ON-LINE DETECTION AND IDENTIFICATION OF FAULTS AND ABNORMALITIES IN SENSORS FOR ULTRA PRECISION PROCESS MONITORING

(**Presented at m^{n}f2013, 1st National Conference on Micro and Nano Fabrication, January 21-23, 2013, CMTI, Bangalore*)

[1]Rajesh Kumar C, [2]Shanmugaraj V, [3]Prakash Vinod and [4]Shashikumar PV

[1]Graduate Engineer, [2]Scientist-E, [3]Scientist-E, [4]Joint Director,
Central Manufacturing Technology Institute, Bangalore
E-mail: rajesh.c@live.com

Abstract: *In Ultra precision machining best results are obtained with on-line monitoring and adaptive control of various process parameters during machining. For a reliable on-line process monitoring and error compensation system, it is necessary to have accurate sensor readings. However, sometimes sensors may become faulty and due to failure it gives erroneous or constant values throughout the process. The problem of sensor validation is therefore a critical part of effective process monitoring. The objective of this study is to develop a sensor fault detection module which will be useful for different error compensation /diagnostic techniques needed for the ultra precision machine. A procedure based on Principal Component Analysis (PCA) is developed, which enables to perform detection and identification of sensor failures. PCA is a data driven modelling that transforms a set of correlated variables into a smaller set of new variables (principal components) that are uncorrelated and retain most of the original information. This new index is proposed in order to detect simple and multiple faults affecting the process and diagnose abnormalities in the original system in a robust way. The PCA model maps the sensor variables into a lower dimensional space and tracks their behaviour using Hotteling T^2 and Q statistics.*

Keywords: *PCA, Scores, Faulty Sensor, HottelingT^2, Artificial Drifts, Q Statistics*

1. INTRODUCTION

Ultra precision machining is highly sensitive to the environmental conditions and operating parameters. To attain sub-micron accuracy levels, these critical parameters need to be monitored time to time. Abrupt changes in these parameters will affect the accuracy of the machining. So various sensors are integrated to the system to monitor the changes in these parameters and compensate the abrupt changes through intelligent modules. Hence the accuracy of the error compensation modules relies heavily on the output of these sensors. Any failure in these sensors leads to erroneous results and affects the accuracy. This work is to develop a procedure, which enables detection and identification of sensor failure. Importance of this module lies in the fact that detecting of the fault can be done while the system is still running, so we can avoid abnormal events and losses, thus avoiding major system breakdowns [1] [3].

An unusual measurement is often caused by a major sensor failure. A faulty sensor usually breaks down the normal correlation with the remaining sensors [2]. This can be due to sensor failure, broken wires, lost contact with the surface, etc. in which case the reading shown by the sensor is not related to the value of the measured physical parameter [4] [6].

If the failure is due to external factors like broken wire or lost contact it can be checked and rectified but if it is due to a sensor failure, it can be detected and identified only by using some fault detection techniques. Generally the sensor failure can be classified into 4 major categories. They are, Bias, complete failure, Drift and Precision degradation[1]. A complete failure can be easily identified since it will give either a constant value throughout the operation. But identifying other types of failure is a difficult task, because it will give some value which may be varying with time to time.

2. FAILURE DETECTION

Once a fault has occurred in a system, the immediate step will be detecting the fault. These faults can be detected using different methods[6]. Fault detection methods can be classified into following:

1. Data-Driven Methods
2. Model-Based Methods

2.1 Data-Driven Methods

Data driven methods use the availability of large amount of data from many process. This data can be transformed and used as a knowledge bank using many different ways. This process is also called as feature extraction. This feature extraction can be done using statistical or non-statistical methods [5]. The important class of non-statistical is neural networks. Statistical feature extraction methods mainly contain Principal Component Analysis (PCA), Fisher Discriminant Analysis (FDA), Partial Least Squares (PLS) and statistical pattern classifiers [6]. The only drawback of data-driven methods is that it largely depends on the quality and quantity of the process data.

2.2 Model-Based Methods

There are different model based fault detection methods. Fault detection using input and output measurements of the system is the basic one among them. In some cases we measure only the output signal, in such a case model based methods such as spectral analysis and band-pass filters are used for fault detection. [6] Among these methods, the most frequently used techniques are parameter estimation and observer based methods. Fault detection using these methods is done by comparing the system's measured variables with the information obtained from the system's mathematical model [6].

It is very difficult to know the exact process model of the system, in such cases data driven techniques can be used for fault detection and diagnosis [1]. Many process control systems have large data archives of the system parameters like pressure, temperature, flow-rate, etc. during normal and faulty operating conditions. This large data can be used to build a statistical model, which can be used to detect the faults [4].

3. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is one of the best data driven technique used to detect the sensor failure and abnormal variation in the sensor measurements. This module can be used to identify the faulty sensor with an abnormal condition is detected. It uses the residual of each sensor at every sample to identify the sensors related to a detected fault. PCA is a technique developed to reduce data dimensionality by extrapolating correlated variables in sets of new uncorrelated variables, keeping variance of the original data. It determines a set of orthogonal vectors called loading vectors, ordered by the amount of variance explained in the loading vectors direction. Given a training set of n observations and m process variables, with mean zero and unit variance, stacked into a matrix X $\in R_{n\times m}$ the loading vectors are of the covariance matrix S[5]

$$S = \frac{1}{N-1} X^T X = V \Lambda V^T \quad (1)$$

Where the diagonal matrix $\Lambda \in R_{mxm}$ contains the nonnegative eigenvalues of decreasing magnitude $(\lambda_1 \geq \lambda_2 \geq \cdots \lambda_m \geq 0)$ and the i^{th} eigenvalue equals the square of the i^{th} singular value (i.e. $\lambda i = \sigma^2 i$) PCA can handle high dimensional, noisy and correlated data by projecting the data onto a lower dimensional subspace which contains most of the variance of original data. Selecting the column of the loading matrix $P \in R_{mxa}$ to correspond to the loading vector associated with a singular values, the projections of the observations in X into the lower-dimensional space are contained in the score matrix T [5]

$$\hat{X} = TP^T = \sum_{i=1}^{a} T_i P_i^T \quad (2)$$

Where T_i is a score vector (orthogonal) which contains information about relationship between samples and pi is a loading vector (orthonormal) which contains information about relationship between variables. Projection into principal components space reduces the original set of variables to a latent variable.Generally, the a principal components should explain the variability of a process through its data X, therefore the difference between X and $\hat{X}$ is the residual matrix E that captures the variations associated with n−a singular values [5]

$$X = TP^T + E = \sum_{i=1}^{a} t_i p_i^T + E = \hat{X} + E \tag{3}$$

It is very important to choose the number of principal components, a, because TP^T represents the principal sources of variability in the process, and E represents the variability corresponding to process noise. There are different techniques to determine the a principal components: Cumulative Percent Variance screen test, parallel analysis and prediction residual sum of squares statistic. Cumulative Percent Variance is used to elaborate this technique [6].

Cumulative Percent Variance (CPV) approach is used to find the number of PC. It's the most commonly used technique to select the number of PC to be used. The cumulative sums of all PC are calculated and by setting the control limit the number of PC to be used is calculated. It is basically selected based on the process [6].

$$PC(i) = \frac{\lambda_i}{\sum_{j=1}^{m} \lambda_j} \tag{4}$$

When the amount of PC's (a) is selected, only the eigenvalues $\lambda_1, \lambda_2 \ldots \lambda_a$ and Eigenvectors $e_1, e_2, \ldots, e_a$ related to these are used.

3.1 Process Statistics with PCA

On-line monitoring of measurement variables can be carry out with the help of the Hotelling's T^2 and Q statistics, this last also known as the squared prediction error (SPE). These two statistics can be used to detect faults for multivariate process data. The Hotelling's T^2 statistic measures the variation within the PCA model while the Q statistic measures the variation outside of the PCA model [2] as given in Fig. 2.

The T^2 statistic for the lower-dimensional space can be calculated for each new observation x by [5]

$$T^2 = x^T P (\sum a)^{-1} P^T x \tag{5}$$

Where Σa contains the non-negative real eigenvalues corresponding to the a principal components and P contains the loading vectors associated with the a columns of V.

The upper confidence limit for T^2 is obtained using the F-distribution [5]

$$T^2_{a,n,\alpha} = \frac{a(n-1)}{n-a} F_{a,n-a,\alpha} \tag{6}$$

Where n is the number of samples in the data, a is the number of principal components and is the level of significance.This statistic can be interpreted as measuring the systematic variations of the process, and a violation of the threshold would indicate that the systematic variations are out of control the portion of the measurement space corresponding to the lowest m−a eigenvalues can be monitored using the squared prediction error (SPE) or Q statistic [5]

$$Q = x^T (I - PP^T) x \tag{7}$$

Where, I is the identity matrix.
The upper confidence limit for the Q can be computed from its approximate distribution [5]

$$Q_a = \theta_1 \left(\frac{h_0 c_\alpha \sqrt{\theta_2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{\frac{1}{h_0}} \tag{8}$$

With:

$$\theta_i = \sum_{j=a+1}^{m} \lambda_j^i \qquad h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$$

Where Cα is the value of the normal distribution with α the level of significance. The Q statistic does not suffer from over-sensitivity to inaccuracies in the smaller singular values and it is associated with noise measurements. A violation of the threshold would indicate that the random noise has significantly changed, or a usual event has occurred that had produced a change in the covariance structure of the model.The individual scores of the observations are compared with the Q value; the observations which are crossing the limits are called outliers. After detecting the outlier due to extreme Q value, inputs that contribute to the extreme statistical value are investigated. Formula for calculating the contributes [6]

$$C = \frac{T(i)}{\lambda_i} v(i,j) \overline{X}_j \tag{9}$$

$$CONT_j = \sum_{i=1}^{m} C_{i,j} \tag{10}$$

C (i, j) = the contribution of each variable to the observation

3.2 Algorithm for Sensor Fault Detection using PCA

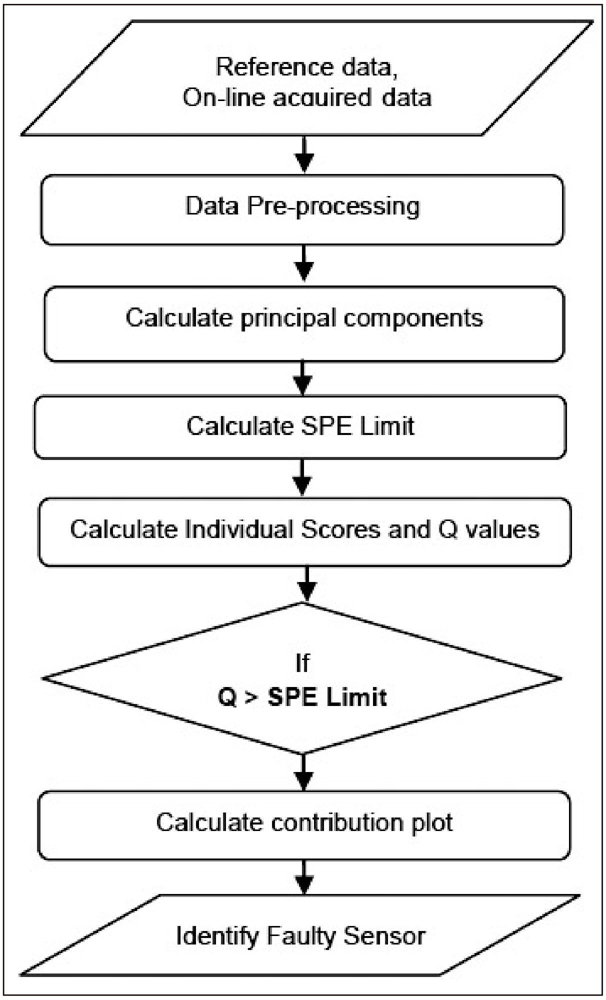

Fig 1. Algorithm of the Sensor Fault Detection Model

4. MODULE DESCRIPTION

A fault detection module has been built and developed based on the model described in the schematic diagram Fig.2. The system is composed of three stages, which are input, processing and output. In input stage, the module collects input data sets. The reference data is retrieved from the data archive and the on-line data is acquired from the sensor. Reference data is collected from the healthy data acquired previously. To ensure that the data is free from abnormalities the reference data is cross checked.

The processing stage is of two steps. In the first step the reference data set is scaled along with the on-line acquired data to make the mean of data set as zero and unit standard deviation. In second step, from the on-line data individual scores are calculated using transform matrix and Q statistic are calculated using residuals. The confidence limits for the scores are calculated using Squared Prediction Error method. In output the individual scores are compared with the SPE limit, the abnormal observations which are above the limit are considered as outliers. In the outliers the sensors which are responsible for the abnormality are identified using contribution plots derived from HottelingT^2 and Q statistics.

4.1 Software Description

The ultra precision truing machine is an open architecture CNC controller with C, C++ and Labview interfaces. Labview software is used to design and develop the fault detection module, as the data is acquired using the NI data acquisition card.

5. EXPERIMENTAL SETUP

Experiments for fault detection have been conducted on an ultra precision turning machine at CMTI using temperature sensors, which are involved in the error compensation system. To improve the accuracy of the compensation system, sensors are mounted on critical parts of the machine like spindle, spindle motor, linear stages, tool post and linear motors of the machine, where change in process is susceptible to the change in temperature. Two set of experiments are carried out for testing the module .i.e. Off-line mode and On-line mode.

5.1 Off-line mode

The input data for Off-line mode experiments are acquired during machining process. Fig.4. shows

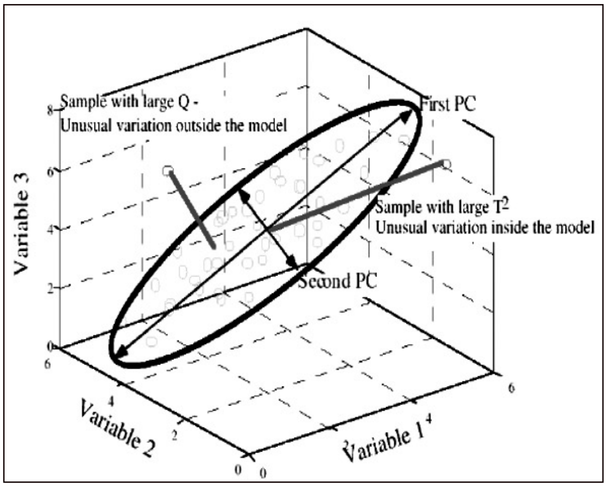

Fig 2. Principal Component model staying on a plane, Showing T^2 and Q Outliers

the temperature plots of the acquired data for a period of 30 minutes. With the various sets of operating conditions the temperature data are acquired. From the sensor output, it is observed that the changes in the parameters are not making any apparent deviation in the temperature. Hence for all operating parameter given below, a common set of reference data is used.

Table1: Operating Parameters of the Conducted Experiments

PARAMETER	RANGE
Spindle Speed	1000-1500 rpm
Depth of cut	5-10 micrometers
Feed Rate	5 mm/min
Room temperature	20 ±0.5 °C

As the variations in temperature between the samples are negligible, data are acquired with optimum sampling rate which gives enough information about the process.

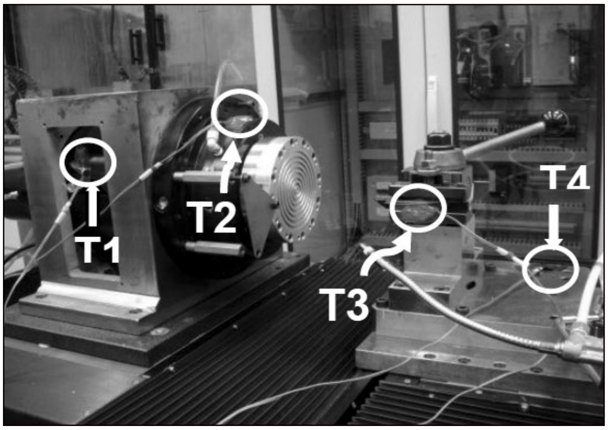

Fig 3. Temperature Sensors Mounted at Various Elements of the Machine (T1, T2, T3 and T4)

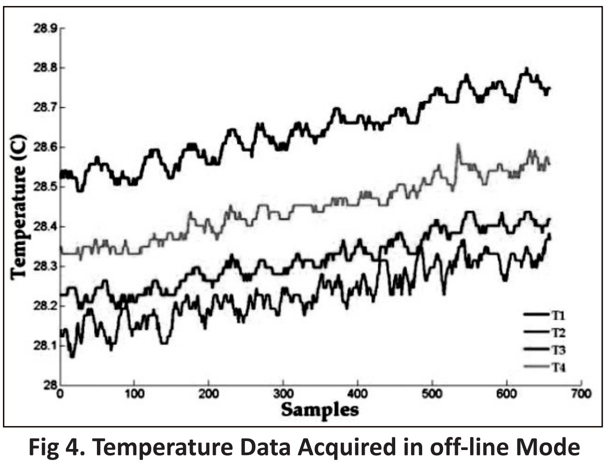

Fig 4. Temperature Data Acquired in off-line Mode

5.2 On-Line Mode

During on-line operation of process, The PCA model is running on-line data. The advantage in the On-line module is that the faults in the sensor are identified during the operation, which helps to avoid inaccurate compensation, thus improves the accuracy of the system.

During machining process, data are acquired and compared with reference data derived in Off-line mode. If outliers are present in any observation, contribution plot are calculated for those abnormal observations. The variable which is responsible for the abnormal Q statistic will contribute more than the other variables of the particular observation.

The output screen of the associated software module (Fig. 7) developed using “NI-Labview” will display the error message as “Malfunction process” on detection of abnormalities in the process. The sensor column in the output screen shows the faulty sensor and the outlier column shows the number of outlier observations present in the data set and contribution plot is presented in graphical representation.

6. RESULTS AND DISCUSSION

To validate the Off-line mode ability to detect and identify the drifting sensor, artificial drifts was applied in temperature sensor T2. The drift simulates a common problem that affects the process sensor and result abnormality. The simulated drift is a ramp that grows to 0.5°C for a temperature variable. This small drift corresponds to an imperceptible in output profile as shown in the Fig. 5.

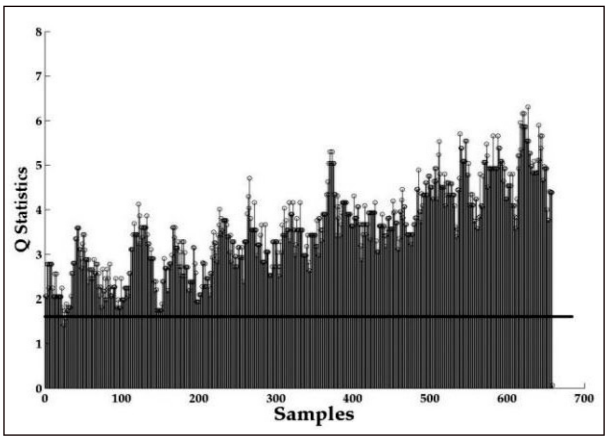

Fig 5. Q Residual Values of Drifted Data

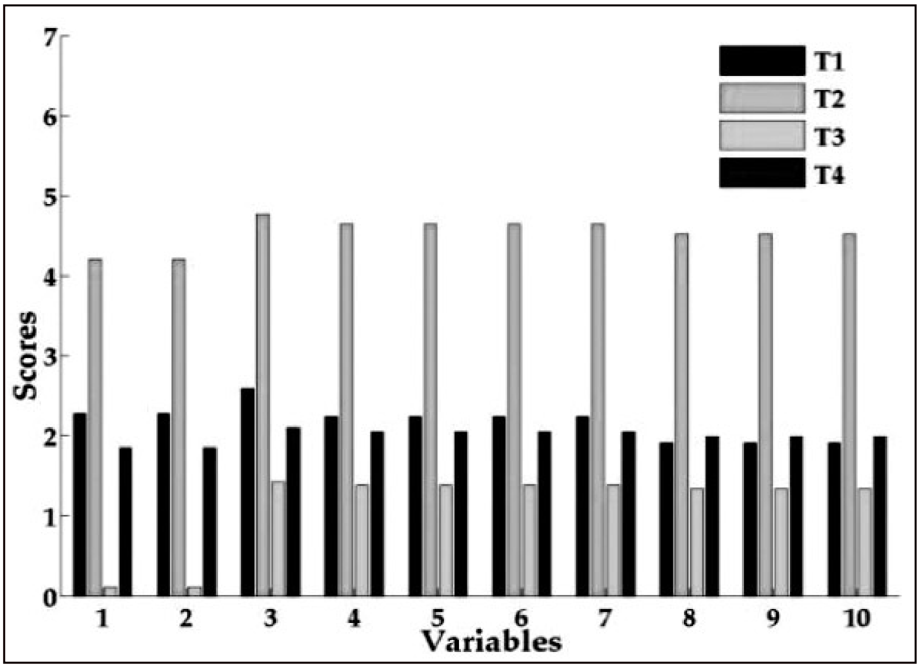

Fig 6. Contributions of Variables to Q Statistics of the Samples from 301 to 310

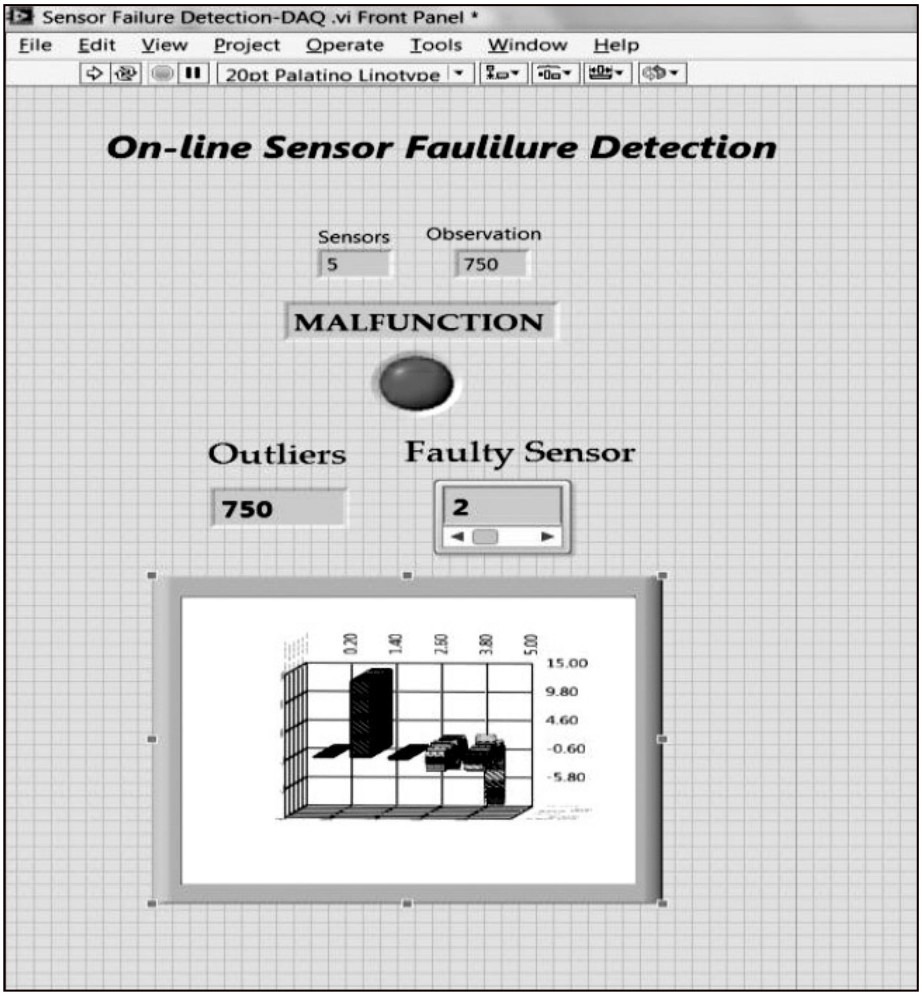

Fig 7. Output Screen of the on-line module

The sensors affected by the drift are determined by the contribution plot. For investigation, Samples from 301-310 are selected. The contribution of each sensor to Q statistics is plotted in Fig.6. From figure it is clear in that in all the observations variable T2 is contributing more to the unusual Q statistic than the other sensors in the observation. This agrees with the fact that the drift was added to the variable T2.

To validate the On-line mode's ability to find malfunctioning of sensor (bias failure), the system is simulated on all the 750 samples and bias is added to the variable 2 since the first sample till the end of this simulation.

This fault is represented by a constant bias of amplitude equal to 6% of the variation domain. As the output screen in Fig. 7 shows all the observations are detected with abnormality and the plot shows the graphical representation of each variables contribution to Q statistics from this it's evident T2 contribute more to the abnormality, hence variable 2 is displayed as faulty variable displayed. This result matches with the fact that bias was added to the variable T2.

7. CONCLUSIONS

In this study, Principal Component Analysis (PCA) based module is presented for detection and identification of faulty sensors used in intelligent ultra precision machining. NI-Labview based software has been also developed for identification of faulty sensors in on-line mode. The performance of the module is evaluated by testing it in Off-line and On-line mode with the ultra precision tuning machine. Artificial drifts and bias were added to the variables and the model has both detected and identified the faulty variables. The results demonstrate the effectiveness of the module for detecting and identifying malfunctioning and complete failure of the sensor in both On-line and Off-line.

Thus the proposed module was determined to be an effective method to monitor the sensors in ultra precision machines and can be used along with the error compensation and process monitoring systems. This module will help to get the best results in Micro/Nano machining.

Sensor fault detection works well as a standalone module. In future it has to be integrated to the controller of the machine, so that it will be directly involved in different error compensation modules. This is observed to be the best module to detect the faults in the sensor with linear data. If the sensors with non-linear data are present in the system, other methods such as non-linear PLS or neural networks may be used.

8. ACKNOWLEDGEMENTS

The authors are grateful to Shri. B.R. Satyan, Director, CMTI for extending support throughout the work. Also, the authors are thankful to Mrs. Usha S, HOD-SVT, Mr. Balashanmugam N, HOD-NMTC, Narendra Reddy T, Scientist-B and Mr. Gopi Krishna S, Scientist-B for their helpful discussions and suggestions throughout this research.

9. REFERENCES

1. Ricardo Dunia,” Identification of Faulty Sensors Using Principal Component Analysis”, AIChE Journal October 1996 Vol. 42, No. 10 2797.
2. Rosani M. L. Penha,” Using Principal Component Analysis Modeling to Monitor Temperature Sensors in a Nuclear Research Reactor”, Centro de Energia Nuclear Institute de Pesquisas Energéticas Nucleares – Ipen São Paulo, SP 05508-900 Brazil.
3. MSRK Boss, “Detection isolation and reconstruction of faulty sensor using Principal Component Analysis”, International Journal of Chemical theory, Vol 12 July 2005 Pg 430-435.
4. Anissa Benaicha, Mohamed Guerfel,” New PCA-Based Methodology For Sensor Fault Detection And Localization”, 8th International Conference of Modeling and Simulation - MOSIM’10 - May 10-12, 2010.
5. Thamara Villegas, “Principal Component Analysis for Fault Detection and Diagnosis. Experience with a pilot plant”, ISBN: 978-960-474-257-8 Advances in Computational Intelligence, Man-Machine Systems and Cybernetics.
6. Vamshi Krishna Kandula, “Fault detection in process control plants using principal component analysis” Louisiana State University and Agricultural and Mechanical College ◘