

# Health Monitoring of Electro-Pneumatic Controlled Systems Using Multivariate Latent Methods: An Experimental Validation

W. M. Hussein, Ahmed Onsy, and Ibrahim El Sherif  
MTC

## ABSTRACT

Electro-Pneumatic systems exhibit highly nonlinear characteristics due to air compressibility, the presence of friction and the nonlinearities of control valves. Monitoring by acquiring the system's transfer function accurately can be difficult for nonlinear systems. This paper outlines a new idea that one can deal with the electro-pneumatic system as a black box, and using a multivariate technique called principal component analysis (PCA) and projection to latent structure discriminant analysis (PLSDA) to provide robust information about the system's condition. The monitoring system has been experimentally validated for an electro-pneumatic printing machine system using vibration, pressure and displacement sensory data integration using PCA-PLSDA algorithm. Experiments were conducted under two pressures for three artificially conditions: normal, throttled, and leaking system. The models were tested under these conditions, and the results showed that the proposed technique can successfully differentiate between these process conditions and can be used to overcome these faulty states using external connection line in a real-time environment.

**KEYWORDS:** Electro-Pneumatic Systems, Condition Based Maintenance, Health Monitoring, Multivariate Analysis, PCA, PLSDA

**CITATION:** Hussein, W., Onsy, A., and El Sherif, I., "Health Monitoring of Electro-Pneumatic Controlled Systems Using Multivariate Latent Methods: An Experimental Validation," *SAE Int. J. Mater. Manf.* 7(1):2014, doi:10.4271/2013-01-9097.

## INTRODUCTION

In recent printing industry, the on-line monitoring of process performance is extremely important for the safety, efficiency and quality printing machines. As the condition of the machine changes due to variations in its system condition, this consequently leads to variations in performance. One of the major systems in printing machines is the electro-pneumatic system, where monitoring the condition of its various parts is crucial to prevent changes in the operating state of the machine. Several methods are currently in use for the on-line monitoring of printing machines, including vibration, pressure, and displacement analysis. Vibration analysis is the most commonly used for machine condition monitoring [1], where vibration occurs in printing machines due to the movements of mechanical parts and air flow variation. Meanwhile pressure and displacement are monitored to track any changes from the design ranges [2, 3]. However, each method has advantages and disadvantages and

combining the information from all three methods should provide a more robust conclusion about the condition of the printing machine. This can be achieved by the application of artificial intelligence or data mining algorithms to fuse sensory data and provide robust information about the system's condition.

Artificial intelligence can be defined as "the science of making machines do things that would require intelligence if done by humans" [4]. To develop a health monitoring system (HMS), the running system condition must be recognized and classified. Researchers have devoted considerable effort to the application of various different soft computing methods to develop HMSs for mechanical geared transmission systems [5, 6] and electro-pneumatic systems [7, 8], and the methods used include neural networks, fuzzy or neuro-fuzzy logic, and multivariate statistical approaches. These methods can provide important tools for the field of intelligent monitoring which can learn, adapt, and make decisions concerning the system they are in charge of [4]. For example, Onsy et al. [5,

[6, 24] used prediction models and sensor fusion algorithms to predict the failure of a geared transmission system based on vibration, acoustic emission and on-line debris analysis features; including adaptive fuzzy logic algorithms [5,6] and model-based parametric method algorithms [24]. Also, Uppal et al. [7] introduced a fault diagnosis system for an electro-pneumatic valve actuator using a fuzzy system, while Sang et al. [8] monitored the process of an electro-pneumatic valve actuator using the multivariate statistical technique of kernel principal component analysis.

Artificial neural networks (ANNs) in combination with other soft computing methods have been the most predominant to date [9, 10]. However, ANN techniques are characterized as black box approaches which model the relationship between different sensory data and a desired response without giving any information about the process. Also some other approaches such as fuzzy logic are sensitive to missing data due to the static and dynamic characteristics of sensors. In addition, issues associated with collinearity and dimensionality needs to be specifically addressed in these techniques. Solving many of these limitations can be addressed through the use of multivariate statistical analysis, which is a data-driven modelling approach [11]. The most fundamental difference between multivariate statistical analysis and the other methods mentioned previously is that the model is now based on non-causal empirical correlations extracted from normal process operating data when common cause variation exists [12]. The simplicity of this approach is that there is no need for a fundamental model of the system and only data from normal operation needs to be used, which is generally available in some form for most machines. Among the approaches used in multivariate analysis are: a projection method called principal component analysis (PCA) [13, 14] and projection to latent structure (PLS) [15]. Many applications of these two techniques have been successfully applied in other fields of process monitoring [9, 16, 17, 18, 19, 20, 21].

Principal component analysis (PCA) is the most widely used data-driven technique for process monitoring which has been studied and applied to industrial systems over the past decade. PCA is an optimal dimensionality reduction technique in terms of capturing the variance of the data, and it accounts for correlations among variables. The lower-dimensional representations of the data produced by PCA can improve the proficiency of detecting and diagnosing faults using multivariate statistics [22].

This paper outlines the use of the multivariate projection approach to develop a framework to monitor the printing machine under study, by integrating the data collected from different accelerometers, pressure and displacement transducers fixed at different machine positions under different operating conditions. Models based on PCA and PLS-DA are developed so that the machine's condition can be monitored using the latent space of these models.

## BASIC THEORY OF MULTIVARIATE PCA AND PLS-DA

### Multivariate PCA Theory

Principal component analysis decomposes the variance and covariance structure of a data matrix by defining linear combinations of the columns in the original matrix. Moreover, PCA extracts information from data sets by computing a smaller data set and other summary information that adequately captures most of the underlying features of the larger data set. The point that needs to be stressed is that the data can be reduced to a size which is more manageable but contains the features that are often of interest. PCA extracts a score matrix,  $T$ , and a loading matrix,  $P$ , from  $X$ . These matrices have the following dimensions [14, 23]:

$$X: N \times K \quad T: N \times A \quad P: K \times A \quad (1)$$

The first column of  $T$  and  $P$  are called by their shorter forms,  $t_1$  and  $p_1$  respectively. The lower case letters indicate that these are vectors, upper case letters indicate matrices. Extracting only one principal component (that is, a single score vector and loading vector) gives:

$$X = t_1 p_1^T + E_1 \quad (2)$$

Extracting a second principal component:

$$X = t_1 p_1^T + t_2 p_2^T + E_2 \quad (3)$$

Principal components are extracted in same manner and then group these score and loading vectors to form matrices  $T$  and  $P$ :

$$T = [t_1 \ t_2 \ \dots \ t_A] \quad \text{and} \quad P = [p_1 \ p_2 \ \dots \ p_A] \quad (4)$$

The eigenvectors of the real symmetric matrix  $X^T X$  give us exactly the loading matrix  $P$ . This is the loading matrix obtained by extracting all principal components.  $P$  is used to compute  $T$ :

$$T P^T = X \quad T P^T P = X P \quad T = X P \quad (5)$$

A more in-depth discussion, which also highlights some geometric concepts of PCA, can be found in [13]. The algorithm used to calculate the PCA is the NIPALS (Nonlinear Iterative Partial Least Squares) algorithm [14, 23].

### Multivariate PLS-DA Theory

The projection to latent structure discriminant analysis PLS-DA aims to find a line or a plane that separates classes of observations based on their values. The PLS-DA model as discussed elsewhere [15, 17] is based on building classes or clusters using knowledge about existing groups (supervised

learning). The model is established using a set of experimental runs (training set). PLS-DA works on two matrices: the X-matrix as input and Y-matrix as output. The X-matrix contains all the sensory data with all the known groups (i.e. normal and faulty states). The Y-matrix contains dummy variables in columns equal to the number of already known classes of data.

In the Y-matrix we assign to each observation from the same group a code number which is constant for the entire group. For example, in cases where we have five groups to cluster, the first group (e.g. normal condition data) is assigned with a code value of “1, 0, 0” for each row (observation) of the same group. Depending on the number of samples or observations from each group, there will be an equal number of code values with the same value “1, 0, 0”. The second group will take another code of “0, 1, 0”, and so on for other groups (i.e. 0, 0, 1 etc...). In this way the approach tries to model X-Y variables by finding the latent directions that best separates the known groups. In other words, sometimes factors are introduced that describe the data from different groups. So, we must assign some distinct levels to these factors to help discriminate between different groups. We call these factors qualitative or dummy variables. Mathematically this can be illustrated by adding new variables “G” to the model with some regression coefficients say “c”, and set  $G = 1$  for data from one group and  $G = 0$  for data from another group in the following way: . Suppose that represents the fitted model, so when  $G = 0$ , the data from the first group is highlighted: and by putting  $G = 1$ , the data from the second group is highlighted: where c estimates the difference in levels between the responses of two groups of data. In this way the data from different groups will be clustered depending on their regression coefficient values [15, 18].

## TEST RIG SETUP

The electro-pneumatic system monitored in this study is the Printmaster QM- 46 printing machine, and the health of the machine can be predicted using PAC-PLC prior to faults appearing in order to save the effort and money spent using traditional techniques for machine monitoring. The pneumatic system of this printing machine (shown in [Figure 1](#)) consists of an electric motor with compressor, electro-pneumatic directional control valve and cylinder. This cylinder controls the connect pressure between the plate to the blanket drums of the printing machine.

The schematic of the experimental setup used to validate the proposed method is shown in [Figure 2](#). The test rig is designed to perform the integration of several components, including mechanical electronics and a computer interface with high computational capacity software. The pneumatic unit consists of a compressor, 5/2 directional control valve with on-off solenoid, and a cylinder with a stroke length of 52.5 mm. The sensors used to monitor the Printmaster QM46 pneumatic system are listed in the [Table 1](#).

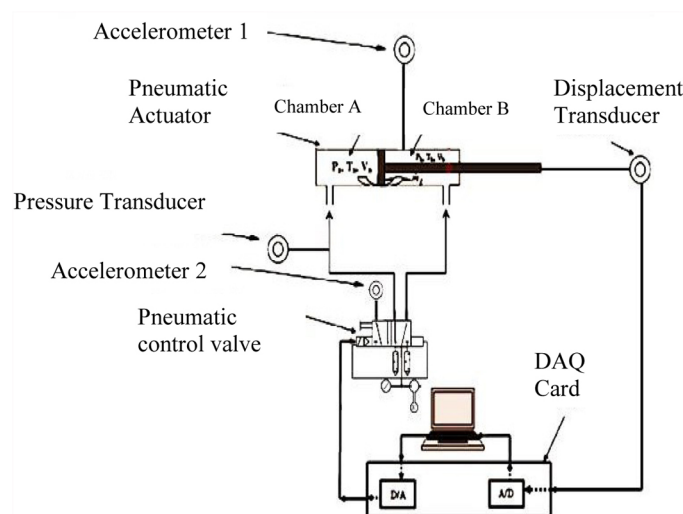


Figure 1. Printmaster QM46 pneumatic system layout

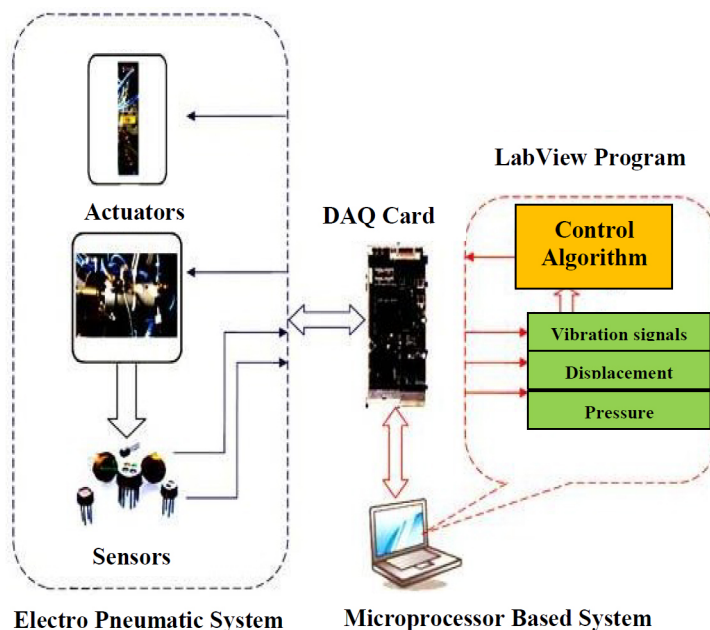


Figure 2. Experimental test rig setup.

Table 1. Sensors Specification

| Transducer                                      | Specification                                                                  |
|-------------------------------------------------|--------------------------------------------------------------------------------|
| Linear variable differential transformer (LVDT) | Displacement range: 0 - 100 mm, Full Range Output: 9.5 V, Accuracy: +/- 0.05%. |
| Pressure                                        | Pressure range: 0 to 10 bars, Full Range Output: 10 V                          |
| ICP Accelerometer 1                             | Model: ACC103, Sensitivity: 10 mV/g                                            |
| ICP Accelerometer 2                             |                                                                                |

All signals were sent to a computer based DAQ system, via a National Instruments (NI) DAQ card model PCI-6036E. The DAQ card had 16 analog inputs, two analog outputs, a sampling rate of 200 KS/s, and an input voltage range of  $\pm 10$  V. The output signal from the DAQ system is used to control the operation on the Printmaster QM46 pneumatic system valve, the change of the input pressure is (4 - 6.4) bar at the two states (normal and faulty). A LabVIEW program is developed for the purpose of system operation and data display and logging.

## HEALTH MONITORING SYSTEM MODEL BUILDING

The multivariate PCA model is a linear model given in matrix form by:  $X = TP^T + E_A$  [14, 23]. However, there exist non-linear PCA versions even by augmenting the original matrix with the non-linear factors or by building a non-linear relation between score factors [14]. To build the model using process variables, let  $X$  include variables containing the information in the sensory data. Figure 3 illustrates how the score plot is built for a simple case of 3-variables and 2-scores. After determining the direction of maximum variation by iterative steps and to get the second orthogonal direction in the same way after subtracting the first component, we rotate the new plane determined by the new score variables  $t_1$  and  $t_2$  and then monitor the movement of the process variables in the reduced dimensional space during printing.

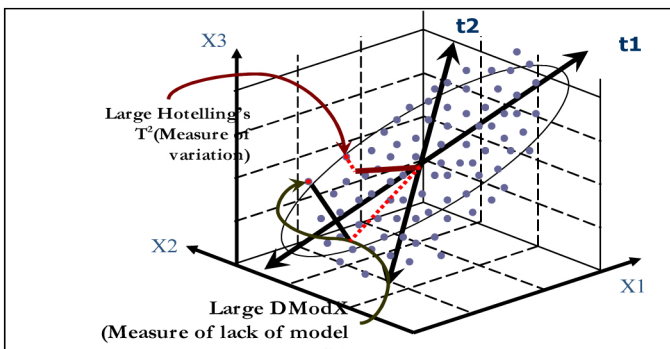


Figure 3. Establishing score plot for three variables

A total of 4 variables were used to generate the input matrix "X". These variables represent the root mean square of the time domain records of the sensory signals (valve vibration, actuator vibration, actuator displacement and system pressure); a data set sample is provided in Table A-1 (Appendix A). All of the models were built using SIMCA-P code developed by Umetrics based on PCA [11]. The PCA model is based on building classes or clusters using existing information inside the data (unsupervised learning). The model is established using a set of experimental runs called the training set that represents the normal operation of the printing machine. Finally, a regression is conducted on the matrix  $X$  based on the NIPALS algorithm; a sample of processed (scaled) data set is provided in Table A-2

(Appendix A). Appendix B provides a simple example code to develop PCA model using Matlab program to illustrate the NIPALS algorithm.

## EXPERIMENTAL RESULTS AND ANALYSIS

To explore the data collected a first model using PCA was built using all the dataset by cross-validation, now obtaining a two component model, with:  $R^2X$  (cum): 0.81,  $Q^2$  (cum) : -0.105. It is seen that the model factor  $Q^2_{cum}$ , is small, but it is still better to first get a general look at the whole dataset without removing any variable or observation. Figure 4 illustrates a scatter plot of the two score vectors ( $t_1$  and  $t_2$ ) of the PCA model using a set of 10 observations. It provides a clear vision of the dispersion of the data, with the normal process data being discriminated from the faulty data along the first two components. This clustering is shown by a straight line separating the different conditions (normal operation of the electro-pneumatic system and different levels of leaking and throttled states of the system).

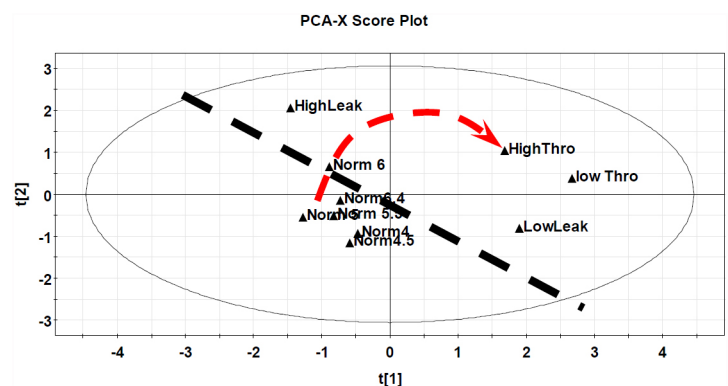


Figure 4. A scatter plot of the two score vectors ( $t_1$  and  $t_2$ ) of the PCA model

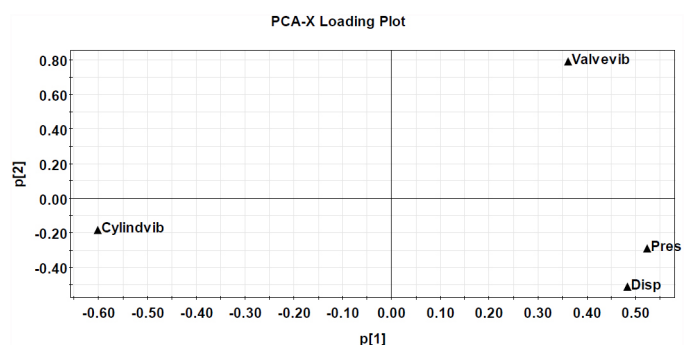


Figure 5. PCA-X loading plot

More analysis is conducted by examining the loading plot which shows the relationships between the different variables. Figure 5 shows the loading plot, clarifying the relations between the 4 variables. In addition, the scores and loading plots are superimposed; this means that variables