

## **AN ADAPTIVE OBSERVER FOR ON-LINE TOOL WEAR ESTIMATION IN TURNING, PART I: THEORY**

KOUROSH DANAI AND A. GALIP ULSOY

*Department of Mechanical Engineering and Applied Mechanics, University of Michigan, Ann Arbor, Michigan 48109-2125, U.S.A.*

*(Received March 1986, and in revised form July 1986)*

On-line sensing of tool wear has been a long-standing goal of the manufacturing engineering community. In the absence of any reliable on-line tool wear sensors, a new model-based approach for tool wear estimation has been proposed. This approach is an adaptive observer, based on force measurement, which uses both parameter and state estimation techniques. The design of the adaptive observer is based upon a dynamic state model of tool wear in turning. This paper (Part I) presents the model, and explains its use as the basis for the adaptive observer design. This model uses flank wear and crater wear as state variables, feed as the input, and the cutting force as the output. The suitability of the model as the basis for adaptive observation is also verified. The implementation of the adaptive observer requires the design of a state observer and a parameter estimator. To obtain the model parameters for tuning the adaptive observer procedures for linearisation of the non-linear model are specified. The implementation of the adaptive observer in turning and experimental results are presented in a companion paper (Part II).

### **1. INTRODUCTION**

The full automation of machine tools has been a goal of manufacturing engineering research for decades [1]. This economically important problem requires reliable techniques for on-line sensing of tool breakage [2, 3]. Unfortunately, despite years of research in this area, a reliable on-line tool wear measurement technique does not exist [4]. The purpose of this paper is to propose a new approach, based on implementing an adaptive observer, for on-line tool wear estimation.

#### **1.1. MOTIVATION**

Since the advent of numerically controlled (NC) machine tools, which provided the physical means to control the cutting process, the goal of complete automation of machine tools has seemed much more promising [5]. The presence of computer numerically controlled (CNC) machine tools has made it feasible to apply additional levels of control to the cutting process by utilising readily available, inexpensive, and reliable computing power. However, control of any process requires sensing capabilities to provide essential feedback data. One of the most significant pieces of feedback data in a cutting process is the state of tool wear [6].

The sensed tool wear data can be used for two major objectives: (i) detection of tool failure due to excessive wear; and (ii) control of the rate of tool wear. First, tool failure detection can be achieved by the on-line monitoring of the tool, thereby, eliminating the function of the machine operator as the tool monitor. Second, the rate of tool wear can be controlled by the manipulation of the cutting conditions, thereby, providing (i) better coordination in the production line by planning the machine shut down for tool changes,

(ii) more efficient use of the tool, and (iii) significant savings of time which is usually spent on the more frequent tool changes caused by unreliable estimates of tool life.

### 1.2. EXISTING APPROACHES

The on-line tool wear measurement problem has been investigated by numerous researchers [7]. The proposed methods can be **categorised** into two groups: direct and indirect. Direct methods, as the name implies, make an assessment of tool wear by either evaluating the worn surface by optical methods, or measuring the material loss of the tool by radiometric techniques. The main difficulty with using optical methods is their long processing time which makes them unsuitable for on-line tool wear measurement, and their limited application to cases where the surface of the tool is visually accessible during the operation [8]. Radiometric techniques have been proposed and demonstrated in the laboratory. These techniques have not been implemented in production, however, due to requirements for special preparation of the tool and potential hazards due to radioactivity [9].

Indirect methods, on the other hand, are based on utilising signals such as force or torque, temperature, tool vibration, or acoustic emissions [10–13]. These techniques which estimate tool wear by correlating it with some other measured process variable use different approaches to find such a correlation. Some approaches rely on a detailed mechanistic model of the cutting process (e.g. [14]), while others use empirical relationships between the measured variable and tool wear (e.g. [15]). The mechanistic approach has contributed greatly to the basic understanding of the cutting process, while the empirical approach has been useful for specific tool-workpiece combinations and constant cutting conditions. Both the mechanistic and empirical approach have certain limitations, however, when applied to on-line tool wear estimation.

The mechanistic approach, which relies on the mathematical description of the physics of cutting, assumes certain wear mechanisms such as diffusion, abrasion, and adhesion as being responsible for tool wear. Due to the inherent complexity of the cutting process, and our incomplete understanding of it, this approach is limited in applicability. Moreover, since the coefficients and exponents of these models change with tool-workpiece combinations and cutting conditions, extensive off-line testing is required for each case. Another limitation in the utilisation of the mechanistic approach is the lack of appropriate sensors. For example, most models developed by this approach (mechanistic models) emphasise the relationship between tool wear and temperature (e.g. [16]). However, in the absence of a practical temperature sensor these models are limited in applicability.

The empirical approach, on the other hand, relies on experimentally observed relationships to detect tool failure or estimate tool wear. Although new approaches based on a multi-sensor strategy have recently been proposed [17], most practical methods based on the empirical approach rely on a single sensor to detect tool failure or estimate tool wear. The empirical methods for tool wear estimation usually consider the relationship between only one pair of variables (e.g. flank wear and force). Therefore, they fail to separate the effect of other variables involved in the process (e.g. the effect of feed on force). This usually causes serious errors in the estimation results when other process variables affect the measured **signal**. In other words, as long as the measured variable is not affected by any other process variable, the method works reliably. These measured variables, however, are usually affected by cutting conditions; therefore, in cases where the cutting conditions are manipulated for control purposes the approach fails to separate this effect from the one caused by tool wear. The empirical approach, like the mechanistic approach, requires accurate information about the equation parameters, therefore, extensive off-line testing is required for each tool-workpiece combination and each specific cutting situation.

### 1.3. THE PROPOSED METHOD

The method proposed here attempts to resolve the disadvantages associated with indirect on-line tool wear estimation methods. This method which is based upon adaptive control theory [18], first, tends to separate the effect of different variables involved in the process (inputs as well as different wear components) by relying on a process model—the inability to separate the effect of different variables is a major drawback in empirical methods. This process model must be defined in terms of input, output, and state variables, such as shown in Fig. 1, and its state variables must represent the various types of tool wear. Second, it tends to be less dependent on off-line testing by using on-line parameter estimation [19]—determination of equation parameters is a definite problem in both mechanistic and empirical methods.

The adaptive observer treats the manufacturing process (turning operation in this case) as a “black box”, and as far as the implementation of the adaptive observer is concerned, the process model is only used to (i) determine the order of the adaptive observer, and (ii) specify the input, output, and state variables in order to discriminate between the contribution due to each one of them. However, before the adaptive observer is implemented for tool wear estimation it has to be designed and tuned. The design of the adaptive observer is performed by simulating the tool wear components of interest in a turning operation. For simulation purposes, a detailed dynamic state model of tool wear has been developed.

The model uses existing mechanistic and/or empirical relationships available in the literature. It is assumed that the form of the relationships represent the general behavior of the system, and that at least the structure of the model is a close approximation to the structure of the actual process. It is also assumed that, while the parameters of the model can change, the general relationships between variables hold under all machining conditions.

The proposed approach uses both on-line state and parameter estimation as illustrated in Fig. 2. The state observer is used to estimate the states of the model which represent

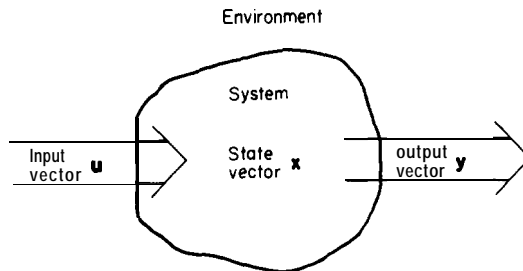


Figure 1. Schematic of a process model showing the input, state and output variables.

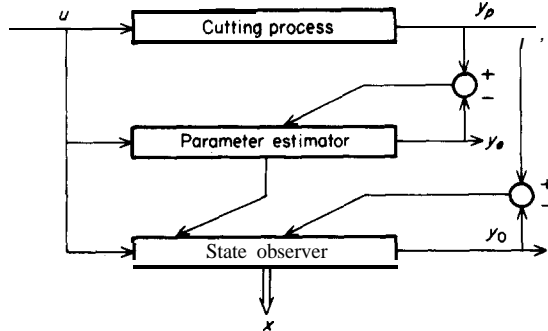


Figure 2. Schematic of an adaptive observer.

the tool wear components of interest. In order to use the state observer the model is transformed into a particular minimal realisation termed the observer form. The state observer estimates the states of this transformed model using measureable input and output signals. Since the parameters of this state observer vary with process conditions and material properties, an on-line parameter estimator is also required. The estimator estimates only the parameters of the model in the observer form. Once the states of the model in the observer form have been estimated, they will have to be transformed into the physically meaningful states of the original form of the model. This will have to be done through off-line measurements.

The proposed approach has several advantages: (i) it is based on a model which is adjusted on-line (by the parameter estimator), therefore, a smaller data base is needed; (ii) it can be applied to various kinds of manufacturing processes and types of tool wear (with the appropriate model) and (iii) it is inexpensive and does not require any modifications in the process components.

#### 1.4. ORGANISATION

This paper represents the first part of a two part study of an adaptive observer for on-line tool wear estimation: Part I (i) describes the process model used for simulation, (ii) evaluates the model as the basis for on-line tool wear estimation and (iii) discusses the design of the adaptive observer. Part II (i) investigates the feasibility of the reconstruction of the wear components from the observed states, (ii) analyses the results and (iii) discusses the experimental set-up and results.

## 2. MODEL OF TOOL WEAR IN TURNING

The proposed method for tool wear estimation is based upon a dynamic state model. This dynamic state model is presented in [20]. Here we only briefly discuss the choice of model variables and introduce the state and output equations.

A dynamic state model, representing a system, is generally defined in terms of input, output, and state variables [21]. The input variables are manipulated to affect the system (see Fig. 1). The output variables are always measurable and provide information about the system. The state variables determine the behavior of the system at any instant in time and are not necessarily measurable.

Mathematically, the equations defining the state variables (state equations) are represented by the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \quad (1)$$

where:  $\mathbf{x}$  is the vector of state variables;  $\mathbf{u}$  is the vector of inputs;  $t$  is time and the functions  $\mathbf{f}$  describe their relationship. Similarly, the equations defining the output variables (output equations) can be expressed as

$$\mathbf{Y} = \mathbf{g}(\mathbf{x}, \mathbf{u}, t) \quad (2)$$

where:  $\mathbf{y}$  is the vector of the output variables, and the functions  $\mathbf{g}$  define its relationships with the states and inputs of the system. Note that the relationship in equation (2) is static, whereas equation (1) represents a dynamic relationship.

As the first step in constructing the dynamic state model, the input, output and state variables have to be selected. The selection of these variables is based on the function of the model as the basis for parameter and state estimation. The inputs to a turning process are the feed ( $f$ ), the cutting speed ( $v$ ), and the depth of cut ( $d$ ). In this model the cutting speed and depth of cut are assumed to be constant, as is typically the case in

many turning operations. Thus, the input variable is selected to be the feed only for purposes of parameter and state estimation. The output is selected to be the cutting force (F), which is directly measurable during the process. The state variables, as a modeling requirement for state estimation, have to represent the wear components of interest. Therefore, they are selected to be two of the most frequently encountered types of wear, namely flank wear ( $w_f$ ) and crater wear ( $w_c$ ).

As the second step in the construction of the model, the state and output equations have to be determined. For this, the relationships reported in the literature for single point turning operations are reviewed and those which define the simplest and most clear relationship between the selected model variables (input, output and state variables) are selected. It should be emphasised here that the selection of state and output equations is a modeling decision and other relationships could be adopted within the same general model development. The significance of each model variable in turning, a detailed literature review on the relationships representing these model variables, and the particular relationships selected as the state/output equations are presented in [20]. The state and output equations are repeated here for easy reference.

State equations:

$$(l_0/v)\dot{w}_{f1} + w_{f1} = K_1 \cos \alpha_r F / (fd) \quad (3)$$

$$\dot{w}_{f2} = K_2 \sqrt{v} \exp [-K_3 / (273 + \theta_f)] \quad (4)$$

$$\dot{w}_c = K_4 F v \exp [-K_5 / (273 + \theta_c)] \quad (5)$$

where:

$$\theta_f = K_6 v^{n_1} f^{n_2} + K_7 w_{f2}^{n_3} \quad (6)$$

$$\theta_c = K_8 F v^{n_4} f^{n_5} d^{n_6} \quad (7)$$

Output equation:

$$F = [K_9 f^{n_7} (1 - K_{10} \alpha_r) - K_{11} - K_{12} v] d + K_{13} d w_f - K_{14} w_c \quad (8)$$

where:  $w_{f1}$  represents the component of flank wear caused by abrasion;  $w_{f2}$  represents the component of flank wear caused by diffusion;  $w_c$  represents the crater wear and  $\theta_f$  and  $\theta_c$  are the tool-work and tool-chip temperatures respectively. The total flank wear  $w_f$  is defined as

$$w_f = w_{f1} + w_{f2}.$$

### 3. MODEL VALIDATION

Before we can proceed with the design of the adaptive observer, however, the suitability of the model for the purpose of on-line tool wear estimation should be determined. We do this by studying the following points, which are discussed in detail in [20]:

1. Simulation results using the non-linear model. As the basis for on-line tool wear estimation it is crucial to have good agreement between the model and the process. The simulation results show very good qualitative agreement with the typical results reported in the literature [20]. They show that (i) model states increase with increasingly heavier cutting conditions, (ii) the diffusive component of flank wear becomes more dominant than the abrasive component at the heavier cutting conditions and (iii) the tool failure is caused by crater wear at heavier cutting conditions.

2. Stability of the system as predicted by the model. The stability of the model was investigated by studying its eigenvalues. In order to obtain the eigenvalues of the model,

the model was linearised. The linearisation method used is explained in the Appendix. The eigenvalues indicate that the model is unstable, as expected due to the continuously increasing nature of the wear components.

3. Controllability of the model as it relates to parameter estimation. Controllability, which is a necessary condition for parameter estimation [22], guarantees that all the modes of the system will be affected by the input variable. Of course, the goal here is to utilise this effect for parameter estimation and not for control. In other words, we are not trying to reduce wear (physically wear cannot be reduced through the manipulation of the input), but rather to affect the wear rate in order to excite all the modes of the system (i.e. to provide a sufficiently rich input for parameter estimation purposes). The controllability of the model was studied by considering the determinant of the controllability matrix. The determinant of the controllability matrix was non-zero for a variety of cutting conditions which implied that the model was controllable at the operating points considered.

4. Observability of the wear components by the cutting force measurement. Observability, which is a necessary condition for both parameter and state estimation, guarantees that all the modes of the system are observable by the output variable. The observability of the model was investigated through the determinant of the observability matrix. Since the determinants were non-zero for the considered operating points, it was concluded that the model was observable at those operating points.

Based on the above study the developed model is considered to be suitable as the basis for parameter and state estimation.

#### 4. THE PROPOSED METHOD

The proposed method uses an adaptive observer to estimate tool wear on-line. The development of such an adaptive observer is discussed in this section. In the first part, it is argued that in order to use a state observer, the process should be representable by a linear time-invariant observable model. Since the cutting process cannot be represented by a linear model, an adaptive observer should be used. In the second part, the different approaches for adaptive observation are discussed and the approach used here for the tool wear estimation problem is described. This approach uses a parameter estimator to update the model parameters. Such a parameter estimator is developed in the third part of this section.

##### 4.1. STATE OBSERVATION

State observers, first introduced by Luenberger [23], estimate the states of a linear time-invariant model. If a machining process can be represented by a linear time-invariant single-input single-output observable deterministic model such as,

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \quad (9)$$

$$y = \mathbf{c}^T \mathbf{x} + \mathbf{D}u \quad (10)$$

where the state vector  $\mathbf{x}$  includes the wear components of interest, then, an observer (state estimator) can be designed to estimate the state of the process [24]. In the above model  $u$  is a scalar input,  $y$  is a scalar output,  $\mathbf{A}$  is a coefficient matrix,  $\mathbf{b}$  and  $\mathbf{c}$  are coefficient vectors, and  $\mathbf{D}$  is a constant coefficient. In order to design the observer, the above model (equations (9) and (10)) can be transformed into discrete-time form [21] by using a

zero-order hold approximation and a constant sampling period,

$$\mathbf{x}(k+1) = \mathbf{P}\mathbf{x}(k) + \mathbf{q}u(k) \quad (11)$$

$$y(k) = \mathbf{c}^T \mathbf{x}(k) + Du(k). \quad (12)$$

The transfer function of the above system can be defined as,

$$\frac{Y(z)}{U(z)} = \frac{b_0 + \sum_{j=1}^m b_j z^{-j}}{1 + \sum_{i=1}^n a_i z^{-i}} + D. \quad (13)$$

The above discrete-time model (equations (11) and (12)) can be transformed into the observer form through a transformation matrix  $\mathbf{T}$  such that,

$$\mathbf{x} = \mathbf{T}\mathbf{x}_0 \quad (14)$$

The observer form for a third-order linear model is represented as [24],

$$\mathbf{x}_0(k+1) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} \mathbf{x}_0(k) + \begin{Bmatrix} b_1 \\ b_2 \\ b_3 \end{Bmatrix} u(k) \quad (15)$$

$$y(k) = [1 \ 0 \ 0] \mathbf{x}_0(k) + Du(k). \quad (16)$$

Now an estimate of the state ( $\hat{\mathbf{x}}_0(k)$ ) can be obtained through a state observer of the form,

$$\hat{\mathbf{x}}_0(k+1) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} \hat{\mathbf{x}}_0(k) + \begin{Bmatrix} b_1 \\ b_2 \\ b_3 \end{Bmatrix} u(k) + \begin{Bmatrix} g_1 \\ g_2 \\ g_3 \end{Bmatrix} [y(k) - \hat{y}(k)] \quad (17)$$

$$\hat{y}(k) = [1 \ 0 \ 0] \hat{\mathbf{x}}_0(k) + Du(k) \quad (18)$$

where the  $g_i$  are the gains of the observer and can be selected based on the desired convergence characteristics of the state observation error ( $e(k) = \mathbf{x}_0(k) - \hat{\mathbf{x}}_0(k)$ ). The equation describing the state observation error,  $e(k)$ , can be obtained by subtracting equation (17) from equation (15) to obtain

$$\mathbf{e}(k+1) = \begin{bmatrix} -(a_1 + g_1) & 1 & 0 \\ -(a_2 + g_2) & 0 & 1 \\ -(a_3 + g_3) & 0 & 0 \end{bmatrix} \mathbf{e}(k). \quad (19)$$

Thus, we can see that the observation error,  $e(k)$ , can be forced to decay at a desirable rate by appropriate selection of  $g_1, g_2$  and  $g_3$ . Also note that the model parameters ( $a_i$  and  $b_i$ ) used in the observer (equations (17) and (18)) are the same as in the model (equations (15) and (16)). Therefore, the assumption that these parameters are known and time-invariant is essential here [22].

#### 4.2. ADAPTIVE OBSERVATION

In the previous part it was shown that an observer can be used to estimate the states of a process if that process can be represented by a linear time-invariant observable model. However, only rarely can a process be fully represented by a linear model. Linear models, which (at best) can provide only an approximate representation of the process,

are generally used for their simplicity. The inadequacy of linear models in representing a process is even more pronounced when machining processes are concerned. These processes, which are extremely complex, can only be represented by highly non-linear models. Moreover, the parameters of these non-linear models depend on the tool-workpiece combination and the cutting conditions; namely the cutting speed, feed, and depth of cut.

In addition to the highly non-linear nature of machining processes, uncertainties associated with unmodeled process variables (e.g., "hard spots" in the workpiece and chip entanglement) make the modeling task very difficult. These unmodeled variables, however, can be eliminated by a careful selection of the workpiece and cutting conditions. It is also assumed here that the measurement noise can be filtered out of the measurement signal. Therefore, stochastic modeling is not required and the machining process can be represented by a deterministic non-linear model. In cases where the process is represented by a non-linear model, the state variables can still be estimated. In these cases, state estimation is done by using an adaptive observer.

Adaptive observers have been investigated by numerous researchers [25-29]. All the proposed adaptive observers, however, appear to have one essential inherent feature. The parameters of the system are identified either implicitly or explicitly. The implicit approach estimates the parameters as part of the observation scheme, while the explicit approach estimates the parameters separately and then uses them in the observer. Overall, the explicit approach has a simpler form and is preferred here for on-line tool wear estimation.

The adaptive observer used here consists of the combined application of parameter estimation and state observation. In this adaptive observer it is assumed that (i) the system parameters vary slowly as compared with the states, and (ii) the initial observation error due to poor initial parameter convergence is tolerable. This adaptive observer has the same form as the linear observer defined by equations (17) and (18) except that its parameters ( $\mathbf{a}_i$  and  $\mathbf{b}_i$ ) will be unknown and will have to be estimated by the parameter estimator [22].

The design of the adaptive observer consists of the design of a linear observer and a parameter estimator. For the design of the state observer the developed non-linear model had to be linearised. The linearisation was performed by the method described in the Appendix. Since fast convergence rate is highly desirable in the observer, it was decided to locate the eigenvalues at the origin of the z-plane (i.e. a deadbeat observer). For this, we would have

$$\mathbf{g}_1 = -\hat{\mathbf{a}}_1, \mathbf{g}_2 = -\hat{\mathbf{a}}_2, \mathbf{g}_3 = -\hat{\mathbf{a}}_3.$$

Of course, one should note that the parameters of the observer (see equation (17)) are not known and have to be estimated by the parameter estimator. Therefore, in cases where parameter estimation is poor one should expect to have poor performance when using this observer.

#### 4.3. PARAMETER ESTIMATION

When the parameters of the linear model in equation (15) are slowly time varying and/or unknown, a parameter estimator can be used to estimate the parameters of the model ( $\mathbf{a}_i$  and  $\mathbf{b}_i$ ) and update them whenever changes occur [22]. Parameter estimation algorithms have different forms, and are usually selected based upon the requirements for the specific problem of interest. These algorithms have a much simpler form when used for cases where the process can be represented by a deterministic model i.e. where the output signal and noise spectra are non-overlapping. Among the various methods for



parameter estimation the method of least squares is used here [30],

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + \frac{\mathbf{P}(k-2)\boldsymbol{\phi}(k-1)}{\boldsymbol{\beta} + \boldsymbol{\phi}(k-1)^T\mathbf{P}(k-2)\boldsymbol{\phi}(k-1)} [y(k) - \boldsymbol{\phi}(k-1)^T\hat{\boldsymbol{\theta}}(k-1)] \quad (20)$$

$$\mathbf{P}(k-1) = \frac{1}{\boldsymbol{\beta}} \left[ \mathbf{P}(k-2) - \frac{\mathbf{P}(k-2)\boldsymbol{\phi}(k-1)\boldsymbol{\phi}(k-1)^T\mathbf{P}(k-2)}{\boldsymbol{\beta} + \boldsymbol{\phi}(k-1)^T\mathbf{P}(k-2)\boldsymbol{\phi}(k-1)} \right] \quad (21)$$

where  $y(k)$  is the value of the measured variable  $y$  at time  $t = k\Delta t$  for  $k = 0, 1, 2, \dots$ ,  $\boldsymbol{\phi}(k)$  is the vector of measured (or known) variables, and  $\hat{\boldsymbol{\theta}}(k)$  is a vector of parameter estimates. The parameter  $\mathbf{P}(k)$  is the  $(n + m, n + m)$  matrix of estimation gains ( $n$  and  $m$  are the same as in equation (13)), and  $\boldsymbol{\beta}$  provides exponential data weighting. The above algorithm recursively updates the estimated parameter vector  $\hat{\boldsymbol{\theta}}(k)$  for any process whose equations can be written in the form,

$$y(k) = \boldsymbol{\phi}(k)^T\boldsymbol{\theta}(k)$$

Thus, the process model must be written in a form that is linear in the unknown parameters, which are the elements of the vector  $B(k)$ .

Parameter estimators have to be designed for the particular problem of interest. The estimated parameters, as already discussed, are the parameters of linear models obtained by linearising the non-linear model in different regions of its trajectory. In order to check the accuracy of the estimated parameters, the parameters of these linear models were needed. The linear models and their parameters were obtained by (i) linearising the non-linear model in several regions of its trajectory for a typical set of cutting conditions, (ii) discretising the obtained continuous-time linear models and obtaining their transfer functions and (iii) obtaining the  $\mathbf{a}_i$  and  $\mathbf{b}_i$  from these transfer functions.

It is interesting to note that for reasonable amounts of wear the parameters do not vary much along the trajectory. These parameters, however, change drastically when excessive amounts of wear develop.

#### 4.3.1. Estimation model

The discrete-time form of the linear model is defined by equations (11) and (12). The transfer function of this model is defined by equation (13). The component  $\mathbf{D}$  in equation (12) represents the direct effect of the input variable,  $\mathbf{u}$ , on the output variable,  $y$ . This effect is termed the “static” value of the output here for easy reference. Since in digital control the output is usually measured first and then the input variable is computed, in most models developed for digital control the  $\mathbf{D}$  component is zero [24]. This means that the “static\*” value of the output is not considered in the model.

In order to eliminate the static value of the output from the model equations (11) and (12) can be written as

$$\mathbf{x}(k+1) = \mathbf{P}\mathbf{x}(k) + \mathbf{q}\mathbf{u}(k) \quad (22)$$

$$\mathbf{A}y(k) = \mathbf{c}\mathbf{x}(k) \quad (23)$$

where

$$\Delta y(k) = y(k) - \mathbf{D}\mathbf{u}(k). \quad (24)$$

In this new model it is assumed that  $\Delta y(k)$  can be computed as the output of the model. Computing  $\mathbf{A}y(k)$  requires that the “static” value of the output be known, so that it can be removed from the output. This static value, which is dependent on the component  $\mathbf{D}$ , is only known when  $\mathbf{D}$  is known or estimated. The component  $\mathbf{D}$  in equation (24) is not known, but it can readily be estimated. At the beginning of the cut, when the tool is still

sharp ( $x(k) = 0$ ), the cutting force is not affected by wear and we can write from equation (12),

$$\frac{y(0)}{u(0)} = D \quad (25)$$

therefore, using the above equation,  $D$  can be estimated and  $Du(k)$ , the static value of the cutting force, can be removed from the cutting force. Now, if we assume that  $D$  stays constant, we can continue to compute  $Du(k)$  during the process by using the value of  $D$  computed at the beginning of the cut. In cases where the input is manipulated for control purposes, equation (25) can be used to re-estimate the value of  $D$  whenever the input is changed. This can be done by measuring the ratio between the change in output and the change in input. In doing this, however, we have to make sure that the value we consider as the change in output is only caused by the change in input and the effect of wear is negligible. Since wear is a slow process and needs time to develop, if we measure the output right before and after the change in input and do not allow any time for wear to develop, we can safely assume that the measured change in output is completely caused by the change in input and the effect of wear is negligible in this interval. Equations (22) and (23) were used as the estimation model in designing the parameter estimator.

For our third-order linear model  $\hat{\theta}$  and  $\phi$  are defined as

$$\hat{\theta}^T(k) = [\hat{a}_1(k) \quad \hat{a}_2(k) \quad \hat{a}_3(k) \quad \hat{b}_1(k) \quad \hat{b}_2(k) \quad \hat{b}_3(k)] \quad (26)$$

$$\phi^T(k) = [-\Delta y(k-1) \quad -A y(k-2) \quad -A y(k-3) \quad u(k-1) \quad u(k-2) \quad u(k-3)] \quad (27)$$

where  $\Delta y(k) = y(k) - Du(k)$ .

Once the estimation model is determined, the parameter estimator can be tuned for the non-linear model. The tuning of the parameter estimator involved the resolution of several problems. These problems were the following:

4.3.1.1. Numerical sensitivity. Parameter estimators are known to be quite sensitive numerically. They particularly have difficulty when the magnitudes of the input,  $u(k)$ , and the output,  $y(k)$ , are very different. In order to reduce the numerical sensitivity of the algorithm the vector  $\phi$  was scaled [31]. This scaled vector has the form

$$\begin{aligned} \phi^T(k) = & [-\Delta y(k-1)/s_1 \quad -\Delta y(k-2)/s_1 \quad -A y(k-3)/s_1, \\ & \times u(k-1)s_2 \quad u(k-2)s_2 \quad u(k-3)s_2] \end{aligned} \quad (28)$$

The scaling of  $\phi$  also requires scaling of  $\theta$ ,

$$\hat{\theta}^T(k) = [\hat{a}_1 s_1 \quad \hat{a}_2 s_1 \quad \hat{a}_3 s_1 \quad \hat{b}_1/s_2 \quad \hat{b}_2/s_2 \quad \hat{b}_3/s_2] \quad (29)$$

where  $s_1$  and  $s_2$  in the above equations are the scale factors. Also, in order to improve the numerical characteristics of the estimator it was decided to use the  $U-D$  factorisation method. This algorithm, which is discussed in detail in [32], improved the performance of the estimator considerably.

4.3.1.2. Lack of excitation. In order to guarantee the convergence of the estimated parameters, the system had to be excited persistently [30]. This would require the existence of a "rich" input to the system. The input used was varied in steps of  $\pm 10\%$  about a nominal value to excite the system. However, the applied input did not improve the performance of the estimator, and the estimated parameters were fairly poor. At this point, the effect of the input change on the parameters of the non-linear model was investigated. This was done by linearising the non-linear model at two different input

levels (the other conditions being the same), and computing their parameters. The computed parameters of the two models were quite **different**, (see Table 1). This meant that the change in input changed the parameters and thus deteriorated the estimation results. To alleviate this problem it was decided to rely on the initial step change in the input to estimate the parameters. In practice, however, random process input noise will generally exist and the estimated parameters are then guaranteed to converge [33]. In the case of simulation and in the absence of any process noise to guarantee the convergence of parameters, parameter constraints were used to avoid any drift in the estimates. The constraints, shown in Table 2, were selected based on the general trend of the parameters obtained from the linear models (see Table 3). Since in the model parameters changed along the trajectory, it was decided to use a "forgetting factor" to track the slowly time-varying parameters. For this,  $\beta$  in equations (20) and (21) was selected as

$$\beta = 0.95$$

The designed parameter estimator was then applied to the non-linear model. The estimated parameters (see Table 3) were reasonably accurate.

The designed observer along with the parameter estimator were applied to the non-linear model. The estimated states from this test are listed in Table 4, which also shows the true states computed from the non-linear model. Based on the results in Table 4, one can state that the observed states are quite accurate. Table 4 also shows the mean and standard deviation of the state estimation error. The observer estimates the states of the model in observer form. Once the states in the observer form have been estimated, they should be transformed into the states of the original model, so they can be interpreted. This transformation will be carried out and discussed in Part II.

TABLE 1

*Estimated parameters of the non-linear model at two different input levels*

$f$	$v$	$d$	$a_1$	$a_2$	$a_3$	$b_1$	$b_2$	$b_3$
0.1	150	2.0	-2.9889	2.9978	-0.9889	15.71587	-31.4317	15.71586
0.125	150	2.0	-2.9881	2.9763	-0.9881	12.8257	-25.6514	12.8257

TABLE 2

*Parameter constraints*

$a_1$	$-3.2 < \hat{a}_1 < -2.8$
$\hat{a}_2$	$2.8 < \hat{a}_2 < 3.2$
$\hat{a}_3$	$-1.2 < \hat{a}_3 < -0.8$
$\hat{b}_2$	$-2.5\hat{b}_1 < \hat{b}_2 < -1.5\hat{b}_1$
$\hat{b}_3$	$0.8\hat{b}_1 < \hat{b}_3 < 1.2\hat{b}_1$

TABLE 3

**Estimated parameters of the non-linear model**

Time (min)	$\mathbf{a},$ $\hat{a}_1$	$\mathbf{a}_2$ $a_{2A}$	$\mathbf{a}_3$ $a_{3A}$	$\mathbf{b}_1$ $\hat{b}_1$	$\mathbf{b}_2$ $\hat{b}_2$	$\mathbf{b}_3$ $\hat{b}_3$
0·1	-2.98	2·95	-0.98	14.58	-29.21	14.63
	-3.00	3.00	-1.00	14.14	-28.28	14.14
1·0	-2.98	2.95	-0.98	13.78	-27.63	13.85
	-3.00	2.99	-1.00	14.14	-28.53	14.26
4·0	-2.98	2.95	-0.98	13·40	-26.88	13.48
	-3.00	2.99	-1.00	14·12	-28·54	14.24
12·0	-2.99	2.96	-0.98	15·20	-30·44	15.24
	-3.00	2.99	-1.00	14.15	-28·52	14.27
15.0	-2.98	2.96	-0.98	15.70	-31.42	15.72
	-3.00	2.99	-1.00	14.16	-28.51	14.28

TABLE 4

**Estimated states of the non-linear model**

Time (min)	$x_{01}$ $\hat{x}_{01}$	$x_{02}$ $\hat{x}_{02}$	$x_{03}$ $\hat{x}_{03}$
0·1	8·6269	-17.2853	8.6586
	4.2419	-8.5839	4.2420
1·0	65·677	-131.750	66.073
	62·893	-127.763	64.806
4·0	100.748	-203.481	102.734
	100.689	-205.971	105.277
12.0	26.660	-57.505	30.851
	26.308	-57.124	30.829
15·0	24.080	-52.172	28.104
	23.380	-50.856	27·504
Mean	1.66	-2.38	0.75
S.D.	1.86	4.22	2.51

## 5. SUMMARY

The use of an adaptive observer for on-line tool wear estimation has been proposed. An overview of the present methods for on-line tool wear estimation has been presented in the introduction. There, the proposed method has been compared with some of the existing tool wear estimation methods.

The design of the adaptive observer relies on a dynamic state model of the cutting process. This dynamic state model has been briefly discussed. The model uses flank wear and crater wear as state variables, feed as the input variable, and the cutting force as the output variable.

The developed non-linear model has been evaluated for its feasibility for parameter and state estimation. This evaluation has been performed by studying (i) the physical

behavior of the model by simulation, (ii) the stability of the model by eigenanalysis, (iii) the controllability characteristics of the model as related to parameter estimation and (iv) the observability of the states using the output (cutting force).

Since the developed model is found to be suitable for parameter and state estimation, the adaptive observer is then designed. The adaptive observer selected for this particular problem is based on the combined application of parameter and state estimation. Some of the issues in the design of the adaptive observer for the developed non-linear model are discussed and finally the adaptive observer is applied to the non-linear model. The results are found to be very satisfactory. The implementation of the adaptive observer in turning, and experimental results are discussed in Part II.

## REFERENCES

1. A. G. ULSOY, Y. KOREN and F. RASMUSSEN 1971 *American Society of Mechanical Engineers Transactions, Journal of Dynamic Systems, Measurement, and Control* 105 107-112. Principal development in the adaptive control of machine tools.
2. J. TLUSTY and G. C. ANDREWS 1983 *Annals of the CIRP* 32, 563-572. A critical review of sensors for unmanned machining.
3. P. K. WRIGHT 1984 presented at the *American Society of Mechanical Engineers Winter Annual Meeting*. Physical models of tool wear for adaptive control in flexible machining cells.
4. S. JETLY 1984 *Manufacturing Engineering* 55-60. Measuring cutting tool wear on-line: some practical considerations.
5. Y. KOREN 1983 *Computer Control of Manufacturing Systems*. New York: McGraw-Hill.
6. D. W. YEN and P. K. WRIGHT 1983 *American Society of Mechanical Engineers Transactions, Journal of Engineering for Industry* 105, 3 1-38. Adaptive control in machining-a new approach based on the physical constraints of tool wear mechanisms.
7. S. K. BIRLA 1980 *Proceedings of the Machine Tool Task Force Conference*, Vol. 4 Sensors for Adaptive Control and Machine Diagnostics.
8. L. K. DANESHMAND and H. A. PAK 1983 presented at the *Third International Conference on Robot Vision and Sensory Controls (Rouisec3)*. Performance monitoring of a computer numerically controlled (CNC) lathe using pattern recognition techniques.
9. S. K. JETLY 1984 *Proceedings of the North American Manufacturing Research Conference*, pp. 255-259. A new radiometric method of measuring drill wear.
10. A. DEFILIPPI and R. IPPOLITO 1969 *Annals of the CIRP* 17, 377-397. Adaptive control in turning: cutting forces and tool wear relationships for P10, P20, P30 carbides.
11. M. P. GROOVER, R. J. KARPOVICH and E. K. LEVY 1977 *International Journal of Production Research* 25, 129-141. A study of the relationship between remote thermocouple temperatures and tool wear in machining.
12. P. MARTIN, B. MUTELS and J. P. DRAIPER 1975 *Proceedings of the 16th International Machine Tool Design and Research Conference*. Influence of lathe tool wear on the vibrations sustained in cutting.
13. E. KANNATEY-ASIBU, JR. and D. A. DORNFELD 1982 *Wear* 76, 247-261. A study of tool wear using statistical analysis of cutting acoustic emission.
14. A. BHATTACHARYYA and I. HAM 1969 *American Society of Mechanical Engineers, Transactions, Journal of Engineering for Industry* 790-798. Analysis of tool wear-Part 1: Theoretical models of flank wear.
15. A. DEFILIPPI and R. IPPOLITO 1972 *Annals of the CIRP* 21, 29-30. Analysis of the correlation among: cutting force variation (vs. time)-chip formation parameters-machining.
16. E. USUI, T. SHIRAKASHI and T. KITAGAWA 1978 *American Society of Mechanical Engineers Transactions, Journal of Engineering for Industry* 100, 236-243. Analytical prediction of three dimensional cutting process-Part 3: Cutting temperature and crater wear of carbide tool.
17. A. G. ULSOY and E. HAN 1985 *Proceedings of the IFAC Conference on Control Science and Technology for Development*. Tool breakage detection using a multi-sensor strategy.
18. Y. D. LANDAU 1979 *Adaptive Control*. New York: Marcel Dekker.
19. R. ISERMANN 1984 *Automation* 20, 387-404. Process fault detection based on modeling and estimation methods-a survey.

20. K. DANAI and A. G. ULSOY 1985 in *Sensors and Controls for Manufacturing*, pp. 137-148. A dynamic state model for on-line tool wear estimation in turning.
21. Y. TAKAHASHI, M. J. RABINS and D. M. AUSLANDER 1972 *Control*. Reading, MA: Addison-Wesley.
22. G. C. GOODWIN and K. S. SIN 1984 *Adaptive Filtering, Prediction, and Control*. Englewood Cliffs, NJ: Prentice-Hall.
23. D. G. LUENBERGER 1964 *Institution of Electronic and Electrical Engineers Transactions on Military Electronics* MIL-80, 74-80. Observing the state of a linear system.
24. K. J. ÅSTRÖM and B. WITTENMARK 1984 *Computer Controlled Systems*. Englewood Cliffs, NJ: Prentice-Hall.
25. R. L. CARROLL and D. P. LINDORFF 1973 *Institution of Electronic and Electrical Engineers Transactions on Automatic Control* 18, 428-434. An adaptive observer for single-input single-output linear systems.
26. P. KUDVA and K. NARENDRA 1973 *International Journal of Control* 18, 1201-1210. Synthesis of an adaptive observer using Lyapunov's direct method.
27. G. LUDERS and K. NARENDRA 1973 *Institution of Electronic and Electrical Engineers Transactions on Automatic Control* 496-499. An adaptive observer and identifier for a linear system.
28. G. KREISSELMEIER 1977 *Institution of Electronic and Electrical Engineers Transactions on Automatic Control* 22, 2-8. Adaptive observers with exponential rate of convergence.
29. P. MURDOCH and E. V. OLIVEROS 1985 *International Journal of Control* 41, 477-491. A discrete adaptive observer for a system of unknown order.
30. K. J. ÅSTRÖM and P. EYKHOFF 1971 *Automatica* 7, 123-162. System identification-A survey.
31. B. WITTENMARK and K. ÅSTRÖM 1984 *Automatica* 20, 595-605. Practical issues in the implementation of self-tuning control.
32. T. F. ELBERT 1984 *Estimation and Control of Systems*. New York: Van Nostrand Reinhold.
33. K. J. ÅSTRÖM and T. BOHLIN 1966 in *Theory of Self-Adaptive Control Systems*, (Ed. P. H. Hammond). Numerical identification of linear dynamic systems from normal operating records. New York: Plenum Press.

#### APPENDIX: LINEARISATION METHOD

As discussed earlier, the non-linear model is used as the basis for the design of the adaptive state observer. Since adaptive observation theory is well developed for linear models, the model in equations (3)-(8) is linearised. The linear model is then used to (i) study the stability, controllability, and observability conditions of the model and (ii) design the adaptive observer.

The non-linear model

$$\dot{\mathbf{x}} = \mathbf{f}'(\mathbf{x}, \mathbf{u}, t) \quad (\text{A1})$$

$$y = \mathbf{g}'(\mathbf{x}, \mathbf{u}, t) \quad (\text{A2})$$

can be linearised about an operating point  $(\mathbf{x}_1$  and  $\mathbf{u}_1)$ . If we define incremental variables, denoted by primes, about this operating point such that,

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}' \quad (\text{A3})$$

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}' \quad (\text{A4})$$

Then, the Taylor series expansion about  $\mathbf{x}_1$  and  $\mathbf{u}_1$  can be used to find a linear model such as,

$$\dot{\mathbf{x}}' = \mathbf{A}'\mathbf{x}' + \mathbf{b}'\mathbf{u}' \quad (\text{A5})$$

$$y' = \mathbf{c}'^T \mathbf{x}' + D'\mathbf{u}' \quad (\text{A6})$$

to approximately define the behavior of the model in the vicinity of the operating point. Note that in the above model the state variables will be  $\mathbf{x}'$  and the input variable will be  $\mathbf{u}'$  which are only incremental changes around the operating point. However, since the

operating points are not known ( $\mathbf{x}_1$  is not known), the knowledge about incremental changes in an unknown value is not meaningful.

The appropriate model for our purpose must represent the gross behavior of the system and not incremental changes about operating points. Such a model will have the form,

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \tag{A7}$$

$$y = \mathbf{c}^T \mathbf{x} + Du \tag{A8}$$

where  $\mathbf{x}$  represents the vector of state variables and  $u$  is the input variable defined by the linear model. This linear model is defined by linearising the non-linear model in a region of its trajectory which always starts at the origin ( $x = 0$  and  $u = 0$ ) and ends at an arbitrary "end point" ( $\mathbf{x}_1 = 0$  and  $u_1 = 0$ ) on the trajectory. In order to obtain such a model  $\mathbf{x}_1$  and  $u_1$  are divided into  $n'$  intervals such that,

$$\Delta \mathbf{x} = \frac{\mathbf{x}_1}{n'} \tag{A9}$$

$$\Delta u = \frac{u_1}{n'} \tag{A10}$$

and the Taylor series expansion is used to linearise the non-linear functions  $\mathbf{f}'(\mathbf{x}, u, t)$  and  $g'(\mathbf{x}, u, t)$  in each interval. The coefficients of this linear model are then defined as,

$$\mathbf{A} = \frac{1}{n'} \left[ \left. \frac{\partial \mathbf{f}'}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=0 \\ u=0}} + \left. \frac{\partial \mathbf{f}'}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=\Delta \mathbf{x} \\ u=\Delta u}} + \dots + \left. \frac{\partial \mathbf{f}'}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=(n'-1)\Delta \mathbf{x} \\ u=(n'-1)\Delta u}} \right] \tag{A11}$$

$$\mathbf{b} = \frac{1}{n'} \left[ \left. \frac{\partial \mathbf{f}'}{\partial u} \right|_{\substack{\mathbf{x}=0 \\ u=0}} + \left. \frac{\partial \mathbf{f}'}{\partial u} \right|_{\substack{\mathbf{x}=\Delta \mathbf{x} \\ u=\Delta u}} + \dots + \left. \frac{\partial \mathbf{f}'}{\partial u} \right|_{\substack{\mathbf{x}=(n'-1)\Delta \mathbf{x} \\ u=(n'-1)\Delta u}} \right] \tag{A12}$$

$$\mathbf{c} = \frac{1}{n'} \left[ \left. \frac{\partial g'}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=0 \\ u=0}} + \left. \frac{\partial g'}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=\Delta \mathbf{x} \\ u=\Delta u}} + \dots + \left. \frac{\partial g'}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=(n'-1)\Delta \mathbf{x} \\ u=(n'-1)\Delta u}} \right] \tag{A13}$$

$$D = \frac{1}{n'} \left[ \left. \frac{\partial g'}{\partial u} \right|_{\substack{\mathbf{x}=0 \\ u=0}} + \left. \frac{\partial g'}{\partial u} \right|_{\substack{\mathbf{x}=\Delta \mathbf{x} \\ u=\Delta u}} + \dots + \left. \frac{\partial g'}{\partial u} \right|_{\substack{\mathbf{x}=(n'-1)\Delta \mathbf{x} \\ u=(n'-1)\Delta u}} \right] \tag{A14}$$

where  $n'$  (the number of intervals) in equations (A11)-(A14) is selected based on the convergence criterion of the method. For this specific non-linear model,  $n'$  was selected to be 200. The partial derivatives in equations (A11)-(A14) are given in [20].