Proceedings of the World Congress on Engineering and Computer Science 2007
WCECS 2007, October 24-26, 2007, San Francisco, USA

# A Principal Component Approach in Diagnosing poor Control loop performance

Haslinda Zabiri, Tran Diep Thanh Thao

*Abstract*— **Principal component analysis, both linear and nonlinear, are used to identify and remove correlations among process variables as an aid to dimensionality reduction, visualization, and exploratory data analysis. While PCA ascertains only linear correlations between variables, NLPCA reveals both linear and nonlinear correlations, without restriction on the character of the nonlinearities present in the data. In this paper, the use of PCA and NLPCA are investigated and compared for nonlinearity detection in regulated systems using routine operating data. Results from simulated and industrial data used in this study clearly show that NLPCA performance supersedes that of PCA in identifying and detecting nonlinearity in poor performing control loops.**

*Index terms*— **Poor control loop diagnosis, PCA, NLPCA**

## I. Introduction

Controller performance monitoring field is one of the main areas that has received much spotlight in the engineering research literature. On the other hand, the diagnosis of poor control loop performance remains an open area [1]. In a poorly performing control loop, oscillations in the process variables may arise due to several reasons including poorly tuned controllers, presence of external oscillatory disturbances, process and/or actuator nonlinearities. The presence of nonlinearities in control loop may have detrimental effect on the controller performance, resulting in oscillations of the process variables, shorten the life of the control valve due to wear and tear, may upset process stability, and poor quality end-products. Nonlinearities in control loop may be present in the process itself or in control valves. This paper is concerned with control valve nonlinearities, which are normally due to faults such as stiction, backlash, saturation, deadzone, ruptured diaphragm, and/or corroded or eroded valve seats [1]. Among these, stiction is one of the common and long-standing problems in process industries [2].

Principal Component Analysis (PCA) is one of the classical multivariate statistical methods within the class of linear methods [3]. PCA essentially detects and characterize optimal lower-dimensional linear structure in a multivariate dataset. It has been widely used in various areas of multivariate analysis, including data validation and fault detection, quality control, correlation and prediction, data visualization [4] as well as oceanography and meteorology [5,6]. PCA uncovers the lower-dimensional hyperplane that optimally characterizes the data, such that the sum of squares of orthogonal deviations of the original data points from the hyperplane is minimized. Due to the linearity assumption underlying the method, PCA is an optimal feature extraction algorithm only if the underlying structure of the data is Gaussian. If the data contains nonlinear lower-dimensional structure, PCA will not be able to work satisfactorily.

A natural nonlinear generalization of PCA for feature extraction problem has been introduced in the early 1990s by [4], which has been called as nonlinear principal component analysis (NLPCA). This neural-network based generalization of PCA adopted the same criterion of optimality as PCA, however the main difference between PCA and NLPCA is that the latter allows nonlinear mapping between the original and the reduced lower dimensional spaces. NLPCA has been shown to perform more satisfactorily than PCA in characterizing the lower-dimensional nonlinear structure of the data [4,5,6]. The applications of NLPCA can be found in chemical engineering, psychology, image compression and climate data, oceanographic data, environmental systems, periodic and wave phenomena [6,7,8,9].

Although both PCA and NLPCA have been widely used in the above mentioned areas, to the author knowledge, they have not been used in diagnosing poor control loop performance. This paper investigates these methods and demonstrates the potential of using NLPCA in controller performance analysis and diagnosis.

This paper is outline as follows: Section II describes the theories of PCA and NLPCA. In Section III, simulation examples to diagnose the causes of poor performance comparison between PCA and NLPCA are presented. Industrial case study is described in Section IV, and finally some conclusions are drawn at the end of the paper.

## II. Linear and Nonlinear PCA

### A. Theory of PCA

Principal component analysis (PCA) has been called one of the most valuable results from applied linear algebra. PCA is used abundantly in all forms of analysis because it is a simple, non-parametric method of extracting relevant information from confusing data sets. With minimal effort PCA provides a roadmap to reduce a complex data set to a lower dimension.

Manuscript received May 31, 2007. This work was supported fully by Universiti Teknologi PETRONAS.

Haslinda Zabiri is with Universiti Teknologi PETRONAS, Bandar Sri Iskandar, Tronoh 31750 Perak, MALAYSIA. (corresponding author phone: 605-368-7625; fax:605-366-6176; e-mail: haslindazabiri@petronas.com.my).

Tran Diep Thanh Thao is with Universiti Teknologi PETRONAS, Bandar Sri Iskandar, Tronoh 31750 Perak, MALAYSIA (e-mail: tran_diethanh@utp.edu.my).

This is to reveal the sometimes hidden, simplified dynamics that often underlie it.

Principal Component Analysis (PCA) identifies patterns in data, and expresses the data in such a way as to highlight their similarities and differences. Once these patterns have been found, the data can be compressed, *i.e.,* number of dimensions reduced, without much loss of information [3].

Assume that a data matrix $\underline{\underline{X}}$ has *n* number of observations and *m* number of variables as shown in Fig. 1.

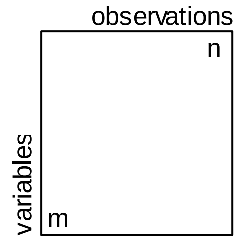



Fig. 1. Notation used in PCA. Extracted from [3].

The observations (columns) in Fig. 1 can be process time points of a continuous process or analytical samples, chemical compounds or reactions, etc. The variables (rows) may be measurements from sensors and instruments in a process (temperatures, flows, pressures, etc.), etc.

PCA allows a linear mapping of data from $\Re^m$ to $\Re^p$, with $p < m$. The optimal transformation of $\underline{\underline{X}}$ via PCA into two matrices is shown in (1).

$$\underline{\underline{X}} = \underline{\underline{T}}\,\underline{\underline{P}}^T + \underline{\underline{e}} \tag{1}$$

$\underline{\underline{T}}$ is called the *score* matrix with dimension $nxp$, where $p$ is the number of principal components $(p < m)$ of $\underline{\underline{X}}$, and $\underline{\underline{P}}$ is the *loadings* matrix with dimensions of $mxp$. The Euclidean norm of the residuals matrix $\left\|\underline{\underline{e}}\right\|$ must be minimized for the given number of principal components for the optimality condition to be satisfied and this is achieved if the columns of $\underline{\underline{P}}$ are the eigenvectors corresponding to the $p$ largest eigenvalues of the covariance matrix of $\underline{\underline{X}}$.

If $\underline{\underline{P}}^T \underline{\underline{P}} = \underline{\underline{I}}$, the linear mapping of PCA is given by (2).

$$\underline{T} = \underline{X}\,\underline{\underline{P}} \tag{2}$$

Where $\underline{X}$ represents a row of $\underline{\underline{X}}$, a single data vector, and $\underline{T}$ represents the corresponding row of $\underline{\underline{T}}$ (the coordinates of $\underline{X}$ is the reduced *p*-dimensional variable space). The loadings $\underline{\underline{P}}$ are the coefficients for the linear transformation, and essentially define the orientation of the principal component plane with respect to the original *m*-variables.

The information lost in this mapping can be assessed by reconstruction of the measurement vector by reversing the projection back to $\Re^m$:

$$\underline{X}' = \underline{T}\,\underline{\underline{P}}^T \tag{3}$$

Where $\underline{X}' = \underline{X} - \underline{e}$ is the reconstructed measurement error.

### *B. Theory of NLPCA*

PCA makes one stringent but powerful assumption: linearity. With this assumption, PCA is limited to re-expressing the data as a linear combination of its basis vectors. Thus, if the data contains nonlinear lower-dimensional structure, PCA will not be able to detect it [7]. To overcome this shortcoming, [4] introduced a nonlinear generalization to PCA. The fundamental difference between PCA and NLPCA is that NLPCA allows arbitrary nonlinear mapping from $\Re^m$ to $\Re^p$ whereas PCA only allows linear mapping. Consider a mapping of the type in (4).

$$\underline{T} = \underline{f}(\underline{X}) \tag{4}$$

In this equation, $\underline{f}$ is a general nonlinear vector function. It is consisted of $p$ individual nonlinear functions; $\underline{f} = \{f_1, f_2, \ldots, f_p\}$, analogous to the columns of $\underline{P}$, such that if $T_i$ represents the *i*th element of $\underline{T}$,

$$T_i = f_i(\underline{X}) \tag{5}$$

The reconstruction of the original data is implemented by a second nonlinear vector function $\underline{g} = \{g_1, g_2, \cdots, g_m\}$:

$$X_j^{'} = g_j(\underline{T}) \tag{6}$$

The loss of information is again measured by $\underline{e} = \underline{X} - \underline{X}'$, and analogous to PCA, the functions $\underline{f}$ and $\underline{g}$ are selected to minimize $\left\|\underline{\underline{e}}\right\|$. For details, please refer to [4].

The NLPCA is solved using a five-layer autoassociative feed-forward neural network [4]. A feed-forward neural network is a nonparametric statistical model. It consisted of a series of parallel layers, each of which contains a number of processing elements, or neurons, such that the output of the *i*th layer is used as input to the $(i+1)$th [10,11]. The input signals are propagated in a feed-forward direction on a layer-by-layer basis, i.e., signal travel only in the forward direction from the first to the last layer. If $y_j^{(i)}$ is the output of the *j*th neuron of the *i*th layer, then

$$y_k^{(i+1)} = \sigma^{(i+1)}\left[\sum_j w_{jk}^{(i+1)} y_j^{(i)} + b_k^{(i+1)}\right] \tag{7}$$

is the output of the *k*th neuron of the $(i+1)$th layer. The elements of the arrays $w_{jk}^{(i)}$ are referred to as the *weights*, and those of the vectors $b_k^{(i+1)}$ are the *biases*.

The weights and biases are adjustable parameters. The transfer function for the $(i+1)$th layer is given by $\sigma^{(i+1)}$. The input layer typically has linear transfer function, and nonlinear transfer functions (generally hyperbolic function) are used in some or all the remaining layers [5].

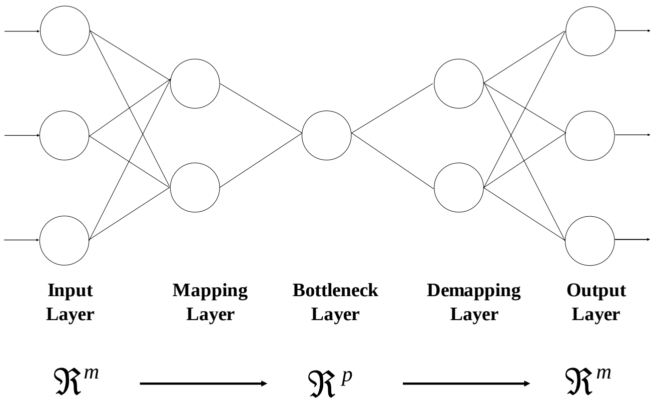



Fig. 2. The five-layer feed-forward auto-associative neural network used to perform NLPCA.

Fig. 2 shows the architecture of the five-layer network for $p=1$ NLPCA approximation to the data set. The number of neurons in the so-called bottleneck layer can be increased accordingly to obtain higher dimensional structure, i.e $p>1$. The input and the output layers contain *m* neurons, and they may have, as in the case of the bottleneck layer, linear transfer functions. The network contains three hidden layers. The mapping and demapping layers must have nonlinear transfer function (hyperbolic tangent) for the network to have the capability of modeling arbitrary nonlinear functions in $\underline{f}$ and $\underline{g}$. They may or may not have the same number of neurons [4].

The weights and biases are optimized using a conjugate gradient algorithm until the sum of squared differences between network input ( $\underline{\underline{X}}$ ) and output( $\underline{\underline{X}}'$ ) is minimized as in (8). As the network is trained to approximate as closely as possible the input data itself, it is said to be auto-associative or self-supervised backpropagation [4,5].

$$E=\sum_{p=1}^{n}\sum_{i=1}^{m}\left(Y_i-Y_i'\right)_p^2 \qquad (8)$$

## III. APPLICATIONS

### A. Simulation Studies

To compare the performance of PCA and NLPCA, the same case study as used in [1] for diagnosing nonlinearities in the control valves are used. A simple single-input, single output system in a feedback control configuration was used for generating simulated data. The first order process with time delay is given by the following transfer function:

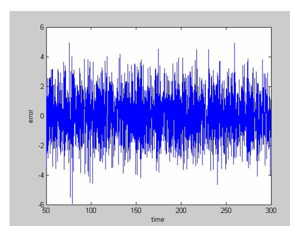
$$G(z^{-1})=\frac{z^{-3}\times(1.45z-1)}{z-0.8} \qquad (9)$$

The process was assumed to be linear, and controlled by a PI controller. An integrated random noise was added to the process. The process output (denoted as *y*) and the controller output (denoted as *u*) were used to detect nonlinearity present in the data for the four cases of (1) well-tuned controller, (2) controller with excessive control action, (3) controller with presence of external oscillatory disturbances, and (4) controller with presence of stiction.

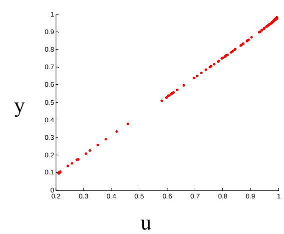
To avoid local minimum, a set of 10 ensembles of NLPCAs process the same dataset [5,7,9] and the average result is taken. Each NLPCA is initialized with random weights and biases. To prevent the NLPCA from over-fitting, 20% of the steady-state data were chosen randomly as test data set [8] using early-stopping criteria. Only $p=1$ is considered because the data set being considered only consisted of *y* and *u*.

*Case 1: Well tune controller*

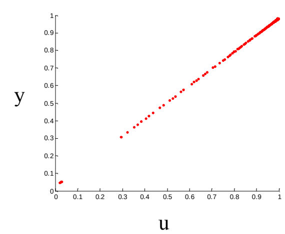
The PI controller parameters for this case were $K_c$=0.15 and $I$=$K_c/\tau_i$=0.15 $s^{-1}$. Fig. 3 shows the result for this case.




(a) Time trend error




(b) NLPCA approximation



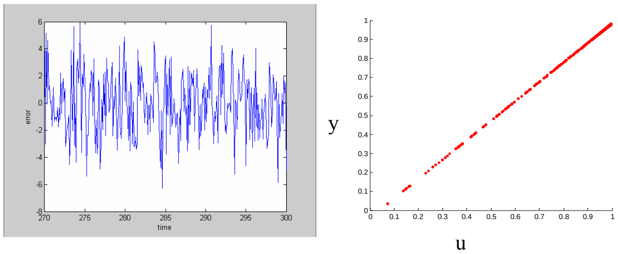

(c) PCA approximation

Fig. 3. PCA vs. NLPCA performance for well-tuned controller.

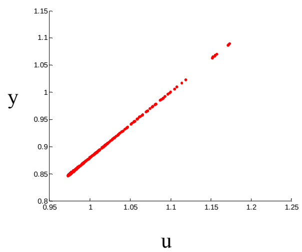
From the figure, it can be clearly seen that both NLPCA and PCA approximations give straight lines (see Fig. 3(b) and (c)). The error signal plot in Fig. 3(a) clearly shown that there was only random noise present in the loop, and hence linear signals are correctly given by both methods.

*Case 2: Controller with excessive integral action*

For this case the controller parameters were set to $K_c$=0.15 and $I$=$K_c/\tau_i$=(0.15/2.5) $s^{-1}$. Compared to case 1, this controller has excessive integral action.Fig 2(a)a shows the error signal plot result for case 2. The presence of relatively large integral action produces oscillations in the process variables.




(a) Time trend error (b) NLPCA approximation




(c) PCA approximation

Fig. 4. PCA vs. NLPCA for controller with excessive integral action.

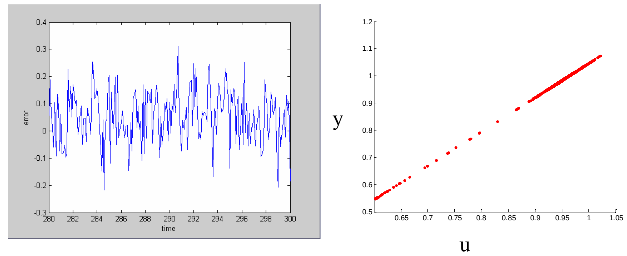
From Fig. 2(b) and (c), both NLPCA and PCA approximations are straight lines, clearly indicating that the error signal is not due to nonlinearity.

*Case 3: Controller with the presence of external oscillatory disturbance*

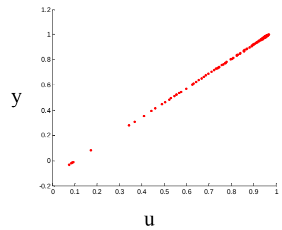
For this case, [1] specified a sinusoid with amplitude 2 and frequency 0.01 to be added to the process output in order to feed external oscillatory disturbances to the process. The error signal plot in Fig. 5(a) clearly shows the presence of the sinusoidal disturbance.
Even though the output is clearly oscillating, NLPCA does not detect any underlying nonlinearity present in the loop. Fig. 5(b) shows straight line approximation when the data is processed with NLPCA. PCA also gives straight line as shown in Fig. 5(c).
The NLPCA and PCA approximations are straight lines. It clearly shows that the reason for oscillation is not due to nonlinearity in the control loop.

*Case 4: Controller with the presence of stiction*

To investigate the presence of stiction, a stiction model developed by [12] was used. To perform the simulation for this particular case, $s$ =3 and $j$ =1 were used. Fig. 6(a) shows the time trend of the control error signal in case 4.

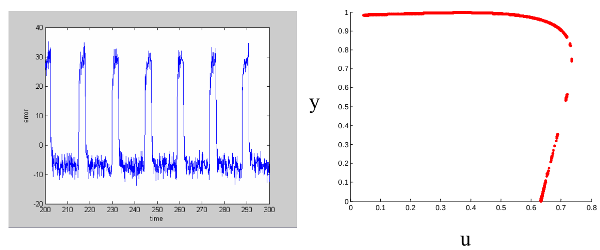



(a) Time trend error (b) NLPCA approximation

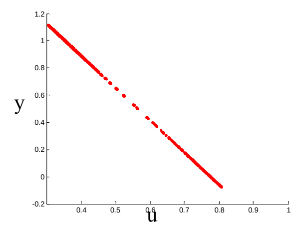



(c) PCA approximation

Fig. 5. PCA vs. NLPCA for controller external oscillatory disturbance.




(a) Time trend error (b) NLPCA approximation




(c) PCA approximation

Fig. 6. PCA vs. NLPCA for controller with presence of stiction.

For this case, NLPCA detected the underlying nonlinearity in the process, and Fig. 6(b) clearly indicating a curve. On the other hand, PCA approximation is still a straight line. PCA, which is linear in nature, can not uncover the nonlinear correlation among variables.

The results obtained for these four case studies show the superior capability of NLPCA in comparison to PCA in detecting nonlinearity. TABLE 1 shows the summary of the results obtained here, and comparison to the Higher-order based methods (HOS-based) developed by [1].

TABLE 1
Comparison between PCA, NLPCA and HOS-based methods in detecting stiction nonlinearity.

| | **PCA** | **NLPCA** | **HOS-based** |
|---|---|---|---|
| **Case 1** | Linear | Linear | Linear |
| **Case 2** | Linear | Linear | Linear |
| **Case 3** | Linear | Linear | Linear |
| **Case 4** | Linear | Nonlinear | Nonlinear |

The results in TABLE 1 clearly indicate the promising ability of NLPCA in the field of diagnosing poor control loop performance.

### B. *Industrial case studies*

Normalized operation data from two chemical processes are shown in Fig. 7-10. Fig. 7 and 8 show the results for applying PCA and NLPCA to the data of a level control loop with stiction in the control valve. Fig. 7 shows the time trend plot of the pv and pv-op plots.

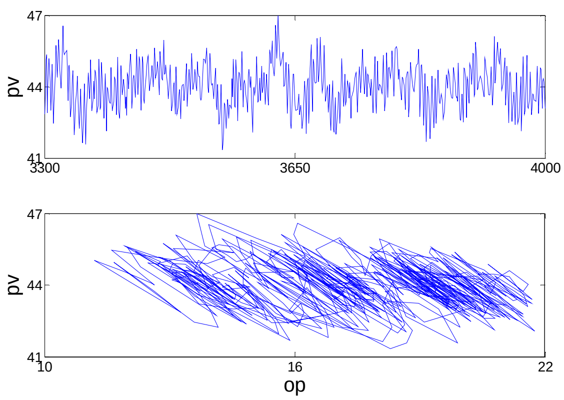


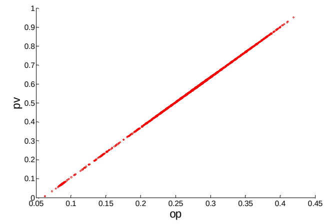

Fig. 7. Level control loop with valve stiction: pv trend and pv-op plots

.

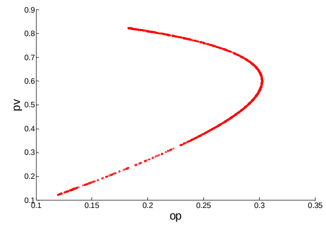
From Fig. 8, it can be clearly observed that PCA failed to detect the stiction in the loop. However, NLPCA correctly concludes that there is stiction in the system.
Fig. 9 shows the time trend plot of pv and pv-op plots for a flow loop with no stiction exists. The corresponding results when applying PCA and NLPCA to the data is shown in Fig. 10. Again, NLPCA performance supersedes that of PCA in detecting stiction as can be clearly observed in Fig. 10(b).

## IV. Conclusion

In this paper, PCA and NLPCA methods have been applied in diagnosing the poor control loop performance. Using the same case studies as in [1] as well as industrial data, it has been shown that NLPCA performance supersedes that of PCA in identifying nonlinearities in the poor loop data. The NLPCA performance is similar as the HOS-based method developed in [1], and this shows that NLPCA is a promising tool for control loop performance analysis.



(a) PCA approximation



(b) NLPCA approximation
Fig. 8. PCA vs NLPCA for Level control loop with valve stiction.

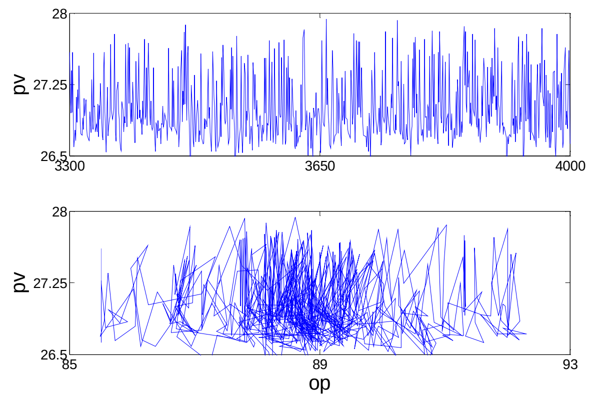
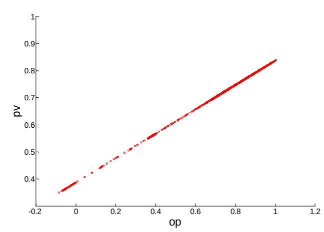
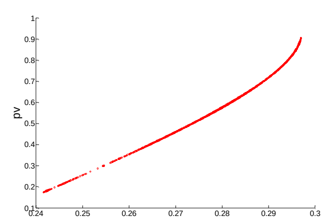



Fig. 9. Flow control loop with no stiction: pv trend and pv-op plots



(a) PCA approximation



(b) NLPCA approximation
Fig. 10. PCA vs NLPCA for Flow control loop with no stiction.

## ACKNOWLEDGMENT

The authors would like to thank Universiti Teknologi PETRONAS for the funding provided for this work. The authors also would like to thank individuals M Akbar Shafawi, Tengku Amansyah Tuan Mat, M Nazeem Nordin and Siti Aishah Sahari for their willingness in providing industrial insights to contribute to this paper.

## REFERENCES

[1] Choudhury, M. A. A. S., Shah, S. L., and Thornhill, N. F., Diagnosis of poor control-loop performance using higher-order statistics, *Automatica*, 40, 2004, pp. 1719-1728.

[2] Kariwala, V., M. A. A. S. Choudhury, H. Douke, S. L. Shah, H. Takada, J. F. Forbes and E. S. Meadows, Detection and diagnosis of plant-wide oscillations: An application study. In: *Proceedings of IEEE APC 2004*. Vancouver, BC, Canada, 2004, http://citeseer.ist.psu.edu/kariwala04detection.html.3

[3] L. Eriksson, E. Johansson, N. Kettaneh-Wold, S. Wold, *Multi- and Megavariate Data Analysis: Principles and Applications*, Umetrics Academy, 2001.

[4] M. A., Kraner, "Nonlinear principal component analysis using autoassociative neural networks," AIChE Journal, vol. 37, 1991, pp. 233-243.

[5] A. Monahan, "Nonlinear principal component analysis by neural networks: theory and application to the Lorenz system," *J. Climate*, vol. 13, pp. 821-835, 2000.

[6] A. Monahan, "Nonlinear principal component analysis: Tropical Indo-Pacific sea surface temperature and sea level pressure," *J. Climate*, vol. 14, pp. 219-233, 2001.

[7] S. S. C. Botelho, R. A. Bern, I. L. Almeida, M. M. Mata, "C-NLPCA: Extracting nonlinear principal components of image datasets," *Anais XII Simposio Brasileiro de Sensoriamento Remoto*, Goiania, Brasil, 16-21 abril 2005, INPE, pp. 3495-3502.

[8] K. Bernhardt, K. W. Wirtz, "Reduction of complex models using data-mining and nonlinear projection techniques," Available: http://citeseer.ist.psu.edu/707321.html

[9] W. Hiesh, "Nonlinear principal component analysis by neural networks," Tellus, 53A:559-615, 2001.

[10] N, Yusoff, S. Srinivasan, "Neural-Networks simulation in predicting ozone concentration," *SOMCHE* 2004.

[11] V. R. Radhakrishnan, A. R. M0hamed, "Neural networks for the identification and control of blast furnace hot metal quality," *Journal of Process Control*, vol. 10, pp. 509-524, 2000.

[12] Choudhury, M. A. A. S., Thornhill, N. F., and Shah, S. L., Modeling valve stiction, *Control Engineering Practice*, 13, pp. 641-658, 2005.