Hybrid Incremental Model for Imbalance Detection in Ultra-high Precision Rotating Devices

Rodolfo Haber, Member, IEEE, and Raúl M. del Toro

Abstract—A novel method based on a hybrid incremental modeling approach has been designed and applied to imbalance detection in ultra-high precision rotating machines. The model is obtained by a two-step iterative process that combines an overall model (least-squares fitting) with a local model (fuzzy k-nearest-neighbour) to take advantage of their complementary capacities. Three normalization strategies of evaluating the effect on accuracy are analyzed. A comparative study demonstrates that the hybrid incremental model provides better error-based performance indices for detecting imbalance than a nonlinear regression model and an adaptive neural-fuzzy inference system model. The suitability of Mahanolobis normalization for hybrid incremental modeling is also demonstrated in this case study. The proposed strategy for imbalance detection is simple, fast, and non-intrusive, reducing the deterioration in the performance of ultra-high precision rotating machines due to vibrations.

Index Terms—modeling, intelligent systems, imbalance detection, ultra-high precision rotating machine.

I. INTRODUCTION

Ultra-high precision rotation devices are today used intensively in micro-scale manufacturing processes. Typical manufacturing operations producing components with micro- or nano-scale features are performed by machine tools with nanometric resolution at the positioning axes [1, 2]. However, the precision of rotational devices is often limited by the complexity and nonlinear behavior of electro-mechanical devices. Intrinsic characteristics such as the nonlinear behavior of the variables involved, a low signal-to-noise ratio, a major influence of environmental factors, and uncertainty in sensory data constrain how the behavior of micro-scale manufacturing processes can be improved.

One research discipline that has seen the greatest advances in this area is soft computing, in which computing, reasoning, and decision-making take advantage of tolerance for imprecision, uncertainty, and approximate reasoning in order to obtain better solutions [3, 4]. Among soft-computing techniques, artificial neural networks and fuzzy systems have been the most widely applied for modeling, control, and diagnosis.

Applications of artificial neural networks are of particular interest because of their ability to accurately model any nonlinear function and their excellent learning capacity [5, 6]. Also, fuzzy systems have become a consolidated technology in the last few decades, although there are still some challenges to be resolved [7]. The potential of artificial neural networks is often combined with the ability of fuzzy systems to represent human thought and robustness in the presence of noise and process uncertainty [8, 9]. Fuzzy systems can also be combined with classical modeling techniques in order to better capture the most important characteristics of processes [10, 11]. Neurofuzzy systems have also demonstrated their ability to represent and control machining processes [12, 13].

One of the most important issues in manufacturing with rotating devices is how to deal with vibrations [14]. Vibrations can cause unwanted motion along the axis, and dynamic forces that arise during the rotation of the device reflect these unwanted movements. A clear example in air-bearing spindles is the increase in forces due to dynamic mass imbalance generating an eccentricity around the axis of rotation, thereby producing vibrations [15].

The main scientific goals of the present work are threefold. The first aim is to design and implement a model-based procedure for detecting spindle imbalance based on a hybrid incremental modeling strategy. The second is to evaluate the influence of data normalization on the modeling accuracy. And the third is to make a comparative study of three different model-based strategies —ANFIS, nonlinear regression, and hybrid incremental modeling— in order to evaluate their behavior and suitability. The study also has a technical goal – the design and implementation of a fast, simple, non-intrusive procedure for imbalance detection, focused on reducing the deterioration of the performance of rotational devices due to vibrations by means of accurately estimating the eccentricity.

The main contributions of the work can be summarized as follows. First, a computationally efficient algorithm is designed and implemented for imbalance detection of an ultra-high precision spindle for micro-turning processes. Second, the influence of normalization on the efficiency of the hybrid...
incremental modeling technique is analyzed, with the comparison of three common normalization strategies (standard, Mahalanobis, and diagonal), and the corresponding conclusions are tested in a real case study. Finally, a comparative study is made of hybrid incremental modeling with two other widely used techniques (ANFIS and regression).

The rest of the paper is organized as follows. Section II presents a brief description of hybrid incremental modeling. Section III analyzes the eccentricity and vibrations due to mass imbalance in rotational devices. The experimental platform and the comparative study are presented in Section IV, and the conclusions are presented in Section IV.

II. HYBRID INCREMENTAL MODELING

Hybrid Incremental Modeling (HIM) aims to provide an approximation of the behavior of a locally nonlinear system [16]. To that end, the strategy uses a global model that captures the basic or general behavior of the system and then superimposes a local model on it that captures the local behavior.

When no prior knowledge of the system is available, a generic model, such as a linear regression or second-order polynomials, can conveniently represent the global behavior of the system. If, however, some prior knowledge of the system is available, it might be possible to use a function that better matches the global behavior.

In our case, it was decided to use a generic strategy to obtain the global model by applying least squares to fit a polynomial of degree $m$. The output of the global model would therefore be as follows:

$$\hat{y}_g(x_i) = f_g(x_i, g(x_i))$$  \hspace{1cm} (1)

where, $x_i$ is the $i$-th input data, and $g(x_i)$ is its output (target) value.

The development of the local model is based on the Fuzzy $k$-Nearest Neighbors (F-$k$NN) approach. Compared to other techniques, F-$k$NN is simple, easily interpretable and can achieve an acceptable accuracy rate [17], [18], [19]. The fuzzy version of $k$-NN averages the value of the points closest to the query point, on the assumption that points close to each other have similar values [20]. However, standard $k$-nearest neighbor methods place equal weights on all the selected neighbors, regardless of their distances from the query data [21]. In this work, the problem is partially addressed by data normalization. Learning in fuzzy $k$-NN is simple, in the sense

The incremental model integrates the two models described above. As previously made clear, F-$k$NN locally adjusts the output value of the polynomial, as it does not capture the nonlinear localized characteristics of the system. The basic model is first trained by obtaining the coefficients of the polynomial of order $m$ that best fit the data.

Thus, let $\hat{y}_g(x)$ be the function that is the output of the basic model. Then the prediction error of the basic model is:

$$\epsilon(x) = g(x) - \hat{y}_g(x)$$  \hspace{1cm} (5)

These errors are the target values of the samples which are passed to the local model for training. It means that F-$k$NN does not use original input-output data, but the errors resulting from the global modeling strategy. Thus, the local modeling done by F-$k$NN is responsible for refining the global model output in regions with localized nonlinear behavior or where the polynomial cannot properly represent the system. It is important to remark that neighborhood size $k$ and fuzzy strength parameter $p$ are not necessary at this stage in the training, which considerably reduces the complexity of the modeling.

The incremental model evaluates a sample of data input $q$, by adding to the output of the basic model the compensation term calculated by the local model:

$$\hat{y}(q) = \hat{y}_g(q) + \hat{g}(q)$$  \hspace{1cm} (2)

A. Data normalization

The performance of the clustering algorithms is influenced by the characteristics of the input-output data, so that data normalization is an indispensable step in many real problems. Normalization addresses the problem that relevance information given to the resulting experimental data by different input systems may be totally incomparable. Due to different causes such as noise, different operating region among others, experimental data may be on different scales, in different ranges, and distributed differently.

There are many possible types of normalization procedures to scale data to a specific range [22]. For the sake of simplicity, we shall here only consider three of them: standard, Mahalanobis, and diagonal.

The most commonly used standard normalization technique is the $z$-score or ZMUV norm (zero-mean, unit-variance) that is calculated using the arithmetic mean and standard deviation of the given input-output data [23]. Denoting by $x_i^j$ the variable $j$ of input point $i$, one obtains the normalized points in the following form:

$$z_i^j = \frac{x_i^j - \bar{\mu}(j)}{\sigma(j)}$$  \hspace{1cm} (3)

$$\bar{\mu}(j) = \frac{\sum x_i^j}{n}$$  \hspace{1cm} (4)

$$\sigma(j) = \sqrt{\frac{\sum (x_i^j - \bar{\mu}(j))^2}{n-1}}$$  \hspace{1cm} (5)

where $\bar{\mu}(j)$ is the mean of the $j$-th variable, $\sigma(j)$ its standard deviation, and $n$ the length of the input data.

The simplest approach to normalization is to use Eq. (3) which is shift and scale invariant. It is also insensitive because
This transformation does not depend directly on either the min or max values given. This scheme can perform well if prior knowledge about the average of data and the variations are available. If we do not have any prior knowledge about the nature of the physical process, then we need to estimate the mean and standard deviation of the data from a given set of input-output data. The weakness of this approach, however, is that it does not take the couplings between the components into account. As a result, while the elements of the transformed vector indeed have zero mean and unit variance, they are still correlated, and the covariance matrix is not the identity matrix.

The sample Mahalanobis transformation [24], given a data matrix \( X = (x_{1}, \ldots, x_{n}) \), is \( z_{i} = S^{1/2}(x_{i} - \mu) \) for \( i = 1, \ldots, n \), where \( S = S_{x} \) is the sample covariance matrix given by

\[
S = \frac{1}{n-1}X^{T}HX
\]

This creates the normalized data matrix

\[
Z = S^{1/2}X^{1/2}H
\]

where \( H \) is a symmetric and idempotent matrix. Rewriting Eq. (7), one can obtain a simple representation for normalization based on Mahalanobis distances:

\[
z'_{i} = S^{1/2}(x'_{i} - \bar{\mu}(j))
\]

Normalization based on the Mahalanobis distance is therefore a weighted version of normalization based on the Euclidean distance, in which the weighting is determined by the range of variability of the sample point, expressed by the covariance matrix. Mahalanobis normalization has the advantage over other normalization techniques in that the contribution of highly correlated data does not differ too much from the contribution of uncorrelated or mildly correlated data [25].

If the covariance matrix is diagonal, it leads to the following diagonal normalization:

\[
S = \text{diag}(\sigma_{1}^{2}, \ldots, \sigma_{n}^{2})
\]

We used the C/Java programming language to implement the algorithm, and all tests were performed on both a Linux kernel and Windows XP.

III. ECCENTRICITY AND VIBRATIONS DUE TO MASS IMBALANCE IN ROTATIONAL DEVICES

There are processes (e.g., micro-turning) with very strict requirements of less than 50 nm in eccentricity of the spindle [26]. The maximum eccentricity is a key performance index of a rotational device's precision. Vibrations due to dynamic imbalance, for example, increase the form error (dimensional error), and lead therefore to less precision of the part being manufactured. Current on-line procedures for compensating the imbalance are expensive, tedious to implement, and far too dependent on the know-how and skill of the operator.

Eccentricity in the shaft of a rotating device occurs when its center of mass differs from its geometric center [27]. One of the commonest causes is mass imbalance of the device, mainly produced by an uneven distribution of its components’ masses. This eccentricity can generate dynamic forces that cause vibrations in synchrony with the device's rotation frequency [28].

The mass imbalance is also characterized by an eccentricity \( e \) away from the rotor's axial axis and an angle \( \phi \). If the rotor has angular velocity \( \omega \), the amplitude of the resultant force \( F \) and its components, \( F_{x} \) and \( F_{y} \), due to the imbalance are given by

\[
F_{x} = me\omega^{2}\cos\phi, \quad F_{y} = me\omega^{2}\sin\phi \Rightarrow F = me\omega^{2}, \phi = \omega t
\]

These forces constitute a harmonic excitation of the rotating device, causing vibrations with the same direction and frequency as the exciting force. In order to estimate these vibrations mathematically, the rotating device could be considered as a spring-mass-damper system, with coefficient of viscosity \( c \) and elasticity \( k \). For the sake of simplicity, the excitation will only be considered in one direction. The elasticity and viscosity of other rotor components (e.g., the bearings) will not be considered explicitly in this analysis.

If \( y \) is the displacement of the non-rotational mass \((M-m)\) from the equilibrium position then the displacement \( y_{m} \) of the imbalance mass \( m \) is

\[
y_{m} = y + e\sin\omega t
\]

The general equation of motion is given by

\[
(M-m)\frac{d^{2}y}{dt^{2}} + m\frac{d^{2}y_{m}}{dt^{2}} = -ky - c\frac{dy}{dt}
\]

Equation (16) can be simplified as follows:

\[
M\frac{d^{2}y}{dt^{2}} + c\frac{dy}{dt} + ky = (me\omega^{2})\sin\omega t
\]

The excitation input is the imbalance force component in the \( y \) direction (\( F_{y} \)). The solution of Eq. (13) has two parts – the homogeneous and the particular solutions. The homogeneous solution describes the transient behaviour of the system. It is a free vibration that can be under-damped, over-damped, or critically damped. In the steady state, the response
of the system is characterized by the particular solution of the equation, which is an oscillatory vibration of the same frequency as the excitation $F_y$ with amplitude $Y$ and phase $\phi$:

$$y = Y \sin(\omega t - \phi)$$

$$Y = \frac{m e \omega^2}{\sqrt{(k - M \omega^2)^2 + (c \omega)^2}} = \frac{m e}{M \bar{\omega}^2}$$

$$\phi = \tan^{-1} \left( \frac{c \omega}{k - M \omega^2} \right) = \tan^{-1} \left( \frac{2 \zeta \omega}{1 - \bar{\omega}^2} \right)$$

$$\bar{\omega} = \frac{\omega}{\omega_n}; \quad \omega_n^2 = \frac{k}{M}; \quad 2 \zeta \omega_n = \frac{c}{M}$$

where $\zeta$ is the damping factor of the system and $\omega_n$ its natural frequency.

From the second derivative of $y$, the acceleration of the motion can be expressed as

$$\ddot{y} = -\omega^2 Y \sin(\omega t - \phi)$$

The above equations represent the relationship between the eccentricity, caused by the mass imbalance, and the vibrations that take place in a rotating device [29]. The amplitudes of both the vibration and its acceleration are proportional to the imbalance mass and its eccentricity.

The methods that are currently used to balance ultra-precision rotation devices are usually based on determining the position and magnitude of the imbalance masses, and then adding compensatory masses to the rotor (of the same magnitude, and diametrically opposite), as shown in Fig. 1. Some machines include manual procedures or special modules with which to carry out the compensation, in this case usually based on measuring the eccentricity of the axis at different points by means of capacitive or inductive sensors, and determining the position of the greatest eccentricity. The compensation involves manually adding masses to this position until the eccentricity is within the bounds required for the corresponding process.

### A. Eccentricity estimation

One way to address the compensation of imbalance in rotating devices is by applying a model-based strategy in conjunction with signal processing strategies. Frequency and time domain information can be appropriately combined to estimate the eccentricity. The model proposed here has three inputs: the spindle rotation speed, $RF$, the overall amplitude of harmonics around the first harmonic of the vibration signal spectrum, $A_{T1}$, and around the second harmonic of that spectrum, $A_{T2}$. The output of the model is the maximum eccentricity of the axis of rotation due to the mass imbalance, $UL_E$. The parameters $AT_1$ and $AT_2$ provide information on the vibration in the frequency domain, and $RF$ provides information about the spindle speed.

![Fig. 2. Outline of the eccentricity estimation procedure.](image)

Figure 2 is an outline of the eccentricity estimation procedure in which the harmonic sequence detection is by means of spectral analysis [30]. More details on the algorithm are given in [27]. The model inputs $AT_1$ and $AT_2$ are calculated as the overall sum of the amplitudes of the sideband harmonics detected around the two main harmonics. Thus the value of each $AT_i$ input is only related to the harmonics generated due to eccentricity.

$$A_{T_i} = \sum_{j=1}^{n} A_{i,j}$$

where $A_{i,j}$ represents the amplitude of each $j$ sideband harmonic around the $i$-th principal harmonic, where $n$ is the total of new harmonics within a sequence.

### IV. EXPERIMENTAL PLATFORM: COMPARATIVE STUDY.

In order to evaluate and implement the imbalance detection procedure experimentally, a multi-sensory system was installed on a spindle mounted on an ultra-high precision lathe. The overall system is shown in Fig. 4: from left to right, the ultra-high precision lathe, the accelerometers installed on the spindle, and the measurement and data acquisition system. Ultra precision lathes are employed for finishing operations on curved and flat surfaces of both brittle and ductile materials,
with very small tolerances. The allow components (e.g., an optical lens) to be manufactured with a mean surface roughness of less than 10 nm and a few hundred nm of form precision.

Vibration signals were measured with two accelerometer sensors rigidly attached to the spindle housing. The sensor model was 352C15 from PCB Piezotronics, which has a sensitivity of 10 mV/g and a bandwidth of 12 kHz. The vibration signals of the X and Y spindle axes are acquired and processed with the high performance processor PXI-8187 from National Instruments, which has a sampling frequency of 50 kHz. The ultra-high precision lathe is a Nanoform 200 from Precitech Inc. It is located in an industrial environment as part of a functioning production line.

The reference eccentricity value was obtained from a measurement system embedded in the lathe’s computer numerical control (CNC). The amplitude and phase of this value correspond to the maximum eccentricity position of the spindle shaft, and are displayed on the CNC’s graphical user interface. These values can only be obtained prior to each manufacturing operation. For the experimental analysis, only the eccentricity amplitude was used as the reference value.

In order to analyze the relationship between the level of vibration and shaft eccentricity, different operating conditions of the spindle were studied: not rotating, and rotating at different speeds and eccentricities. Some of these experimental operating conditions are listed in Table I. The eccentricity was produced by manually adding imbalance masses to the spindle.

According to expert operator criteria, the acceptable tolerance for spindle shaft eccentricity was 50 nm. Values of less than 50 nm were therefore not considered relevant to the process.

Table I. Operating conditions of experiments for the study of spindle vibrations.

<table>
<thead>
<tr>
<th>Spindle speed [rev/min]</th>
<th>Eccentricity [nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>14, 192, 205, 309</td>
</tr>
<tr>
<td>2000</td>
<td>17, 635</td>
</tr>
<tr>
<td>3000</td>
<td>17, 39, 98, 162, 1358</td>
</tr>
<tr>
<td>5000</td>
<td>56, 2874</td>
</tr>
</tbody>
</table>

A. Hybrid incremental model

For the sake of computational simplicity the set of parameters selected was \((m, k, p) = (1, 1, 2)\). This means a first-order polynomial (linear), only one neighbor for the neighborhood size, and a value of 2 for the fuzzy strength parameter. The global model is obtained after the training. The parameters for the first-order model were 
\[
[-0.0000001 \quad 0.970536 \quad 0.025334 \quad -0.063098].
\]

The output of the model is obtained by applying Eq. (2).

B. Experimental results

In order to evaluate the behavior of the model two important issues were considered in the study. First, the influence of data normalization on the accuracy of the hybrid incremental strategy was analyzed. Second, a comparative study was made with a regression model and an adaptive neural-fuzzy inference system (ANFIS).

A regression model is used in the comparative study. A single-input single-output model relates the eccentricity \(( \text{ULE} )\) of the spindle shaft with the power ratio \(( \text{PRA} )\) of the detected sequence around the spectrum's principal harmonic. The model for estimating the eccentricity is obtained by inversion.

\[
PRA = \frac{1}{2} \left[ \tanh(m \cdot \text{ULE} + n) + 1 \right]
\]  

\[
\text{ULE} = \begin{cases} 
UL_{\text{min}}, & PRA \leq PRA_{\text{min}} \\
\frac{1}{m} \left[ \frac{1}{2} \ln \left( \frac{PRA}{1-PRA} \right) - n \right], & PRA > PRA_{\text{min}} 
\end{cases}
\]

\[
PRA_{\text{min}} = \frac{1}{2} \left[ \tanh(m \cdot \text{ULE}_{\text{min}} + n) + 1 \right]
\]

where \(m=2.220126\) and \(n=-1.59696\) are linear fit coefficients. The minimum value \(\text{ULE}_{\text{min}}\) was set to 0.010 µm.

The Adaptive Network-based Fuzzy Inference System (ANFIS) was also selected to perform the comparative study. ANFIS as well as being a pioneer in the field, is the simplest such system computationally, and thus feasible for real-time applications [31].

Each model input has three generalized bell-type membership functions. The universes of discourse of the input variables are \(AT_{1} \in [0, 50], AT_{2} \in [0, 142] (mg)\), \(RF \in [16.67, 83.33] (rev/s)\), and the universe of discourse of the output variable \(\text{ULE} \in [0.013, 0.478] (\mu m)\). The main training parameters were 200 iterations, step size 0.001, and error 10⁻². The rule base consists of 27 rules. For instance, the first rule is defined by:

\[
R1: \quad \text{IF } (AT_{1} \text{ is Low}) \quad \text{AND } (AT_{2} \text{ is Low}) \quad \text{AND } (RF \text{ is Low}) \\
\quad \text{THEN } (UL_{M} \text{ is } UL_{R1})
\]
The output of each rule is a linear combination of the three inputs, and the model output is the sum of all the rules:

\[ U_{E_i}^{R_i} = a_i \cdot AT_1 + b_i \cdot AT_2 + c_i \cdot RF + d_i \]

(20)

where \( a_i, b_i, c_i, d_i \) are the coefficients obtained by the training corresponding to the \( i \)-th rule, and \( U_{E_i}^{R_i} \) is the output of the \( i \)-th rule.

Five performance indices were considered in the experimental study – the sum of squared errors \( SSE \), the sum of average errors \( SAE \), the mean square error \( MSE \), the mean absolute error \( MAE \), and the maximum error. The results are listed in Table II.

The first three rows of Table II present the behavior of the hybrid incremental model with and without normalization. It is interesting that HIM without normalization has better accuracy and better performance than standard normalization and diagonal normalization, the two of which have results that are very similar to each other. Mahalanobis normalization provides the best results of the three normalization methods considered, reducing the \( MAE \) by about 12% with respect to ANFIS and by a factor of about 5 with respect to the regression model.

### Table II. Comparison of the normalization strategies and modeling techniques considered in this study.

<table>
<thead>
<tr>
<th>Algorithm/Error Criterion (%)</th>
<th>( SSE ) ( x 10^4 )</th>
<th>( SAE )</th>
<th>( MSE )</th>
<th>( MAE )</th>
<th>( Max Error )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without normalization</td>
<td>1.14</td>
<td>364.64</td>
<td>20.91</td>
<td>14.02</td>
<td>63.22</td>
</tr>
<tr>
<td>Standard score</td>
<td>1.18</td>
<td>384.64</td>
<td>21.31</td>
<td>14.79</td>
<td>67.26</td>
</tr>
<tr>
<td>Mahalanobis</td>
<td>0.69</td>
<td>321.48</td>
<td>16.40</td>
<td>12.36</td>
<td>40.57</td>
</tr>
<tr>
<td>ANFIS</td>
<td>0.83</td>
<td>371.92</td>
<td>17.89</td>
<td>14.30</td>
<td>44.67</td>
</tr>
<tr>
<td>Regression</td>
<td>64.3</td>
<td>2013.6</td>
<td>157.25</td>
<td>77.44</td>
<td>93.48</td>
</tr>
</tbody>
</table>

Figure 5 shows the behavior of the models considered in this study relative to the eccentricity measured on the experimental platform described in Sec. IV. The results of the three models considered in this study (ANFIS, nonlinear regression, and HIM) are also presented in Table II. The HIM outperformed both the ANFIS and the regression models. With the Mahalanobis normalization, it improved the \( MAE \) by about 12% with respect to ANFIS and by a factor of about 5 with respect to the regression model.

### V. CONCLUDING REMARKS.

This paper presents the design of a model-based procedure for detecting spindle imbalance based on a hybrid incremental modeling strategy. The influence of data normalization on the modeling accuracy is also analyzed. The comparative study of three different model-based strategies –ANFIS, nonlinear regression, and hybrid incremental modeling– in order to evaluate their behavior and suitability corroborates the advantage of using hybrid incremental models.

Future works from technical viewpoint are focused on a fast, simple, non-intrusive procedure for imbalance detection, focused on reducing the deterioration of the performance of rotational devices due to vibrations.

### REFERENCES


