CHAPTER 3

3. Tool Condition Monitoring

3.1 Introduction

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

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

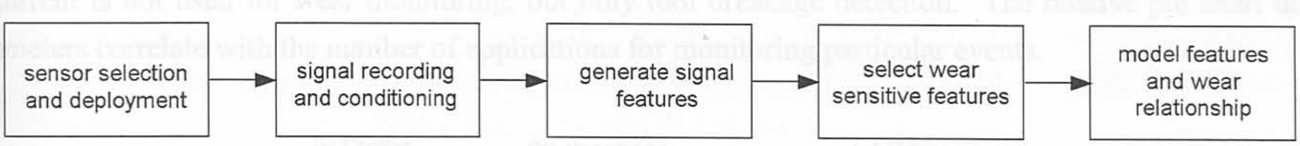

Figure 3.1: TCM steps

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

3.2 Sensors for general process monitoring

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

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

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

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

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

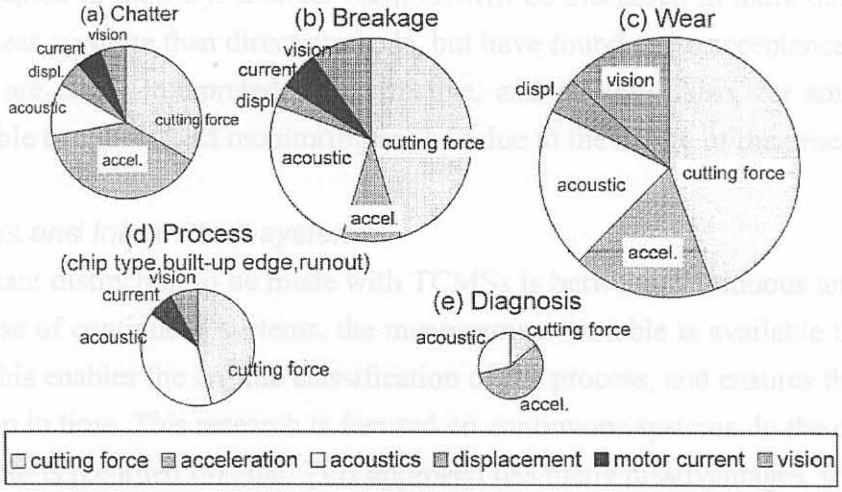

Figure 3.2: Sensor application in manufacturing process monitoring [97]

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

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

3.3 Sensor-based tool wear monitoring

3.3.1 Introduction

A. Direct and Indirect systems

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

B. Continuous and Intermittent systems

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

C. Sensor requirements for tool wear monitoring

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

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

3.3.2 Force-based monitoring

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

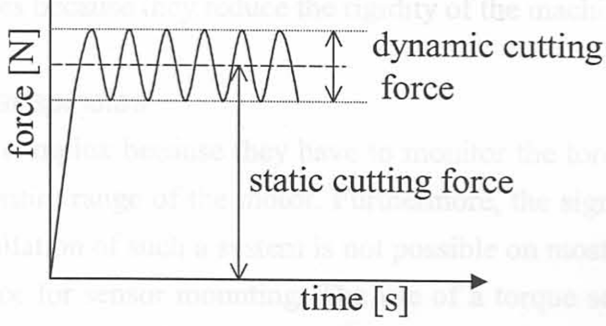

Figure 3.3: Static and dynamic forces

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

A. Direct measurement dynamometers

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

B. Plates and rings

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

C. Pins, extension sensors

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

D. Measurement of displacement

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

E. Force-measuring bearings

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

F. Force and torque at spindles

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

3.3.3 Measurement of motor current

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

3.3.4 Acceleration

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

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

3.3.5 Acoustic emission

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

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

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

Araujo *et al.* [115] investigated the sliding friction as a possible source of AE during metal cutting. The AErms in different frequency ranges was collected for different widths of cut and also with the tool rubbing against the workpiece (without cutting). It was found that the level of AE remains almost constant for all width of cut conditions, and hence it can be concluded that the main mechanism for AE during metal cutting is the sliding friction between the tool and workpiece. Consequently, an increase or decrease of AE can be expected with tool wear depending on the effect on the sliding friction due to tool wear. It is also believed that the cutting temperatures will affect the level of AE due to thermal expansion effects. The effect of plastic deformation with other materials is currently under investigation. Chio and Liang [116] investigated AE with tool wear and chatter effects in turning. A model is presented that can predict the chatter AErms amplitude with certain severities of flank wear. Good correlation was found between the model and experimental results. Kim *et al.* [117] reports on the use of AE to monitor the tool life during a gear shaping process. The AErms is collected and used in a software program to predict the remaining tool life.

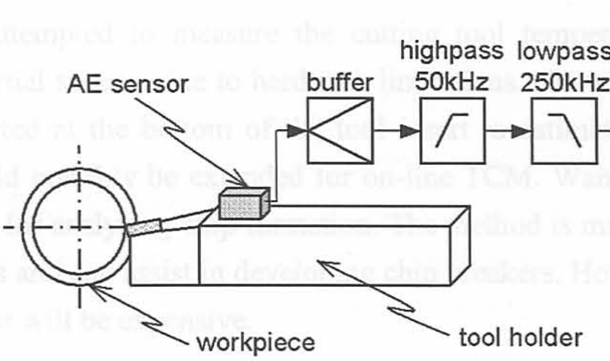

Figure 3.4: Steps for collecting AE during turning

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

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

3.3.6 Temperature monitoring

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

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

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

3.3.7 Ultrasonic methods

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

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

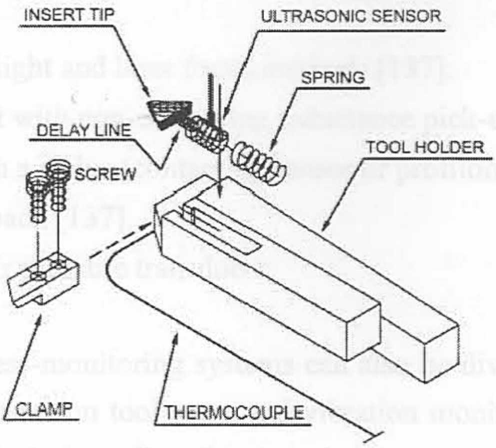

Figure 3.5: Ultrasonic TCM approach [97]

3.3.8 Vision systems

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

3.3.9 Surface roughness monitoring

A. Introduction

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

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

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

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

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

B. Surface roughness analysis and tool wear

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

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

C. Vibration monitoring and surface roughness analysis

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

One problem is that chatter between the tool and workpiece causes large vibrations that cannot be superimposed on the surface roughness. Another problem is that loose metal parts and other external factors easily distort signals from the sensors. However, the method has been successfully implemented in dry turning with ferrous metals by Jang *et al.* [139]. In this case the kinematics of the machine tool are taken into account to estimate the roughness from vibration signals collected during cutting. They suggested that further research be done in this field.

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

3.3.10 Other methods

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

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

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

3.4 Decision making in sensor-assisted TCM

3.4.1 Introduction

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

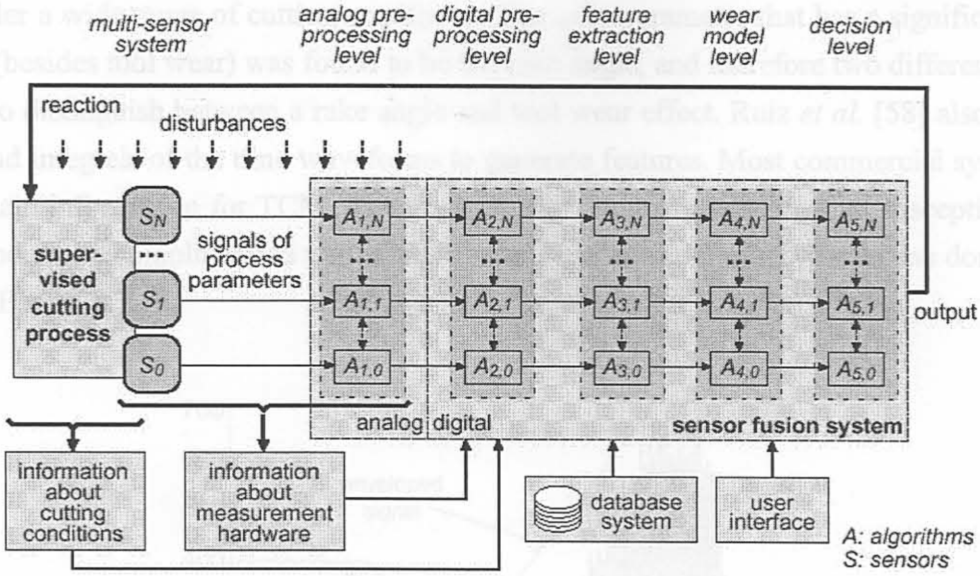

Figure 3.6: Sensor fusion architecture for TCM by Sick [6]

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

3.4.2 Feature extraction

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

A. Time domain

Features extracted from the time domain are mostly basic values such as the signal average, mean, or a root mean square (rms) value. Other techniques include the shape of enveloping signals, threshold crossings, ratios between time domain signals, peak values and polynomial approximations of time domain signals. It has been found that some of the time domain features are very useful and they are easy to implement. Bayramoglu and Dungel [150] investigated the use of several different force ratios (calculated from the static cutting forces). It was found that certain force ratios can be used to monitor tool wear under a wide range of cutting conditions. The only parameter that has a significant influence on the ratios (besides tool wear) was found to be the rake angle, and therefore two different force ratios are required to distinguish between a rake angle and tool wear effect. Ruiz *et al.* [58] also report using derivatives and integrals of the time waveforms to generate features. Most commercial systems rely on the time domain information for TCM. The time domain features are somewhat susceptible to disturbances and should be complimented with features from another domain. Typical time domain features are shown in Figure 3.7.

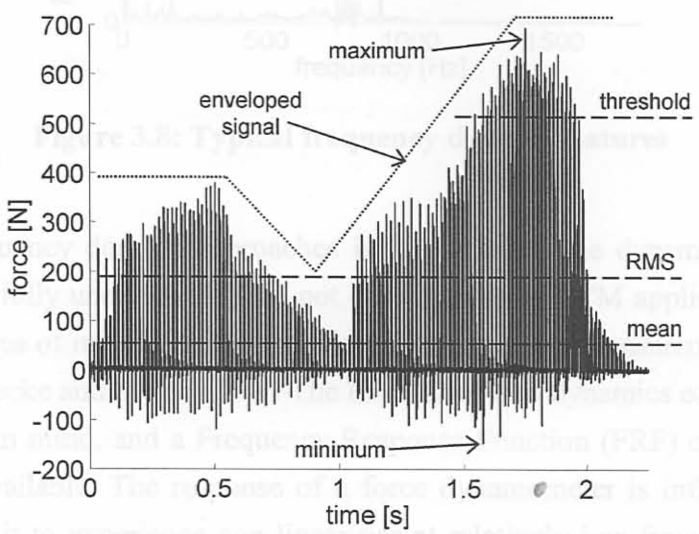

Figure 3.7: Typical time domain features

B. Frequency domain

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

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

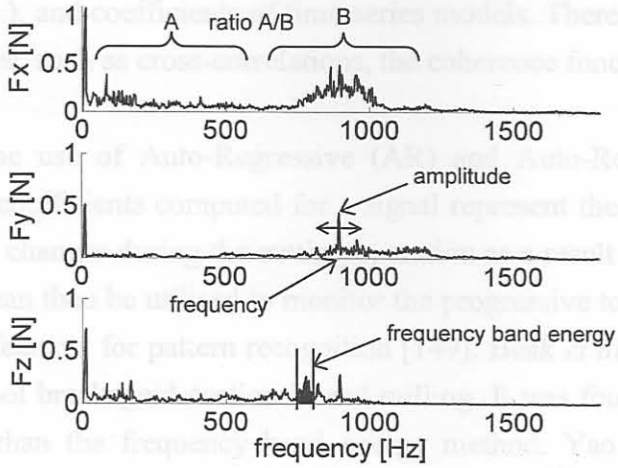

Figure 3.8: Typical frequency domain features

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

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

C. Statistical processing

In the case of statistical features, signals are assumed to have a probabilistic distribution. Hence, the signal is regarded as a random process. Generally, machining processes are non-stationary but are assumed stationary for the short periods during which these features are calculated. Several statistical features have been investigated for TCM and can be applied to machining operations. The main features are those that describe the probability distribution of a random process (variance, standard deviation, skewness, kurtosis *etc.*), and coefficients of time series models. There are also various other miscellaneous statistical features, such as cross-correlations, the coherence function and harmonic mean.

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

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

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

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

D. Time-frequency domain

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

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

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

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

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

D. Other

There are also a few other techniques that cannot be categorised as either time or frequency domain. One interesting technique is the use of entropy functions. An example is described by Fu *et al.* [169], where the entropy of the frequency spectrum is calculated. The result is one value that is used for pattern classification of different faults that may occur during machining. It is stated that the advantage is the entropy function's insensitivity towards new geometries of cutting. This might however be more related to the characteristic of the FFT than the entropy function! Chungchoo and Saini [170] use the total energy and total entropy of force signals in the frequency domain for TCM. It is stated that the entropy is more related to the distribution of energy in the spectrum, and is relatively insensitive to changing machining parameters. However, the total energy was found to be the only parameter sensitive to progressive tool wear. The entropy is often also used as a time domain feature, and also represents the energy contained within the wave, thus more or less the same as the signal rms.

3.4.3 Feature selection

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

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

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

Ruiz *et al.* [58] proposed the use of a discrimination power for feature selection in a TCM application. The method is similar to that of the SOF. An automated version is proposed that also checks for linear correlation between features. It is difficult to assess the success rate of the automated procedure because the experiments / simulations are not described in enough detail. Quan *et al.* [176] reported the use of the correlation coefficient to assist in feature selection. Lee *et al.* [177] describe the use of the ANOVA to determine the best force ratio for TCM statistically. Several ratios between the three main cutting forces are computed and the influence of controllable parameters (*e.g.* machining conditions) on these ratios are investigated by means of the ANOVA technique. The use of ANOVA as well as the correlation coefficient was also reported by Scheffer [164]. Du [33] describes the use of a blackboard system, which is a knowledge-based approach for feature selection and decision-making. An advantage is the fact that a physical interpretation of feature can be linked to phenomena in the machining operation. The method is also flexible, but suffers from the disadvantage of requiring a large amount of data and expertise to establish the knowledge-based rules. Some of these techniques can also be automated for a faster implementation. In the opinion of the author, engineering judgement plays a vital role in feature selection for TCM. Some of the techniques and their role in feature selection will become apparent in the chapters that follow.

3.4.4 Wear model / Decision making

A. Time domain signature

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

B. Trending, threshold

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

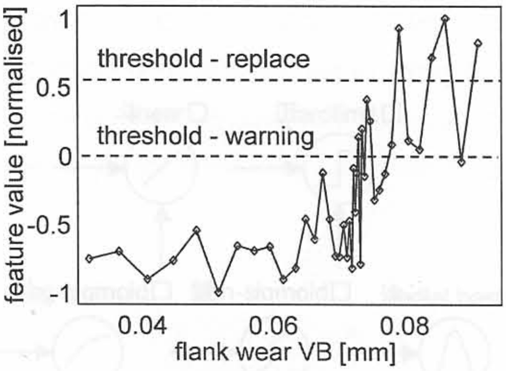

Figure 3.9: Thresholds

C. Neural networks

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

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

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

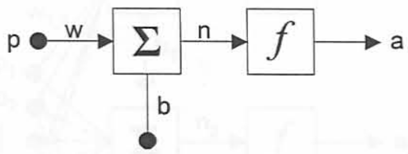

Figure 3.10: Single neuron with bias

The value for a is determined by:

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

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

Figure 3.11: Common activation functions

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

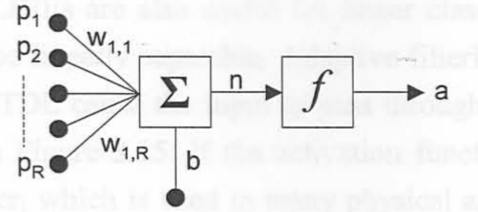

Figure 3.12: Vector input to a neuron [196]

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

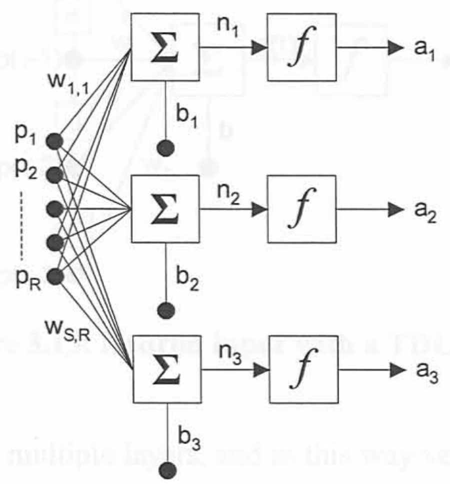

Figure 3.13: A layer of neurons [196]

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

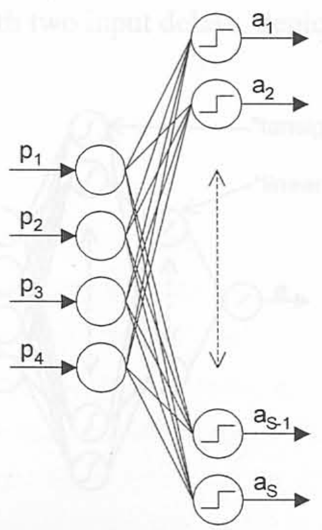

Figure 3.14: A layer of neurons with "hardlimit" activation functions

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

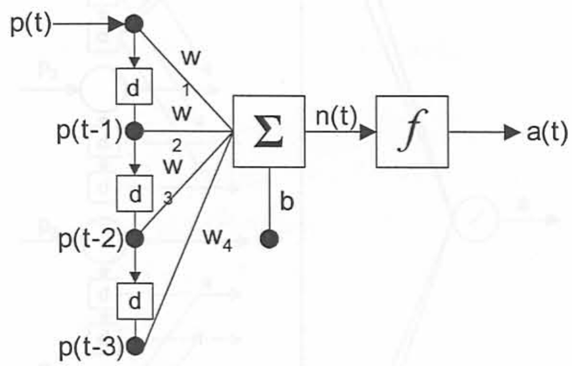

Figure 3.15: Neuron input with a TDL [196]

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

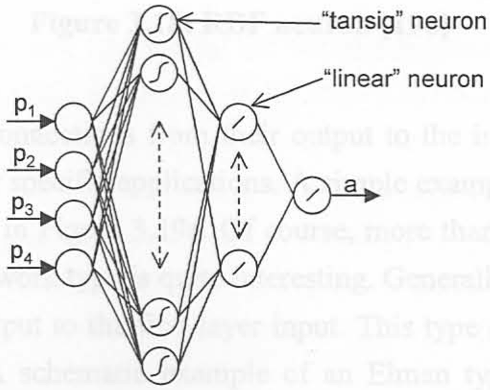

Figure 3.16: Multilayer FF network

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

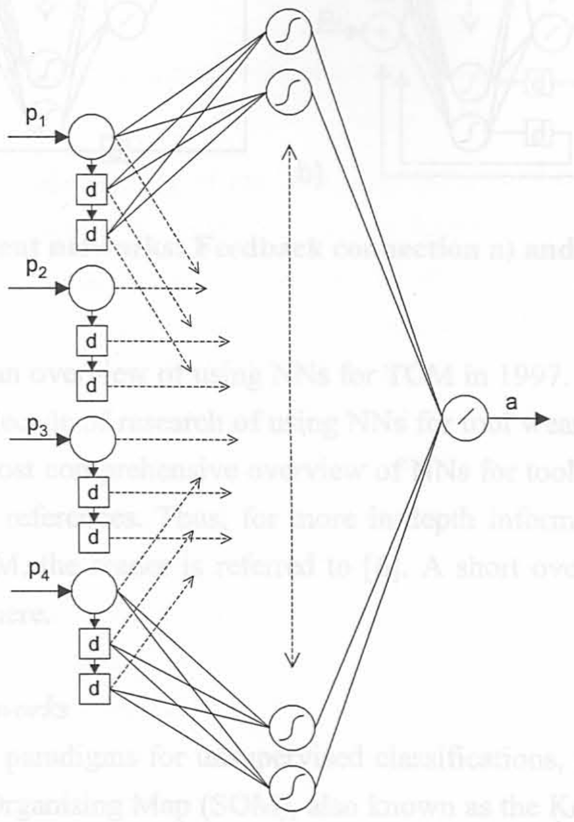

Figure 3.17: Multilayer FF network with TDLs

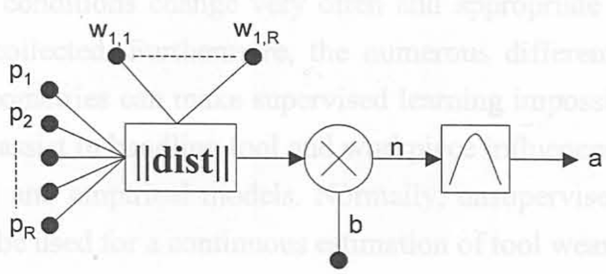

Figure 3.18: RBF neuron [196]

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

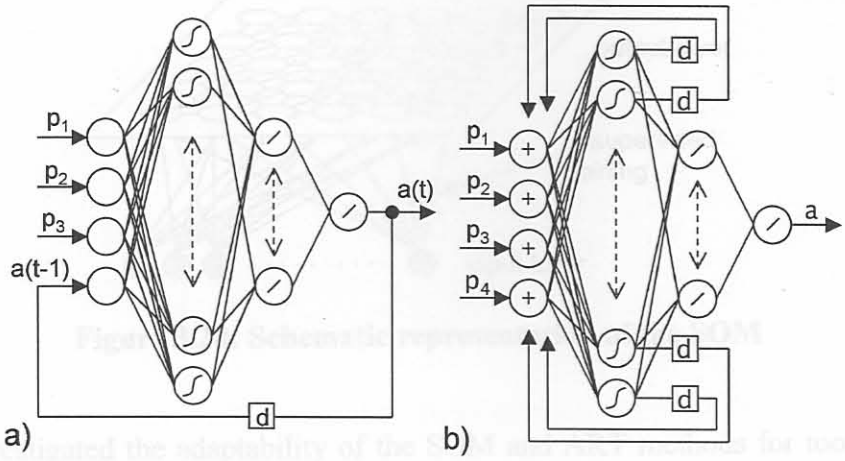

Figure 3.19: Recurrent networks: Feedback connection a) and Elman network b)

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

C.1 Unsupervised networks

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

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

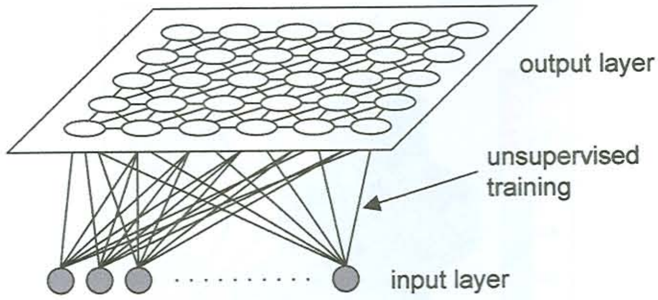

Figure 3.20: Schematic representation of the SOM

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

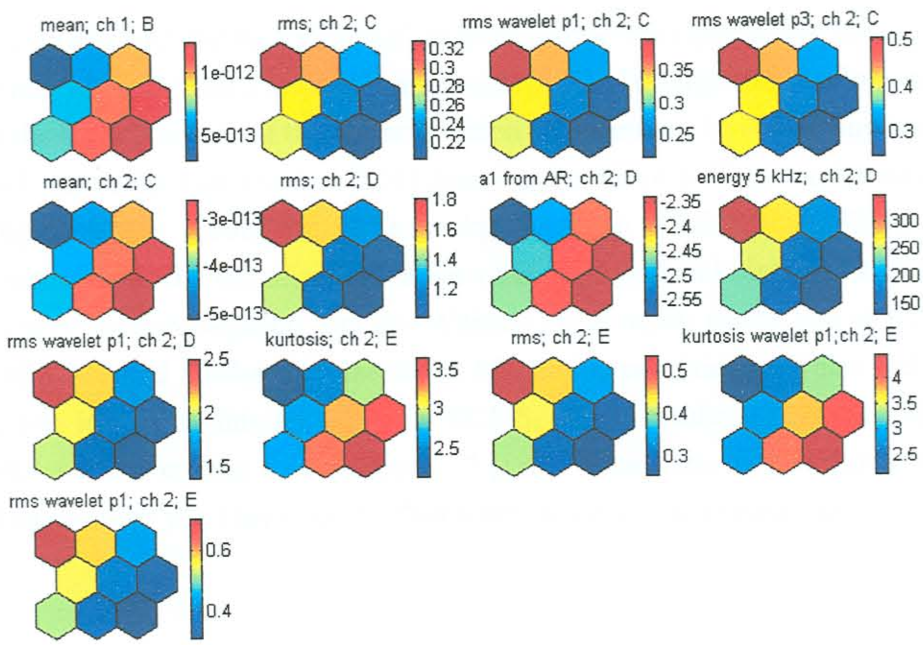

Figure 3.21: Data mining with SOM [165]

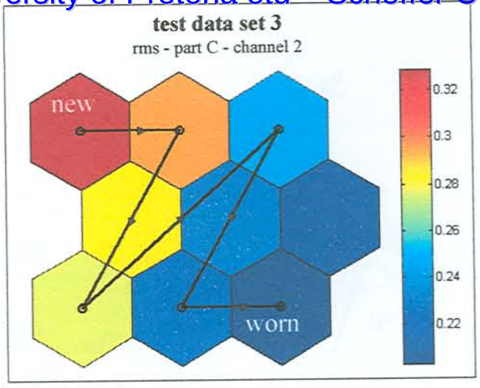

Figure 3.22: Wear classification with SOM [165]

C.2 Supervised networks

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

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

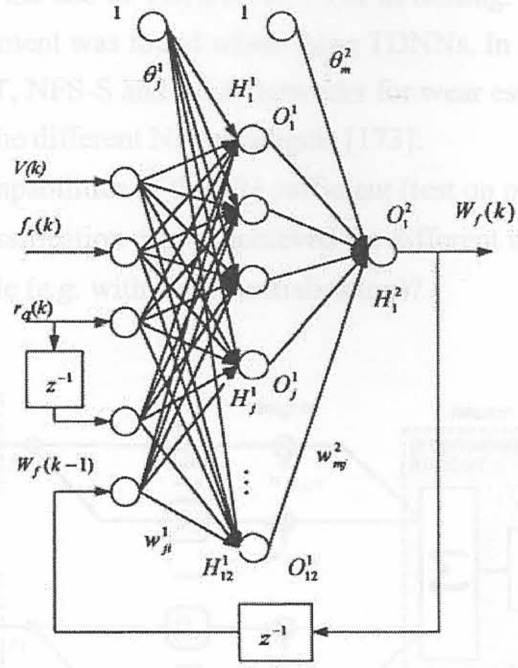

Figure 3.23: NN formulation with TDLs and feedback [189]

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

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

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

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

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

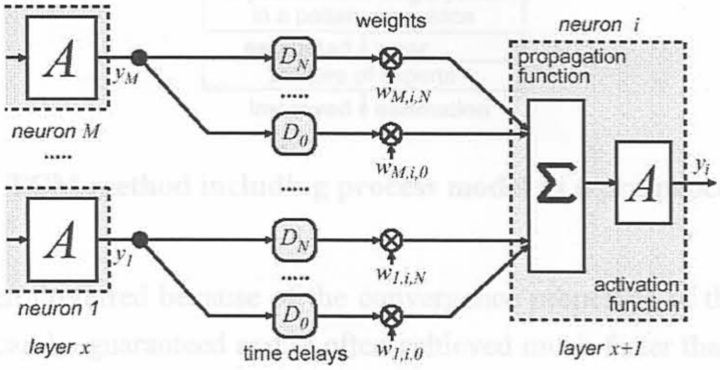

Figure 3.24: TDLs for a TCM application [171]

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

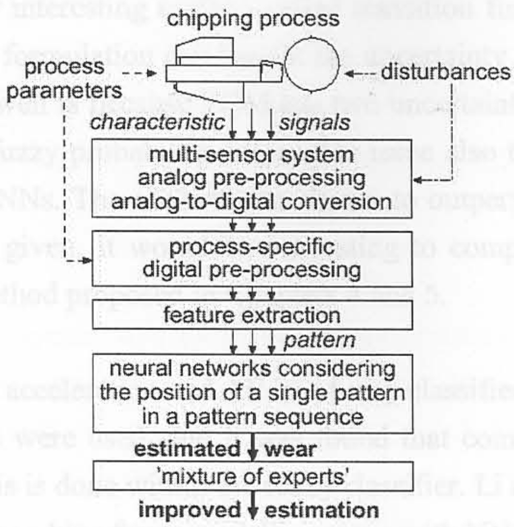

Figure 3.25: TCM method including process model as a pre-processing step [6]

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

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

D. Fuzzy logic

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

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

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

E. Other methods

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

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

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

3.5 Conclusion

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