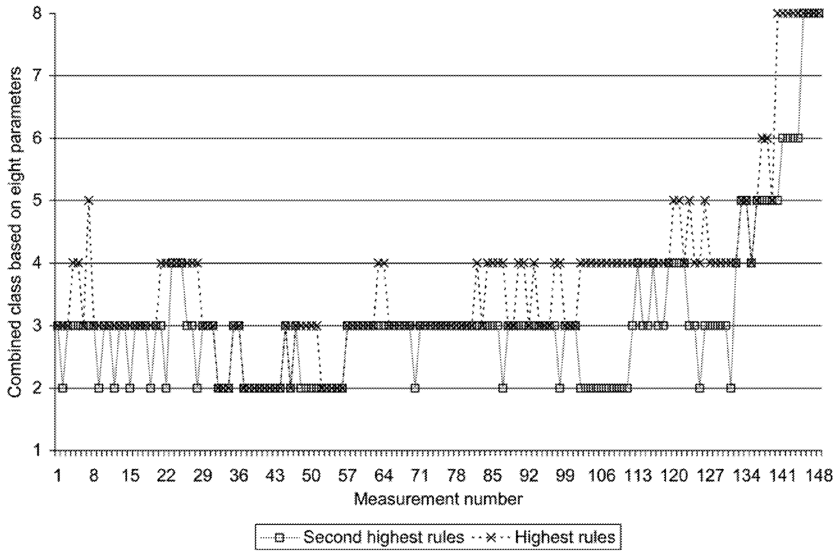


FIG. 5. Comparison of simplified rules for combined classification of tool wear.

reliable is to monitor a number of signals and parameters simultaneously, instead of monitoring and relying on just one single signal. Figure 5 shows an example of tool wear monitoring where the diagnosis of the state of the tool is based on three different signals (vibration, sound and acoustic emission), and altogether eight statistical parameters (rms, mean deviation and maximum for vibration and sound, and rms and mean deviation for acoustic emission) have been derived from these signals. One of the curves is based on the highest indication of the eight monitored parameters. In the other curve the final conclusion is based on a very simple idea, i.e. it is assumed that if only one parameter indicates that the tool is worn, this is not considered sufficient but instead it is expected that at least two parameters have to give this indication. This in fact is the same as saying that the second highest class of the calculated classes of eight signals is the one that rules. Both curves have been calculated using a higher order regression function. In the regression analysis, the ninth-order polynomial with the sixth- and third-order terms and a term of unity (because of normalization) has been used. In addition, in both curves, the factor q has been chosen as 0.99 and the classes have been defined so that j is 1 and k is 0.5 in (5.1)–(5.5). In both curves, the calculation of the mean value and standard deviation is based on the 25 first values. The tool in question is tool number 42 which is one of the tools for which data are shown in Fig. 2. In this case the simple classification rule, where the indication of at least two parameters is needed for the conclusion that the tool is worn, seems to give a more consistent result, i.e. the combined class does not vary a lot and reaches class 8 in the last measurements when the tool is worn.

Figure 6 shows the influence of the definition of the parameter j in (5.1)–(5.5) defining the fuzzy limits. In the three chosen cases, j has been 1, 2 and 3 while k has been kept constant, i.e. 0.5 in all of the cases. In the example shown in Fig. 6, the combination with $j = 1$ seems to work best and to distinguish best the difference between worn and unworn tools, since with $j = 1$ class 8 is reached at the end of the drill life, whereas with $j = 2$ only class 6 is reached, and with $j = 3$ only class 5 is reached. In practice, it would probably be advantageous to choose the parameters in a conservative way, i.e. in such a way

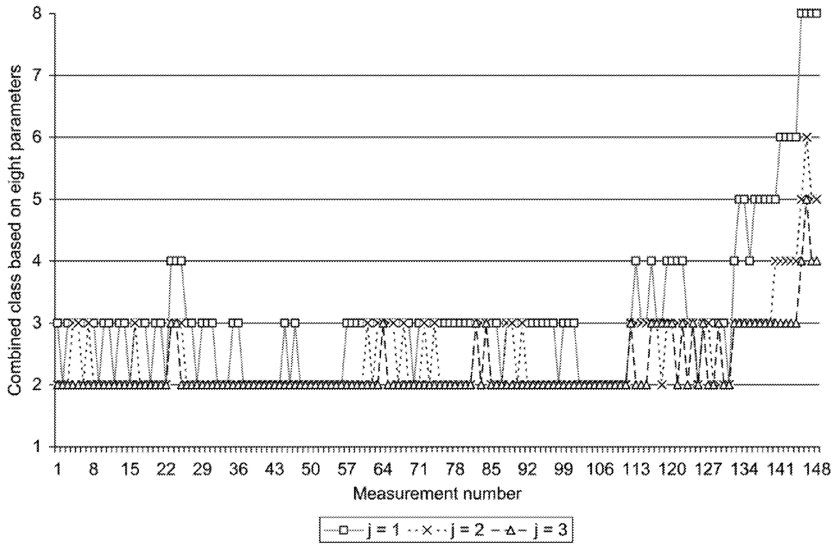


FIG. 6. Comparison of three different automatically defined classification limits.

that the indication of a worn tool will be seen rather too early than too late, and hence lower values of j are suggested in order to be on the safe side. All the other parameters remain the same as in the previous example. In this case, the classification rule of relying on two indicators is used in all the three cases.

Figure 7 shows a comparison of the ninth- and third-order regression functions together with a classified case where no regression function has been used. From the figure it is apparent that the use of a regression function makes the analysis smoother, i.e. it filters out the effect of individual measurements which in tool wear monitoring tend to vary a lot. There is not that much difference between the ninth- and third-order regression functions with limited number of terms and hence either one could be used. Logically, the use of a regression function also gives the ability of predicting the future, i.e. it makes it possible to give prognosis of the future development of the monitored signal. In this sense, the ninth order has been better in some tests because it reacts to changes more quickly than the third-order function but the difference is not very big. In the analyses shown in Fig. 7, the simple classification rule of following the second highest value has been used, the factor q has been 0.9 and j and k have been the same as in the previous example (i.e. $j = 1$ and $k = 0.5$).

Figure 8 shows the influence of q, the factor that emphasizes the current data at the cost of older data. In the two examples, the values of 0.9 and 0.99 have been used. Naturally, a lower value of q makes the curve more sensitive. In the example of the third-order regression function, the simple rule of relying on the second highest parameter value and the values $j = 1$ and $k = 0.5$ have been used. It should be noted that the use of a higher order function together with a higher value of q has a similar kind of influence as using a lower value of q together with a lower order regression function, i.e. the sensitivity of the response to changes in the measured signal or parameter is rather similar. However, it can be argued that a higher order regression function can possibly mimic the shape of the simplified wear curve shown in Fig. 1 more closely and thus give a somewhat better prognosis of the remaining lifetime of the tool in question, see e.g. Fig. 5.

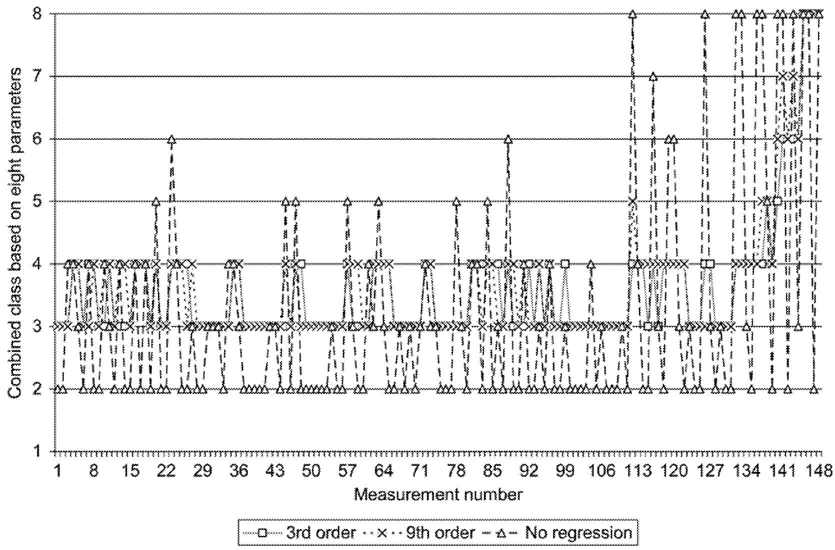


FIG. 7. Diagnosis with third- and ninth-order regression functions compared with analysis without a regression function.

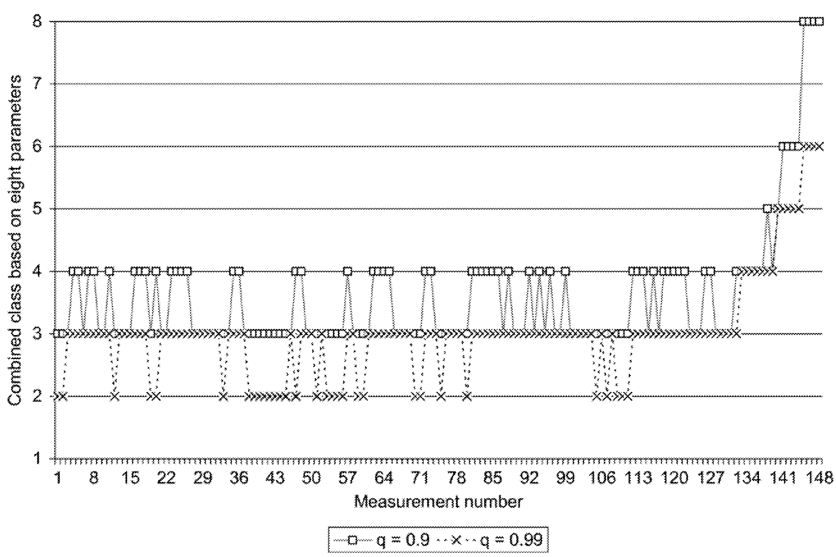


FIG. 8. Comparison of the influence of the factor q which defines how much emphasis more recent data has in the regression analysis.

6. Conclusion

Tool wear and breakage monitoring is important for economical reasons. Without reliable monitoring methods, the unmanned use of flexible manufacturing systems is not possible and the quality of products cannot be guaranteed. In order to be able to diagnose and to do prognosis when the tool is worn, the progressive development of tool wear and its influence on the monitored parameters are described. The use of a higher order polynomial regression function is suggested. With the use of the regression function, it is possible to reduce the amount of data that needs to be saved when automated monitoring is performed. The regression analysis also reduces the variation of noisy data from a typical machining process based on simple statistical parameters calculated from time domain data. Since there is a lot of variation from test to test in these simple parameters, it is also suggested that fuzzy logic can be used to classify these parameters. An automatic approach for the definition of classification limits has been developed and tested. Finally, it is concluded that the diagnosis whether a tool is worn and should be changed should in practice be based on a number of signals and a number of calculated parameters instead of one monitoring signal and only one calculated parameter. This conclusion is simply based on the test results where usually there is no similar variation in all of the parameters even though there might be great variation in one measured signal and parameters calculated from it. Although there are quite a number of steps in the approach, the methodology is easy to program and does not need a very powerful processor or a lot of memory in this device. In addition, all the steps can be performed automatically, i.e. the user does not need to interfere in the process when the developed computer program is used.

Acknowledgements

This work is a part of the project Multiple Intelligent Diagnostics for MAchiNery (MindMan) funded by the Tekes (National Technology Agency of Finland) and industry Metso Paper, Rautaruukki Raahe Steel Works, Fläkt Woods Oy, Danisco Sweeteners, Oy Sinebrychoff Ab and HK Ruokatalo Oyj. The efforts of research scientists Natalia Siren and Seppo Tenitz in programming and running of the test cases is greatly appreciated.

REFERENCES

EL-WARDANY, T. I., GAO, D. & ELBESTAW, M. A. (1996) Tool condition monitoring in drilling using vibration signal analysis. *Int. J. Mach. Tool Manu.*, **36**, 687–711.

HOLMBERG, K. (1991) Tribological bases for accelerated testing. *Operational Reliability and Systematic Maintenance* (K. Holmberg & A. Folkeson eds). London and New York: Elsevier Applied Science.

JANTUNEN, E. (2002) A summary of methods applied to tool condition monitoring in drilling. *Int. J. Mach. Tool Manu.*, **42**, 997–1010.

JANTUNEN, E. (2003) Prognosis of wear progress based on regression analysis of condition monitoring parameters. *Proceedings of the 16th International Congress of COMADEM*. pp. 481–491. Växjö, Sweden: Växjö University Press.

JANTUNEN, E. & JOKINEN, H. (1996) Automated on-line diagnosis of cutting tool condition. *Int. J. Flex. Autom. Integr. Manuf.*, **4**, 273–287.

JANTUNEN, E. & POIKONEN, A. (1993) Dynamics in monitoring gear faults. *Proceedings of the 47th Meeting of Mechanical Failures Prevention Group*, Virginia Beach, VA. Willowbrook, Illinois: Vibration Institute.

LIN, S. C. & TING, C. J. (1995) Tool wear monitoring in drilling using force signals. *Wear*, **180**, 53–60.

MILTON, J. S. & ARNOLD, J. C. (1995) Introduction to probability and statistics. *Principles and Applications for Engineering and the Computing Sciences*, 3rd edn. Singapore: McGraw-Hill.

ONSØYEN, E. (1991) Accelerated testing of components exposed to wear. *Operational Reliability and Systematic Maintenance* (K. Holmberg & A. Folkeson eds). London and New York: Elsevier Applied Science.

PAN, H., CHEN, Y. & ORADY, E. (1993) Monitoring methods of tool wear in a drilling process. *Proceedings of the 2nd Industrial Engineering Research Conference*. pp. 380–384. Institute of Industrial Engineers.

PSK 5705 STANDARD (2004) *Condition Monitoring, Vibration Measurement, Planning of Condition Monitoring*. Helsinki: PSK Standardisointiyhdistysry.

RAO, V. B. & RAO, H. V. (1993) *C++ Neural Networks and Fuzzy Logic*. New York: Management Information Source, Inc.

REHORN, A. G., JIANG, J. & ORBAN, P. E. (2004) State-of-the-art methods and results in tool condition monitoring: a review. *Int. J. Adv. Manuf. Tech.* (published online 11 August 2004).

SUBRAMANIAN, K. & COOK, N. H. (1977) Sensing of drill wear and prediction of drill life. *J. Eng. Ind.*, **101**, 295–301.

THANGARAJ, A. & WRIGHT, P. K. (1988) Computer-assisted prediction of drill-failure using in-process measurements of thrust force. *J. Eng. Ind. Trans. ASME*, **110**, 192–200.

WILLIAMS, J. H., DAVIES, A. & DRAKE, P. R. (1994) *Condition-Based Maintenance and Machine Diagnostics*. London: Chapman & Hall.

PUBLICATION VIII

Flexible hierarchical neuro-fuzzy system for prognosis

In: Proceedings of COMADEM 2000, 13th International Congress on Condition Monitoring and Diagnostic Engineering Management. Ed. by H.C. Pusey & Raj B.K.N. Rao. Houston, USA, December 3–8, 2000. Pp. 699–708.

FLEXIBLE HIERARCHICAL NEURO-FUZZY SYSTEM FOR PROGNOSIS

Erkki Jantunen
Technical Research Centre of Finland
Manufacturing Technology
P.O.Box 1702
FIN-02044 VTT, Finland
Erkki.Jantunen@vtt.fi

Abstract: An easy to configure hierarchical neuro-fuzzy system has been defined for the configuration of a prognosis system for condition monitoring of machinery. The system consists of a number of modules: data acquisition, signal processing, data handling, fuzzy classifier and a neural net for diagnosis. Data acquisition is based on the use of an AD card, and signal processing on the use of traditional FFT. The fuzzy classifier together with the neural network is organised in a hierarchical structure, which enables the easy configuration of the whole system. The approach is especially flexible in the sense that the total number of parameters the system can handle is not limited in practice. In the hierarchical structure the individual sub-models are restricted to handling eight fuzzy inputs simultaneously. In the system, the type of neural networks can be chosen from a list of choices based on the desired type of behaviour. In a normal case, parts of the hierarchical system are configured based on crisp information. Similarly, the features of neural nets are not used in all of the sub-models and they can be substituted with arithmetic expressions if there is no need for handling non-linear information or the behaviour is well known and can be easily defined otherwise.

Keywords: Condition monitoring, Diagnosis, Fuzzy logic, Neural networks, Prognosis, Signal processing

Introduction: Condition monitoring of rotating machinery has become increasingly popular in recent years as a result of better understanding of the financial values involved [1]. However, when organising condition monitoring in the plant environment, even though the transducers are still not cheap and cabling can be even more expensive, the problem in practice is the amount of work often involved in analysing the monitoring data. The analysis work is also very demanding, and it takes time to train people to a sufficient level of experience so that analysts become real professionals. Basically for the above reasons, quite a number of attempts have been made to automate the whole analysis and diagnosis procedure. The first kind of automatic analysis tools were rule-based expert systems. The rule-based approach as such can be considered in principle to be rather generic, assuming the developers have taken into account all possible situations which can occur with the machinery in question. However, herein already lies the problem in practice: only well-defined situations can be handled, and this in turn pushes the solution towards working only with very simplistic machinery. One way to overcome this restriction is to use so-called case-based reasoning, where the principle is to develop a system which can document all possible problems or cases and the corresponding information from the transducers. Assuming that suitable information from analysis is available, this kind of approach should lead to quite a reliable result if there is time and resources to do the definition work. Since the features of condition monitoring signals can be rather complicated to analyse, and it is not always easy to know when a fault is present, different types of neural networks or statistical approaches have been used for this classification task. Basically the idea with the use of neural nets and other numerical methods is usually rather simple, i.e. let the net see a sufficient number of cases and it can then learn how the measured parameters are linked and consequently learn how faults can be recognised. Again, it is rather easy to say where the problem lies, i.e. how the system can be fed enough information from a set of transducers so that the whole range of interesting faults are covered in a remarkable set of running conditions. It is not the purpose of this introduction to try to cover the wide field of artificial intelligence (AI) and of knowledge based approaches to diagnosis of condition monitoring signals. Instead the idea is merely to show, using some examples, how solving one problem might lead to a range of other problems. There are so many approaches and none of them, although they work well in certain cases, are suitable for every kind of purpose, and that is why there is still room for new ideas and attempts in this demanding field of engineering and maintenance. The approach described in the following could be described as an attempt to combine a number of techniques referred to above in the most suitable way that would make the system easy to use, reliable, and wide in scope.

Principles of the approach: The system consists of a number of modules: data acquisition, signal processing, data handling, a fuzzy classifier and a neural net for diagnosis. Data acquisition is based on the use of an AD card, which can be configured to work with a number of sensors including, for example, vibration transducers. The system can also handle the on/off type of crisp information. Signal processing is mainly based on the use of traditional FFT (Fast Fourier Transform) together with ordinary statistical parameters. The novelty in signal processing and data handling lies in the use of regression analysis functions which make it possible to monitor a great number of different kinds of components, e.g. the tools in a machining process, without running into problems with available computer hard disk space. In the approach fuzzy logic, neural networks and case based reasoning are combined to build a system where the user can easily, through a graphical user interface, use and configure the system. The fuzzy classifier together with the neural network is organised in a hierarchical structure, which enables the easy configuration of the whole system. The approach is especially flexible in the sense that the total number of parameters the system can handle is not really limited in practice. In the hierarchical structure the individual sub-models are restricted to handling eight fuzzy inputs simultaneously (see Figure 1). The user can construct the whole diagnosis model through a graphical user interface. In practice, the most time consuming task is not the configuration of the system but the adjustment of the limits of the fuzzy classes, which again takes place through an easy to use graphical user interface with built-in editing features, such as copying. In the system the type of neural networks can be chosen from a list of choices based on the desired type of behaviour. In a normal case, parts of the hierarchical system are configured based on crisp information, and in these sub-models the fuzzy classifier does not have its normal function but is merely used as such to render the treatment of data similar in all cases. Similarly, the features of neural nets are not used in all of the sub-models and they can be substituted with arithmetic functions or expressions if there is no need for handling non-linear information or the behaviour is well known and can be easily defined otherwise. The system has been programmed using Visual Basic programming language in a Windows operating system environment and is based on the use of multiple windows [2]. The major advantages in the proposed approach are its flexibility of working with different types of machinery and the possibility to copy parts of the model (=sub-models) from one industrial plant to another where similar components are used.

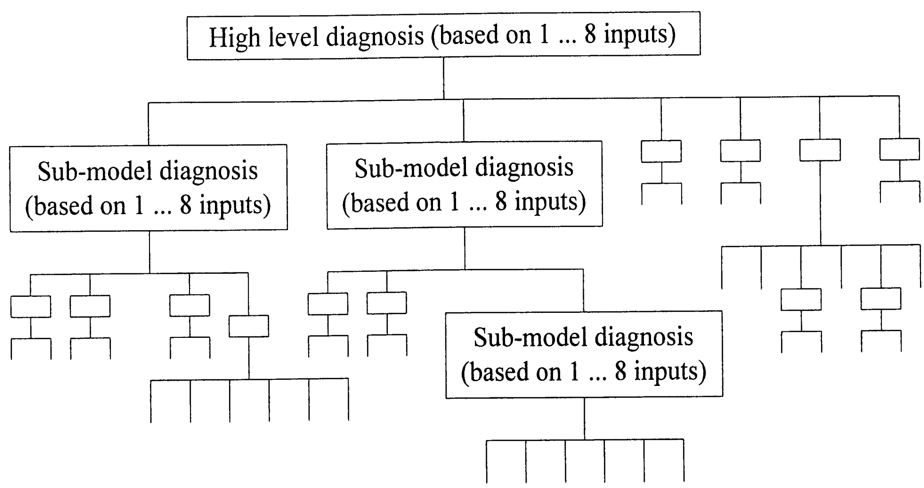


Figure 1. The structure of the hierarchical neuro-fuzzy system.

Data acquisition: Data acquisition is based on the use of an AD card, which can be configured to work with a number of sensors including, for example, vibration, sound, acoustic emission, pressure, current, voltage, power, speed of rotation and strain. The AD card is configured using a graphical user interface. The user is expected to define such parameters as the sampling rate, number of channels, type of windowing function, amplification/sensitivity, name of sensor, type (vibration, pressure, strain etc.) of sensor, units definition, type of averaging and number of averages. The idea is for the system to be capable of supporting a number of AD cards from a number of manufactures, although to date it has been configured to support only two models from different manufacturers.

Signal analysis: Signal processing is based on the use of traditional (spectrum, cepstrum) FFT (Fast Fourier Transform) together with statistical (root mean square, average, maximum, minimum, skewness and kurtosis) parameters and also the on/off type of information [3]. The user has a choice of these parameters and can assign from one to eight of them to the system at the lowest defined classification level in a suitable optimal mix. The novelty in signal processing and data handling lies in the use of regression analysis functions which make it possible to monitor a great number of different kinds of components, e.g. the tools in a machining process [4]. When using regression analysis techniques, only the regression analysis coefficients are stored in the database. This markedly reduces the amount of data to be stored, especially if the system is used to monitor such a complex target that this might become a problem. The available regression functions are first-, second- and third-order polynomials, and a logarithmic function developed to indicate or follow the progress of wear and thus to be suitable for prognosis of the remaining lifetime of the machine component [5].

Database: All data used by the system is stored in an Access database. The neuro-fuzzy diagnosis part of the database consists of five tables, as shown in Table I along with the function of each. The database, although very easily described, is actually the key element of the whole system. All communication internally is through this database, i.e. the definition of the structure of the system is there, as is all the measured data saved there, taken from there for diagnosis, and the results of the diagnosis. The consequence of the above is naturally that the database size can with time become immense, although both signal analysis techniques and regression analysis are used to reduce the amount of data.

Table I. Tables of the database of the neuro-fuzzy diagnosis module.

Table	Description
Hierarchy	Describes the hierarchy of the system
Measurement-Data-Fuzzy	Gives the measurement results
Measurement-Conditions	Describes the measurement conditions
Text	Texts that the program uses for communication in different languages
Diagnosis	The results of fuzzy classification

Fuzzy classifier: In the approach, fuzzy logic, neural nets and case-based reasoning are combined to build a system which the user can easily configure through a graphical user interface. The fuzzy classifier, together with the neural network, is organised in a hierarchical structure which enables easy configuration of the whole system. The fuzzy classifier acts as a pre-processor to the neural net [6]. The approach is especially flexible in the sense that the total number of parameters the system can handle is not limited in practice (i.e. with the limitations given below there can be a total of 4681 lines in the hierarchy table). In the hierarchical structure the individual sub-models are restricted to handling eight fuzzy inputs simultaneously. The number of hierarchy levels in the system is limited to four. The user can construct the model through a graphical user interface. In practice the most time-consuming task is not configuring the system, but adjusting the limits of the fuzzy classes, which again takes place through an easy-to-use graphical user interface show in Figure 2.

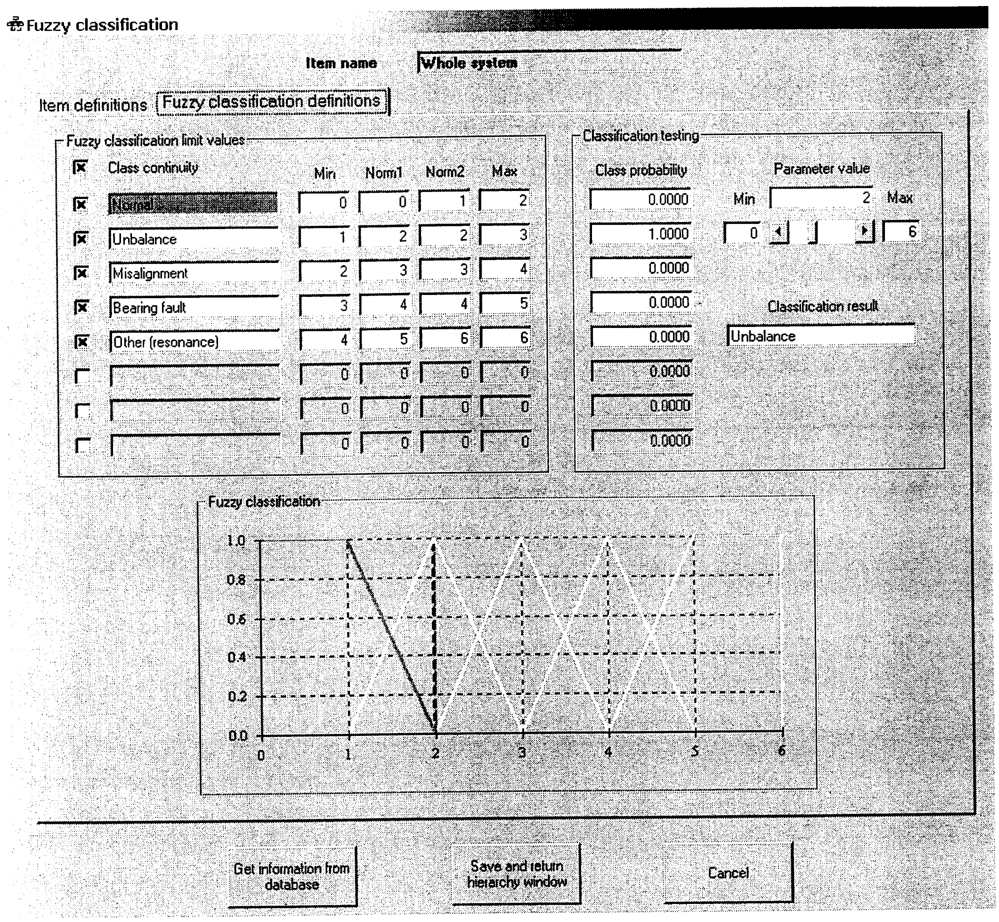


Figure 2. User interface of the fuzzy classification module.

The number of fuzzy classes can vary from one to eight. The classes must be continuous but the user can turn off checking for continuity while changes are being made, and then turn it on again. In the example shown in Figure 2, the number of fuzzy classes is five. For all of these inputs the user is expected to give four values which define the limits of that specific fuzzy class. However, two of these values actually define two values for the next class, i.e. only two additional values are needed for definition of the next class. In practice this means that two plus n times two values are needed in the definition of n classes. The user interface shows the fuzzy classes graphically. It also gives the user the opportunity to test what happens when the parameter being classified gets a certain value, i.e. into which class it falls. It should be noted that all the measured parameters go through the fuzzy interface, and similarly all the results of fuzzy classification of lower sub-models pass through this interface. However, not all the classes need to be fuzzy, i.e. it is possible to define sharp limits between the classes. Sharp limits are often used when the results of fuzzy classification are passed further on, or when on/off type information is being handled. The example shown in Figure 2 is from a higher level, i.e. it is not the lowest level that handles parameters from condition monitoring signals or process status information. The example shows the fuzzy interface at the level where the diagnosis system distinguishes between a number of typical faults that can be diagnosed with the use of vibration measurements.

Diagnosis: In the system, the type of neural networks can be chosen from a list of choices based on the desired type of behaviour. In a normal case parts of the hierarchical system are configured based on crisp information, and in these sub-models the fuzzy classifier does not have its normal function but is merely used as such to make the treatment of data similar in all cases. Similarly, the features of neural nets are not used in all of the sub-models, and they can be substituted with arithmetic functions if there is no need to handle non-linear information. If the behaviour is well known, i.e. it is implicitly defined what combination of results a fuzzy sub-model means, this information can be defined into the system through the interface shown in Figure 3. When the updating routines of the system are started, the hierarchy is first optimised during which all unnecessary nodes of the hierarchy tree are deleted. After optimisation the system goes through all the nodes and levels, starting at the bottom. If classification between the levels is based on neural nets or other algorithms, the whole classification process proceeds automatically. However, if the user has chosen to specify that a certain combination in a sub-model should be translated or classified to correspond to a specific situation, i.e. to a certain number, this combination may not yet have been defined. Should this be the case, the system will stop and ask the user to make this specification. In the case of neural nets there is some variation depending on the type of nets used. In the case of a traditional feed forward network, it is assumed that the user will train the sub-model first so that it can handle all possible situations [7]. In the case of self-organised maps it is possible to let the system organise itself, so that after a learning period it can handle various situations. The aim is especially to configure a specific version of the QSOM routine [8] so it can be used as a self-organising map. For each of the sub-models, the system shows on the interface the corresponding interpretation of that model using a colour code. It also shows the result both as a number and as plain text if the cursor is moved to that point on the interface (see Figure 4). The idea is that when the system is running continuously, the user can easily identify where the indication of a fault or something peculiar appears in the system. More specifically, the system shows the item the user is looking at, and gives information about the fault. Naturally all this information has had to be defined for the system, and if the number of connected channels is high this might be quite a task. However, to make the system definition more effective in the case of complicated systems, it is possible to copy information from one parameter to another.

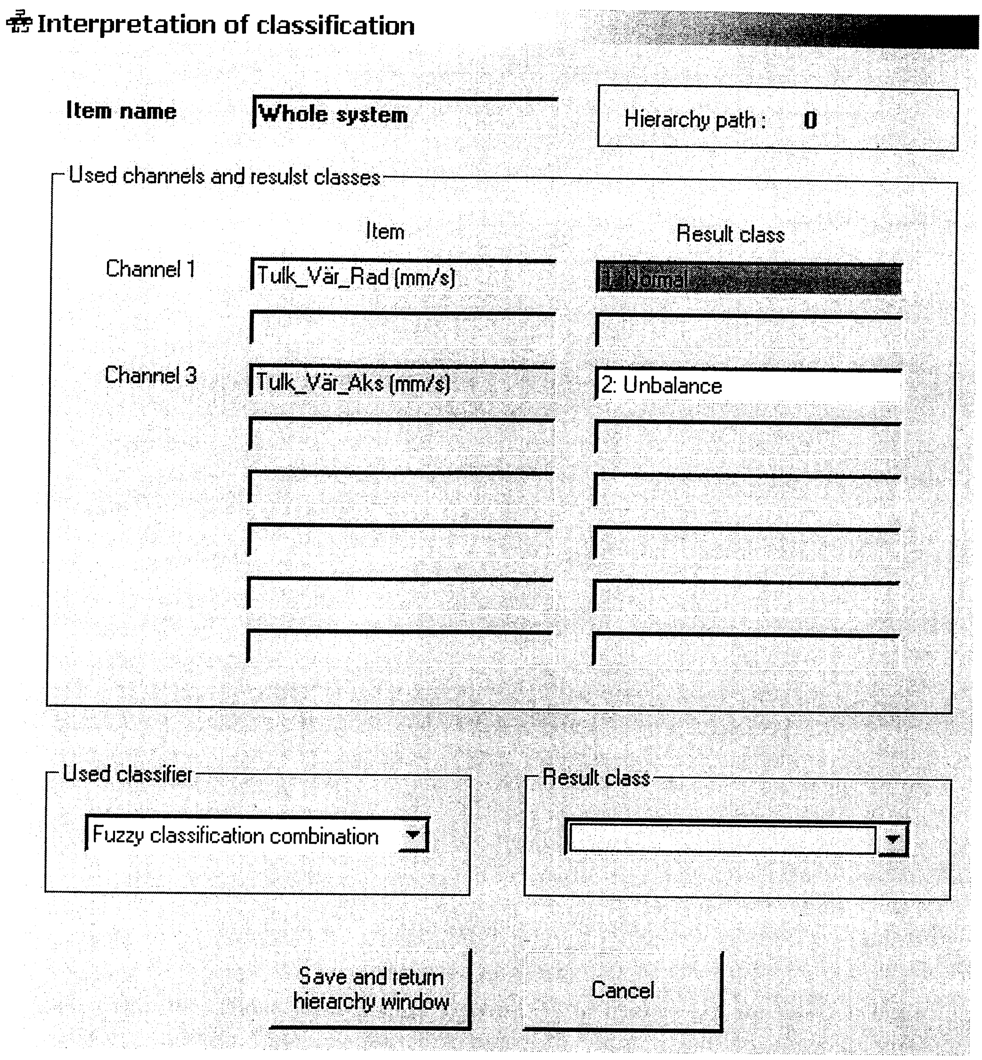


Figure 3. The interpretation of fuzzy classification.

The copy and paste technique is very practical and saves a lot of time if the machinery to be monitored has, for example, a number of rolling bearings that are monitored using a number of acceleration sensors. In principle, all of these are monitored using basically the same set-up and fuzzy limits at the start, so it is easy to copy the definition of bearing monitoring for all of these bearings. In practice, the way these bearings behave may vary, which affects what sort of parameters should be used and what the exact fuzzy limits are.

However, it is a lot easier to do a little fine tuning than to define the same thing a number of times from scratch.

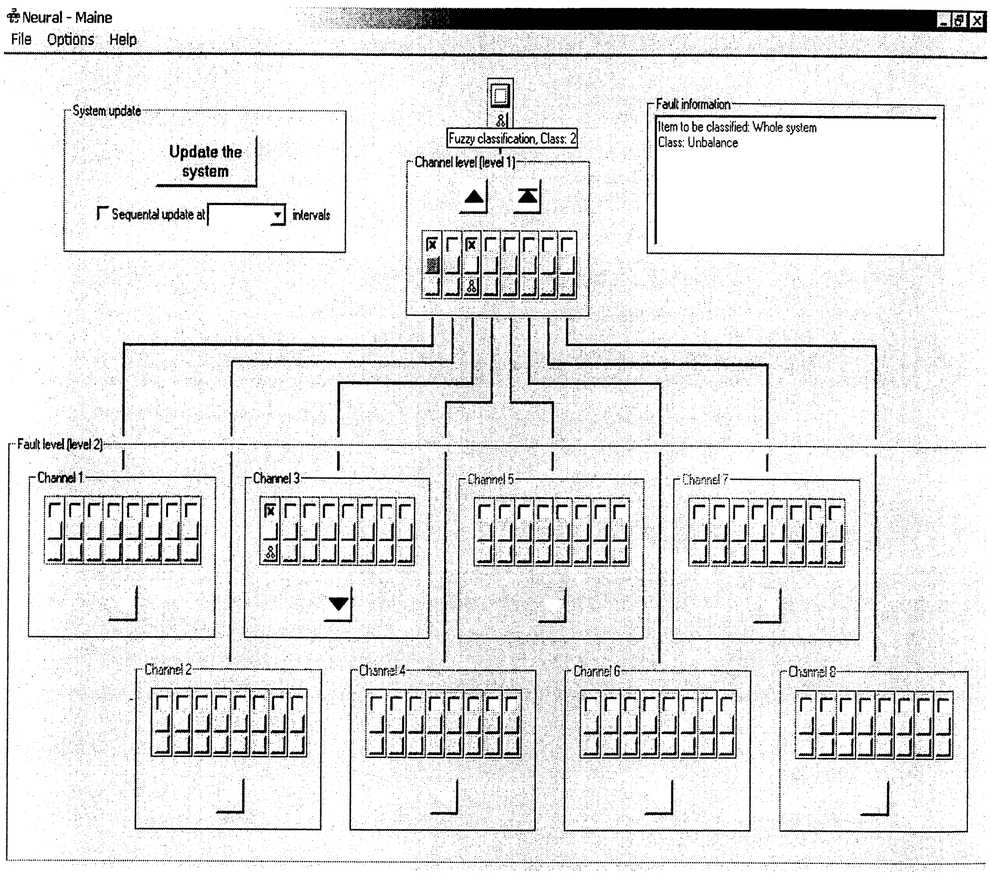


Figure 4. User interface of the hierarchical neuro-fuzzy prognosis system.

The time it takes for the system to go through all the sub-models with all parameters naturally depends on the number of parameters defined, and on the power of the computer. With a typical Pentium-type PC it takes only a few seconds if only a few channels are connected, or several minutes if a number of sub-models are connected. The user interface shows how the system is progressing. Naturally, if these times are compared with off-line monitoring lasting 2 weeks they are ridiculously short, but to a hurried user used to quick responses with a PC, they may feel a lot longer. Especially during the training and definition phase it may be frustrating to wait for the system to update, but it is possible to concentrate only on sub-models of interest by clicking off those parts that are not of interest. Even though a channel is turned off, the definition for that channel will be held in the database unless purposely altered or deleted. Because the structure of the whole system is large it is impossible to show it all at once; consequently the user can move around in the model using the arrows shown in Figure 4. The user interface shows the level and channel in question.

Further development: The system is undergoing a testing phase in e.g. the food and manufacturing industries, power production and the monitoring of conveyors and lifts. The neural network part of the system cannot be regarded as complete due to the number of approaches that could be installed. Connections to different kinds of measuring equipment could widen the scope of the system. In many cases it would be logical to use the system on the World Wide Web as this could lower the cost of ownership, and offer ease of upgrading and a larger number of sub-models in the library [9]. Naturally the most important thing is to take into account feedback from industrial users, especially concerning any bugs in the system, and their views on how to make the system easier, faster and more logical to use.

Conclusion: An easy to configure hierarchical neuro-fuzzy system has been defined for the configuration of a prognosis system for condition monitoring of machinery. The system consists of a number of modules: data acquisition, signal processing, data handling, a fuzzy classifier and neural networks for diagnosis. Data acquisition is based on the use of an AD card, which can be configured to work with a number of sensors including, for example, vibration, sound and pressure. The system can also handle the on/off type of crisp information. Signal processing is based on the use of traditional FFT (Fast Fourier Transform) together with statistical time domain parameters. The novelty in signal processing and data handling lies in the use of regression analysis functions which make it possible to monitor a great number of different kinds of components, like the tools in a machining process. The fuzzy classifier together with the neural network is organised in a hierarchical structure, which enables easy configuration of the whole system. The approach is especially flexible in the sense that the total number of parameters the system can handle is not limited in practice. In the hierarchical structure the individual sub-models are restricted to handling eight fuzzy inputs simultaneously. The user can construct the model through a graphical user interface. In practice the most time-consuming task is not the configuration of the system but the adjustment of the limits of the fuzzy classes, which again takes place through an easy-to-use graphical user interface. In the system the type of neural networks can be chosen from a list of choices based on the desired type of behaviour. In a normal case parts of the hierarchical system are configured based on crisp information, and in these sub-models the fuzzy classifier does not have its normal function but is merely used as such to render the treatment of data similar in all cases. Similarly, the features of neural nets are not used in all of the sub-models and they can be substituted with arithmetic functions if there is no need for handling non-linear information, or if the behaviour is well known and can be easily defined otherwise. The major advantages in the proposed approach are its flexibility of working with different types of machinery and the possibility to copy parts of the model (=sub-models) from one industrial plant to another where similar components are used.

Acknowledgements: The work described in this paper was carried out as part of the Eureka Project EU1250 (Neural-Maine) On-Line Multi-Sensor Diagnostic Analysis for Maintenance using Neural Networks. The contribution of the project team, T. Arokivi, J. Halme, A. Helle, T. Runtti, J. Virtanen and K. Vähä-Pietilä, is gratefully acknowledged. The support and guidance of the industrial project partners YIT Service Ltd, NeuroVision Ltd, Helsinki Water, Gustav Paulig Ltd, Finnish Post Ltd, Hartwall Plc, Fortum Power and Heat, Kone Elevators, Raisio Group Ltd and Fazer Bakeries Ltd Oululainen Bakery is also gratefully acknowledged.

References:

[1] Rao B.K.N., Handbook of Condition Monitoring, Elsevier Science Ltd., Oxford, 1996, ISBN 1-85617-234-1

[2] Schneiderman B., Designing the User Interface, Strategies for Effective Human-Computer Interaction, Addison-Wesley Longman Inc., 1998, ISBN 0-201-69497-2

[3] DeFatta J., Lucas J.G. and Hodgkiss W.S., Digital Signal Processing, A System Desing Approach, John Wiley & Sons, Inc., New York, 1988, ISBN 0-471-63765-3

[4] Jantunen E. and Jokinen H., Reduction of Data Needed in an Expert System for Condition Monitoring of FMS Using Regression Analysis Techniques, Proceedings of Comadem '96 1996. University of Sheffield, Sheffield

[5] Milton J.S. and Arnold J.C., Probability and Statistic in the Engineering and Computing Sciences, McGraw-Hill, 1986, New York, ISBN 0-07-042351-2

[6] Rao V.B. and Rao H.V., C++ Neural Networks and Fuzzy Logic, Management Information Source Inc., New York, 1993, ISBN 1-55828-298-x

[7] Schalkoff R., Pattern Recognition, Statistical, Structural and Neural Approaches, John Wiley & Sons, Inc., 1992, ISBN 0-471-52974-5

[8] Lumme V.E., Practical Neural Net Applications in Condition Monitoring, Proceedings of Comadem '97, Vol. 2 1997. Espoo, Finland, ISBN 951-38-4563-X, pp. 162- 171

[9] Rahmel D., Building Web Database Applications with Visual Studio 6, Osborne/McGraw-Hill, Berkeley, 2000, ISBN 0-07-212094-0

Published by

Series title, number and report code of publication

VTT Publications 590
VTT–PUBS–590

Author(s)
Jantunen, Erkki

Title

Indirect multisignal monitoring and diagnosis of drill wear

Abstract

A machine tool utilisation rate can be improved by an advanced condition monitoring system using modern sensor and signal processing techniques. A drilling test and analysis program for indirect tool wear measurement forms the basis of this thesis. For monitoring the drill wear a number of monitoring methods such as vibration, acoustic emission, sound, spindle power and axial force were tested. The signals were analysed in the time domain using statistical methods such as root mean square (rms) value and maximum. The signals were further analysed using Fast Fourier Transform (FFT) to determine their frequency contents. The effectiveness of the best sensors and analysis methods for predicting the remaining lifetime of a tool in use has been defined. The results show that vibration, sound and acoustic emission measurements are more reliable for tool wear monitoring than the most commonly used measurements of power consumption, current and force. The relationships between analysed signals and tool wear form a basis for the diagnosis system. Higher order polynomial regression functions with a limited number of terms have been developed and used to mimic drill wear development and monitoring parameters that follow this trend. Regression analysis solves the problem of how to save measuring data for a number of tools so as to follow the trend of the measuring signal; it also makes it possible to give a prognosis of the remaining lifetime of the drill. A simplified dynamic model has been developed to gain a better understanding of why certain monitoring methods work better than others. The simulation model also serves the testing of the developed automatic diagnostic method, which is based on the use of simplified fuzzy logic. The simplified fuzzy approach makes it possible to combine a number of measuring parameters and thus improves the reliability of diagnosis. In order to facilitate the handling of varying drilling conditions and work piece materials, the use of neural networks has been introduced in the developed approach. The scientific contribution of the thesis can be summarised as the development of an automatically adaptive diagnostic tool for drill wear detection. The new approach is based on the use of simplified fuzzy logic and higher order polynomial regression analysis, and it relies on monitoring methods that have been tested in this thesis. The diagnosis program does not require a lot of memory or processing power and consequently is capable of handling a great number of tools in a machining centre.

Keywords
drill wear, condition monitoring, signal analysis, polynomial regression analysis, fuzzy logic, diagnosis

Activity unit
VTT Industrial Systems, Metallimiehenkuja 6, P.O.Box 1702, FI–02044 VTT, Finland

ISBN			Project number
951–38–6692–0 (soft back ed.) 951–38–6693–9 (URL:http://www.vtt.fi/inf/pdf/)			
Date December 2005	Language English	Pages 80 p. + app. 110 p.	Price D
Name of project BRST-CT98-5429, EUREKA MAINE EU 744, EUREKA Project EU 1250, MindMan		Commissioned by EU, Tekes, Finnish industry, VTT	
Series title and ISSN VTT Publications 1235–0621 (soft back ed.) 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)		Sold by VTT Information Service P.O.Box 1000, FI–02044 VTT, Finland Phone internat. +358 20 722 4404 Fax +358 20 722 4374	

VTT PUBLICATIONS 590 Indirect multisignal monitoring and diagnosis of drill wear

A machine tool utilisation rate can be improved by an advanced condition monitoring system using modern sensor and signal-processing techniques. A drilling test and analysis program for indirect tool wear measurement forms the basis of this thesis. The results show that vibration, sound and acoustic emission measurements are more reliable for tool wear monitoring than the most commonly used measurements of power consumption, current and force. The scientific contribution of the thesis can be summarised as the development of an automatically adaptive diagnostic tool for drill wear detection. The new approach is based on the use of simplified fuzzy logic and higher order polynomial regression analysis, and it relies on monitoring methods that have been tested in this thesis. The diagnosis program does not require a lot of memory or processing power and consequently is capable of handling a great number of tools in a machining centre.

Tätä julkaisua myy	Denna publikation säljs av	This publication is available from
VTT TIETOPALVELU PL 1000 02044 VTT Puh. 020 722 4404 Faksi 020 722 4374	VTT INFORMATIONSTJÄNST PB 1000 02044 VTT Tel. 020 722 4404 Fax 020 722 4374	VTT INFORMATION SERVICE P.O.Box 1000 FI–02044 VTT, Finland Phone internat. +358 20 722 4404 Fax +358 20 722 4374

ISBN 951- 38- 6692- 0 (soft back ed.)
ISSN 1235- 0621 (soft back ed.)
ISBN 951- 38- 6693- 9 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455- 0849 (URL: http://www.vtt.fi/inf/pdf/)

Erkki Jantunen