A brief version of this paper has been accepted (June 2003) and will be appeared in Automatica soon

Diagnosis of Poor Control Loop Performance using Higher Order Statistics

M. A. A. Shoukat Choudhury, Sirish L. Shah, *
Department of Chemical and Materials Engineering
University of Alberta, Edmonton, AB, T6G 2G6, Canada,

Nina Thornhill
Department of Electronic and Electrical Engineering
University College, London, UK WC1E 7JE

Abstract

Higher Order Statistical (HOS) techniques were first proposed over four decades ago. This paper is concerned with higher order statistical analysis of closed loop data for diagnosing the causes of poor control loop performance. The main contributions of this work are to utilize HOS tools such as cumulants, bispectrum and bicoherence to develop two new indices: the Non-Gaussianity Index (NGI) and the Non-Linearity Index (NLI) for detecting and quantifying non-Gaussianity and nonlinearity that may be present in regulated systems, and to use routine operating data to diagnose the source of nonlinearity. The new indices together with some graphical plots have been found to be useful in diagnosing the causes of poor performance of control loops. Successful applications of the proposed method are demonstrated on simulated as well as industrial data. This study clearly shows that HOS based methods are promising for closed loop performance monitoring.

Keywords: diagnosis, stiction, bispectrum, cumulants, bicoherence, control loop performance

1 Introduction

The field of controller performance monitoring has received much attention in the engineering research literature. However, the diagnosis of poor performance remains an open area. Performance diagnosis requires identification of the causes of poor performance, as for example due to poor controller tuning, presence of disturbances, process and/or actuator non-linearities. If there are some non-linearities in the control loop, the controller may not perform at the desired level. Non-linearities degrade the performance of the controller in several ways. For example, they may produce oscillations in process variables, shorten the life of the control valve, may upset process stability, and in most cases lead to inferior quality end-products thus causing larger rejection rates and reduced profitability. The non-linearities may be present in the process itself or in the

*Author to whom all correspondence should be addressed. Email:Sirish.Shah@ualberta.ca
[0] *A preliminary version of this paper was presented at the 15^{th} IFAC world congress, 2002, Barcelona, Spain*

actuators or control valves. This study is concerned with actuator nonlinearities. Actuator or valve nonlinearities are typically due to faults such as stiction, backlash, saturation, deadzone, ruptured diaphragm, and/or corroded or eroded valve seats.

Classical signal processing tools such as power spectrum utilize only the first and second order moments, i.e., the mean and covariance. Such tools are mainly useful for analyzing signals from linear processes. In case of non-linear signals, one needs to look at other methods of characterizing their statistical properties (Rao and Gabr, 1980; Hinich, 1982; Nikias and Petropulu, 1993; Choudhury *et al.*, 2002). This necessitates the use of higher order statistical tools. The third and fourth order moments or cumulants and their frequency domain counterparts are found to be useful in analyzing non-linearities in communication signals, radar signals, nonlinear ocean wave analysis, seismic signal analysis, speech signal analysis and mechanical machine condition monitoring (Kim and Powers, 1979; Nikias and Petropulu, 1993; Fackrell, 1996; Collis *et al.*, 1998). Although HOS techniques have been widely used in the above mentioned areas, they have not been used in solving problems in process control. This paper introduces HOS methods and shows the potential of using these tools in controller performance analysis and diagnosis.

In this work, the diagnosis of the causes of poor performance of control loops is performed by analyzing closed loop data using HOS based methods. If the process is oscillating and the cause is due to poor or tightly tuned controllers or due to a linear external oscillatory disturbance, the proposed method does not detect any nonlinearity. If the method detects nonlinearity in the control error signal, the cause is most likely due to nonlinearities in the control valve or in the process itself. Once the nonlinearity is detected, the process output (*pv*) vs. the controller output (*op*) plot and the manipulated variable (*mv*) vs. the controller output (*op*) plot can be used to locate its source, e.g., due to valve stiction and/or backlash. Since most of the time, processes run under regulatory mode the method has been developed based on routine operating data obtained under regulatory control. Therefore, it does not require any additional process excitation or set point change.

2 Higher Order Statistics: Brief Preliminaries

The first and second order statistics (e.g., mean, variance, autocorrelation, power spectrum) are popular signal processing tools and have been used extensively for the analysis of process data. However second order statistics are only sufficient for describing linear processes. In practice, there are many situations when the process deviates from linearity and exhibits nonlinear behavior. Such type of processes can be conveniently studied using HOS. There are three main reasons for using HOS: to extract information due to deviations from Gaussianity, to recover the true phase character of the signals, and to detect and quantify nonlinearities in the time series (Nikias and Petropulu, 1993).

Time domain data itself is a good source of information. Many statistical measures, e.g., moments, cumulants, auto-correlation, cross-correlation have been developed to measure the time domain information in such data. Almost all type of data are usually collected as samples at regular intervals of time. In statistical analysis, it is often assumed that the time series or the signal is stationary and ergodic. This paper makes the use of these assumptions. Since this study mainly deals with the frequency domain HOS analysis, the time domain HOS measures, e.g., cumulants, will not be discussed here. However, for a tutorial description of time domain HOS, readers are referred to (Stuart and Ord, 1987; Fackrell, 1996; Choudhury and Shah, 2001).

Not all the information content of a signal can be necessarily and easily obtained from time domain statistical analysis of the data. Transforming the signal from time to frequency domain can expose the periodicities of the signal, can detect the nonlinearities present in the signal and can also aid in understanding the signal generating process. Just as the power spectrum is the frequency domain counterpart of the second order moment of a signal and represents the decomposition or spread of the signal energy over the frequency channels obtained from the Fast Fourier Transform, the bispectrum is the frequency domain representation of the third order cumulants. It is defined as

$$B(f_1, f_2) \triangleq DDFT[c_3(\tau_1, \tau_2)] \equiv E[X(f_1)X(f_2)X^*(f_1 + f_2)] \tag{1}$$

where, $B(f_1, f_2)$ is the bispectrum in the bifrequency (f_1, f_2), DDFT stands for Double Discrete Fourier Transformation, $c_3(\tau_1, \tau_2)$ is the third order cumulant, τ's are the time-lag variables, $X(f)$ is the discrete Fourier transform of any time series $x(k)$, and '*' denotes complex conjugate.

Equation 1 shows that the bispectrum is a complex quantity having both magnitude and phase. It can be plotted against two independent frequency variables, f_1 and f_2 in a three dimensional (3d) plot. Just as the discrete power spectrum has a point of symmetry at the folding frequency, the discrete bispectrum also has 12 regions of symmetries in the (f_1,f_2) plane (Rosenblatt and Van Ness, 1965; Nikias and Petropulu, 1993). The bispectrum in one region, the principal domain, gives sufficient information. The other regions of the (f_1,f_2) plane are redundant. Each point in such a plot represents the bispectral content of the signal at the bifrequency, (f_1, f_2). In fact, the bispectrum at point $(B(f_1, f_2), f_1, f_2)$ measures the interaction between frequencies f_1 and f_2. This interaction between frequencies can be related to the non-linearities present in the signal generating systems (Fackrell, 1996) and therein lies the core of its usefulness in the detection and diagnosis of non-linearities.

In order to remove the undesired property of the variance of the estimated bispectrum (Hinich, 1982), the bispectrum can be normalized in a such way that it gives a new measure called bicoherence whose variance is independent of the signal energy (Fackrell, 1996). Bicoherence is defined as:

$$bic^2(f_1, f_2) \triangleq \frac{|B(f_1, f_2)|^2}{E[|X(f_1)X(f_2)|^2]E[|X(f_1+f_2)|^2]} \tag{2}$$

where 'bic' is known as the bicoherence function. A useful feature of bicoherence function is that it is bounded between 0 and 1.

The underlying methods for bispectrum/bicoherence estimation are extensions of the power spectrum estimation methods. There are two broad non-parametric approaches: the indirect method, based on estimating the cumulant functions and then taking the Fourier Transform; and the direct method, based on Welch's segment averaging approach. For details about these methods, see (Nikias and Petropulu, 1993; Choudhury and Shah, 2001). Unless otherwise stated, the direct method of bispectrum estimation with a data length of 4096, a segment length of 64, a 50% overlap, Hanning window with a length of 64, and a DFT length of 128 has been used throughout this work.

3 Test of Gaussianity and Linearity based on the Bicoherence

The presence of nonlinearities in the control loop is one of the main reasons for poor performance of a linear controller designed based on linear control theory. The nonlinearity may be due to the presence of nonlinearities such as stiction, deadzone, hysteresis in the control valve or the nonlinear nature of the process itself. Such a nonlinear system often produces a non-Gaussian and nonlinear time series. The test of Gaussianity of a signal or the test of presence of nonlinearity in a system is a useful diagnostic aid towards determining the poor performance of a control loop. This section describes the development and derivation of the test to check a signal's Gaussianity and nonlinearity.

A discrete ergodic stationary time series, $x(k)$, is called linear, if it can be represented by

$$x(k) = \sum_{n=0}^{M-1} h(n)e(k-n) \tag{3}$$

where, $e(k)$ is a sequence of independent identically distributed random variables with $E[e(k)] = 0$, $\sigma_e^2 = E[e^2(k)]$, and $\mu_3 = E[e^3(k)]$. For this case, the following frequency domain relationships can be obtained.

$$\text{The power spectrum: } P(f) = \sigma_e^2 \mid H(f) \mid^2 \equiv |X(f)X^*(f)| \tag{4}$$

$$\text{and the bispectrum: } B(f_1, f_2) = \mu_3 H(f_1)H(f_2)H^*(f_1+f_2) \tag{5}$$

where, $H(f) = \sum_{n}^{M-1} h(n)e^{-inf}$. Equation 2 can be rewritten as

$$\begin{aligned} bic^2(f_1, f_2) &\triangleq \frac{|B(f_1, f_2)|^2}{E[|X(f_1)X^*(f_1)||X(f_2)X^*(f_2)|] \;\; E[|X(f_1+f_2)X^*(f_1+f_2)|]} \\ &\equiv \frac{|B(f_1, f_2)|^2}{E[|P(f_1)||P(f_2)|]E[|P(f_1+f_2)|]} \end{aligned} \tag{6}$$

For the linear time series, substituting the expressions from equation 4 and 5, it can be shown that

$$bic^2(f_1, f_2) = \frac{\mu_3^2}{\sigma_e^6} \tag{7}$$

Equation 7 shows that for any linear signal, x, the squared bicoherence will be independent of the bifrequencies i.e., a constant in the bifrequency plane. If the squared bicoherence is zero, the signal x is Gaussian because the skewness or μ_3 is also zero in such a case. Strictly speaking, such a signal should be called non-skewed with a symmetric distribution instead of Gaussian. However, in this paper and also in most of the HOS literature (Kim and Powers, 1979; Rao and Gabr, 1980; Hinich, 1982; Nikias and Petropulu, 1993; Fackrell, 1996; Collis *et al.*, 1998) the two terms - nonskewed and Gaussian - have been used interchangeably. To check whether the squared bicoherence is constant or not, two tests are required. One is for testing the zero squared bicoherence which shows that the signal is Gaussian and thereby the signal generating process is linear. The other is to test for a non-zero constant squared bicoherence which shows that the signal is non-Gaussian but the signal generating process is linear.

The bicoherence is a complex quantity with real and imaginary parts. The magnitude of the squared bicoherence can be obtained as

$$bic^2 = \Re(bic)^2 + \Im(bic)^2 \tag{8}$$

where $\Re$ and $\Im$ are real and imaginary parts, respectively. It is well known in the HOS literature that the bicoherence is a complex normal variable, i.e., both the estimates of real and imaginary parts of the bicoherence are normally distributed (Hinich, 1982) and asymptotically independent, i.e., the estimate at a particular bifrequency is independent of the estimates of its neighboring bifrequencies (Fackrell, 1996). Therefore, the squared bicoherence at each frequency is a chi-squared (χ^2) distributed variable with 2 degrees of freedom. Hinich (Hinich, 1982) showed that the signal of interest is Gaussian if the scaled skewness function, a function similar to bicoherence, is asymptotically centrally χ^2 distributed with 2 degrees of freedom. This information was used by (Fackrell, 1996) to test bicoherence at each frequency in the principal domain. The disadvantage of this test is that while applying to each of the bifrequencies in the principal domain of squared bicoherence plot, the probability of false detection accumulates because of a large number of bifrequencies in the principal domain and thus it overestimates the number of bifrequencies where the bicoherence magnitude is significant. A modified test with better statistical properties but no frequency resolution is formulated by averaging the squared bicoherence over the triangle of the principal domain. The test can be summarized as follows:

- Null Hypothesis, H_o: The signal is Gaussian
- Alternate Hypothesis, H_1: The signal is not Gaussian.

Under the null hypothesis, the test for the average squared bicoherence can be based on the following equation:

$$P(2KL\overline{\hat{bic^2}} > c_\alpha^{\chi^2}) = \alpha \tag{9}$$

where, $c_\alpha^{\chi^2}$ is the critical value calculated from the central χ^2 distribution table for a significance level of α at 2L degrees of freedom since $\overline{\hat{bic^2}} = \sum_{i=1}^{L} \hat{bic}_i^2$ and L is the number of bifrequencies inside the principal domain of the bispectrum, K is the number of segments used in data segmentation during bicoherence estimation.

If the number of bifrequencies in the principal domain is very large (more than 100) the normal approximation of the χ^2 distribution can be used. The approximation is given by (Abramowitz and Stegun, 1972):

$$c_\alpha^{\chi^2} = \frac{1}{2}[c_\alpha^z + \sqrt{2dof - 1}]^2 \tag{10}$$

where $c_\alpha^{\chi^2}$ and c_α^z are the critical values of χ^2 and standard normal distribution at a significance level of α, respectively and dof is the degrees of freedom. Now substituting equation 10 into 9 with 2L degrees of freedom, it can be shown that

$$P(\overline{\hat{bic^2}} > \frac{1}{4KL}[c_\alpha^z + \sqrt{4L - 1}]^2) = \alpha \tag{11}$$

This equation can be rewritten as

$$P(\overline{\hat{bic^2}} - \overline{bic^2{}_{crit}} > 0) = \alpha \tag{12}$$

$$or, \qquad P(NGI > 0) = \alpha \tag{13}$$

where, $\overline{bic^2{}_{crit}} = \frac{1}{4KL}[c_\alpha^z + \sqrt{4L-1}]^2$ and $NGI \triangleq \overline{\hat{bic^2}} - \overline{bic^2}_{crit}$

NGI stands for the Non-Gaussianity Index. Therefore, at a confidence level of α the following rule based decision can be obtained:

- **if $NGI \leq 0$, the signal is GAUSSIAN**
- **if $NGI > 0$, the signal is NON-GAUSSIAN**

Therefore, a signal is Gaussian (non-skewed) at a confidence level of α if the NGI is less than or equal to zero. This index has been defined to automate the decision.

If the signal is found to be Gaussian, the signal generating process is assumed to be linear. In the case of non-Gaussian signal the signal generating process should be tested for its linearity. As shown in equation 7, if the signal is non-Gaussian and linear the magnitude of the squared bicoherence should be a non-zero constant at all bifrequencies in the principal domain. The constancy of the squared bicoherence (skewness) was tested by Rao and Gabr (1980) using an F test. . Hinich reported that this test is very vulnerable to outliers. He suggested a method based on the Sample Interquartile Range (SIQR) of the χ^2 distribution of the squared skewness function (Hinich, 1982). But this test depends on the sample size of the time series and also the SIQR is not an ideal measure for constant bicoherence because it is easy to see that SIQR of the squared bicoherence can be zero though all squared bicoherence are not equal (Yuan, 1999). A simple way to check the constancy of squared bicoherence is to have a look at the 3d squared bicoherence plot and observe the flatness of the plot. But this can be tedious and cumbersome for a large number of loops. Alternatively, if the squared bicoherence is of a constant magnitude at all bifrequencies in the principal domain, the variance of the estimated bicoherence should be zero. To check the flatness of the plot or the constancy of the squared bicoherence, the maximum squared bicoherence can be compared with the average squared bicoherence plus two or three times the standard deviation of the estimated squared bicoherence. At a 95% confidence level if the maximum squared bicoherence, $\hat{bic^2}_{max}$ is less than $(\overline{\hat{bic^2}} + 2\sigma_{\hat{bic^2}})$, the magnitudes of squared bicoherence are assumed to be a constant or the surface is flat. The automatic detection of this can be performed using the following Nonlinearity Index (NLI), which is defined as:

$$NLI \triangleq \left| \hat{bic^2}_{max} - (\overline{\hat{bic^2}} + 2\sigma_{\hat{bic^2}}) \right| \tag{14}$$

where, $\sigma_{\hat{bic^2}}$ is the standard deviation of the estimated squared bicoherence and $\overline{\hat{bic^2}}$ is the average of the estimated squared bicoherence. Ideally, the NLI should be 0 for a linear process. This is because if the squared bicoherence is a constant at all frequencies, the variance will be zero and both the maximum and the mean will be same. Therefore, it can be concluded that:

- **if $NLI = 0$, the signal generating process is LINEAR**
- **if $NLI > 0$, the signal generating process is NONLINEAR**

Since the squared bicoherence is bounded between 0 and 1, the nonlinearity index (NLI) is also bounded between 0 and 1.

Practical Implementation: To make this work in practice it is difficult to obtain an exact zero value for NGI for Gaussian signals. Therefore, we select a threshold value, β, of NGI such that $NGI < \beta$ implies a Gaussian signal. To the best knowledge and experience of the authors, for $\alpha = 0.05$, an NGI value of less than 0.001 can be assumed to be zero. Consequently, if $NGI \leq 0.001$ the signal can be assumed to be Gaussian at a 95% confidence level. For NLI, a value less than 0.01 is assumed to be zero and consequently, the process is considered to be linear at a 95% confidence level. The larger the NLI the higher is the extent of nonlinearity. The detailed diagnosis procedure can be summarized in a rule based decision flow diagram shown in figure 1.

4 Illustrative Example

4.1 Bicoherence of a linear and nonlinear signal

Two signals, y_{linear} and $y_{nonlinear}$, were generated using the following equations.

$$\begin{aligned} x(k) &= 3H(q^{-1})d_1(k) \\ y_{linear} &= x(k) + d_2(k) \qquad (15) \\ y_{nonlinear} &= x(k) + 0.1x(k)^2 + d_2(k) \qquad (16) \end{aligned}$$

where, $d_1(k)$ and $d_2(k)$ are zero mean white noises with variance 1 and 0.001 respectively, and $H(q^{-1})$ is a narrow pass Butterworth filter with a frequency range 0.095 to 0.105 in a 0 to 0.5 normalized frequency scale such that $f = 1$ is the sampling frequency.

The objective of this example is to demonstrate the power of the bicoherence in the detection of nonlinearity. By merely looking at the time trend of the signals (the left panel of figure 2), it is not possible to differentiate between them. Also, the power spectrums (the middle panel of figure 2) or the second order moments look alike and are unable to detect the non-linearity present in the second signal. The right panel of figure 2 shows the three dimensional bicoherence plots. For y_{linear}, the test result is $NGI = -0.0028$. Clearly the NGI index indicates that the signal is Gaussian, and therefore the non-linearity test result is not required here. On the other hand for $y_{nonlinear}$, the NGI equals 0.002 detecting the non-Gaussianity of the signal. The nonlinearity test gives $NLI = 0.37$ which clearly indicates the presence of nonlinearity in this signal. From the bicoherence plot, the peak position in the principal domain is approximately at the (0.1,0.1) bifrequency. This means that the nonlinearity in the signal is due to interaction of these two frequencies. Examining the signal generating system shows that the band pass filtered signal has the frequency range [0.095 to 0.105]. This signal was squared to introduce a nonlinearity. Therefore, the nonlinearity is due to the multiplication of two signals, each of them having a frequency of approximately 0.1. In the bicoherence plot the frequencies identified are also 0.1 and 0.1. Therefore, the HOS based method correctly identifies the presence of a nonlinearity and the frequencies of nonlinear interactions.

4.2 Bicoherence of a nonlinear sinusoid signal with noise

An input signal was constructed by adding two sinusoids, each of them having a different frequency and phase. That is,

$$\begin{aligned} x^{'}(k) &= sin(2\pi f_1 k + \pi/3) + sin(2\pi f_2 k + \pi/8) \\ x(k) &= x^{'}(k) + d(k) \qquad (17) \\ y(k) &= x^{'}(k) + 0.05x^{'}(k)^2 + d(k) \qquad (18) \end{aligned}$$

where, $f_1 = 0.12$, $f_2 = 0.30$ on the normalized frequency scale, and $d(k)$ is a white noise sequence with variance 0.04 .

The left panel of the figure 3 shows the time series while the middle panel shows the power spectrum of the signal x and y, respectively. Neither of these plots help in distinguishing the two signals. However, the use of higher order statistics can successfully detect the nonlinearities present in y. The right panel of figure 3 shows the three dimensional squared bicoherence plots of x and y, respectively. For the signal x, $NGI = 0.0008$, clearly indicating that the signal is Gaussian and linear. On the other hand, for y the test results are $NGI = 0.016$ and $NLI = 0.233$. Thus the nonlinearity present in y is correctly detected. This example also shows the sensitivity of the proposed indices to the presence of nonlinearity in the signal. The presence of as small as 5% of the nonlinear square term in the signal y has been detected. The peaks in the bifrequency plane can be explained by rewriting the expression for y as:

$$\begin{aligned} y(k) = {} & \sin(2\pi f_1 k + \pi/3) + \sin(2\pi f_2 k + \pi/8) + 0.05[1 - \cos(2(2\pi f_1 k + \pi/3)) \\ & - \cos(2(2\pi f_2 k + \pi/8)) + \cos(2\pi(f_1 - f_2)k + \pi/3 - \pi/8) \\ & - \cos(2\pi(f_1 + f_2)k + \pi/3 + \pi/8)] + d(k) \qquad (19) \end{aligned}$$

The nonlinearities can be caused by the interactions of any two of the signals with frequencies f_1, f_2, $2f_1$, $2f_2$, $f_2 - f_1$, and $f_1 + f_2$. For the output signal y, the squared bicoherence plot shows peaks at (0.12,0.12), (0.12,0.18), (0.30,0.30), and (0.12,0.30) bifrequencies. These bifrequencies correspond to (f_1, f_1),$(f_1, f_2 - f_1)$, (f_2, f_2),and (f_1, f_2), respectively. Note that since only 5% of the nonlinear term was added, the peak for the frequency $f_1 + f_2$ is not visible in the bicoherence plot due to its small size. Therefore, the bicoherence plot correctly identifies the frequency interactions that resulted from the presence of nonlinearity in the signal.

5 Simulation Examples to Diagnose the Causes of Poor Performance

As mentioned earlier in this paper the poor performance of a control loop may be due to a variety of reasons, for example poorly tuned controllers, presence of oscillatory disturbances, and nonlinearities. The purpose of this simulation example is to demonstrate the application of HOS based techniques in diagnosing the causes of poor performance. If the method does not detect any nonlinearity then the focus of the diagnosis should be on controller tuning or on the possible presence of an external oscillatory disturbance. If the method detects nonlinearity then the nonlinearity should be isolated or localized. Is it in the valve or in the process? This study assumes that the process is linear.

A simple single-input, single-output (SISO) system in a feedback control configuration (figure 4) was used for generating simulated data. The first order process with time delay is given by the following transfer function:

$$G(z^{-1}) = \frac{z^{-3}(1.45 - z^{-1})}{1 - 0.8z^{-1}} \tag{20}$$

The process is under regulatory control and is controlled by a PI controller. An integrated white noise generated by integrating random noise with a standard deviation of 0.224 was added to the process. The simulation was performed for 6000 sampling intervals. To remove the effect of transients the first few hundred data points were discarded and the last 4096 points of the error signal to the controller (*sp-pv*) were analyzed to detect the nonlinearity present in the system for the following cases:

5.1 Well tuned controller

The PI controller parameters for this case were $K_c = 0.15$ and $I = K_c/\tau_i = 0.15$ $second^{-1}$. The nonlinear 'stiction model' block was removed from the simulation block diagram. The top row of figure 5 shows the results for this case. The proposed test yields a value of $NGI = -0.0008$. This indicates that the error signal is Gaussian and linear. The corresponding bicoherence plot is flat.

5.2 Controller with excessive integral action

For this case the controller parameters were set to $K_c = 0.15$ and $I = K_c/\tau_i = (0.15/2.5)$ $second^{-1}$. Compared to the previous case, this controller has excessive integral action. The second row of the figure 5 shows the results for this case. The presence of relatively large integral action produces large oscillations in the process variables. An NGI value of -0.0007 indicates the Gaussianity and linearity of the system. It indicates that the poor performance is not due to nonlinearities. Also, since there are no external oscillatory disturbances, a suitable diagnosis is a poorly tuned controller.

5.3 Presence of an external oscillatory disturbance

A sinusoid with amplitude 2 and frequency 0.01 was added to the process output in figure 4 in order to feed an external oscillatory disturbances to the process. The results for this case are shown in the third row of figure 5. Horch's correlation method (Horch, 1999) of diagnosing the oscillation or more specifically valve stiction gives an odd correlation function between *op* and *pv* for this case, thereby falsely detecting the presence of stiction in the control loop. The proposed test gives an $NGI = -0.0003$. It clearly shows that the reason for the oscillation is not due to any nonlinearity in the system. The bicoherence plot is also flat.

5.4 Presence of stiction

A stiction model developed by (Choudhury *et al.*, 2003*a*; Choudhury *et al.*, 2003*b*) was used to perform this simulation. The model consists of two parameters -namely deadband plus stickband, s and slip jump, j. Figure 6 summarizes the model algorithm. It can be briefly described as:

- First, the controller output (mA) is provided to the look-up table where it is converted to valve travel in %.
- If this is less then 0 or more than 100, the valve is saturated.
- If the signal is within 0 to 100% range, it calculates the slope of the controller output signal.
- Then, the change of the direction of the slope of the input signal is taken into consideration. If the sign of the slope changes or remains zero for two consecutive instants, the valve is assumed to be stuck and does not move.
- When the cumulative change of the input signal is more than the amount of the stickband (say, "s"), the valve slips and starts moving.
- The parameter, "j" signifies the slip jump start of the control valve immediately after it overcomes the deadband plus stickband. It accounts for the offset between the valve input and output signals.
- Finally, the output is again converted back to a mA signal using a look-up table based on the valve characteristics.

To perform the simulation for this particular case, $s = 3$ and $j = 1$ were used. Note that in order to initiate limit cycling or oscillations in a simple first order time delay process in presence of valve stiction, a set point change at the beginning of the simulation is required. Then the process is left to operate under regulatory control. The last row of the figure 5 shows the time trend of the control error signal, the bicoherence plot and the *pv*-*op* plot. The presence of stiction produces oscillations in the process. The values of NGI and NLI are 0.05 and 0.048, clearly detecting the presence of nonlinearity in the process signal. After detecting the nonlinearity, the process variable versus controller output plot, i.e., *pv* - *op* plot can be used to diagnose the type of nonlinearity. Usually, the presence of distinct elliptical loops with sharp turn around points is an indication of the presence of stiction in the valve. Note that for other cases there are no such distinct cycles in *pv*-*op* plot (see the right panel of the figure 5). If the valve position (*mv*) is available, the *mv*-*op* plot can be conveniently and more accurately used to identify the type of nonlinearity in the valve.

6 Industrial Case Studies

The proposed method has been successfully applied in the detection and isolation of process/actuator faults for some industrial control loops. Two of these case studies are reported here.

6.1 Case 1: Dryer temperature control loop

This is a temperature control loop on a furnace feed dryer system at the Tech-Cominco mine in Trail, British Columbia, Canada. The temperature of the dryer combustion chamber is controlled by manipulating the flow rate of natural gas to the combustion chamber. The minimum variance performance index of this loop was very poor. Figure 7(a) shows time trends of the controlled variable, the controller output and the set point. It shows clear oscillations both in the controlled variable and the controller output. Figure 7(b) shows the bicoherence plot. The NGI and NLI values obtained for this loop were 0.006 and 0.197 respectively, clearly indicating the presence of nonlinearity in the loop. The presence of distinct cycles in the characteristic *pv*-*op* plot (figure 7(c)) together with the pattern obtained in figure 7(d) characterize the presence of backlash and stiction in the control valve. Therefore, this analysis was able to confirm the cause of poor loop performance due to the presence of valve stiction.

6.2 Case 2: Flow control loops

This analysis is for two flow control loops at Celanese Canada Ltd., a chemical complex located in Edmonton, Canada. Data was collected with a sampling interval of 1 min over two time periods: from April 10 to 17, 2001 and following the annual maintenance shutdown of the plant, from July 1 to 15, 2001. Results of both these loops are discussed below:

6.2.1 Flow loop 1

This is a recycle flow control loop. The detailed diagnostic plots are shown in figure 8. Time series of the data collected in April and July are shown in figures 8(a) and 8(b). The *op* trend in figure 8(a) shows that the valve movement was very slow and insignificant compared to the change in the error signal, (*pv*-*sp*). The values of NGI and NLI are 0.01 and 0.10 respectively. An NGI value of 0.01 shows that the signal is Non-Gaussian. The NLI value of 0.10 indicates the presence of nonlinearity in the error signal. The *op* time trend in figure 8(a) shows that a small change in *op* caused a big change in the *pv* value (note the range of y-axis for *op*, 49.4 to 50). Therefore, it was suggested that the nonlinearity in this loop was most likely due to an oversized valve. This 6 inch valve was replaced by a 3 inch valve during the annual maintenance shutdown of the plant (in May, 2001). In order to confirm the result of the analysis, additional data was collected in July and the results of the 'post-maintenance' data analysis are shown in the right part of the figure 8. For the new data set: $NGI = 0.0005$ (less than 0.001) and $NLI = 0.05$. These values indicate Gaussian linear system characteristics.

6.2.2 Flow loop 2

This is also a flow control loop at the outlet of a pump located at the bottom of a distillation column. Analysis of the April, 2001 data of this loop revealed that this loop had nonlinearity problems with $NGI = 0.032$ and $NLI = 0.13$. The diagnostic plots are shown in the top panel of the figure 9. The test results correctly detected the presence of a significant nonlinearity. The *pv* - *op* characteristic plot indicated a type of nonlinear characteristic in the process or the valve that had not been observed before. During the annual maintenance, the plant instrument personnel noticed that the valve seat and the plug were severely corroded. The valve was replaced. The results of the 'post-maintenance' analysis are shown in the lower panel of the figure 9. Now $NGI = 0.04$ and $NLI = 0.06$ indicating yet again the presence of a nonlinearity but in a substantially reduced form. The *pv* - *op* plot still shows unfamiliar patterns for unknown sources of nonlinearities. However the overall controller performance of this loop has improved significantly to the point where additional analysis was deemed unnecessary.

7 Conclusions

This paper has provided a brief introduction to Higher Order Statistics based data analysis tools. Two new indices, the Non-Gaussianity Index (NGI) and the Non-Linearity Index (NLI), based on HOS theory have been developed to detect and quantify signal non-Gaussianity and non-linearity. These indices together with specific patterns in the process output (*pv*) vs. the controller output (*op*) plot can be conveniently used to diagnose the causes of poor control loop performance. The method has been successfully applied to many industrial data sets and three such studies have been presented here. In both cases the results of the analysis were confirmed by the plant engineers. Current work in progress is focussed on detection and quantification of stiction.

8 Acknowledgements

Financial support to the first author in the form of a Canadian International Development Agency (CIDA) scholarship is gratefully acknowledged. The project has also been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Matrikon Consulting Inc., and the Alberta Science and Research Authority (ASRA) in the form of an NSERC-Matrikon-ASRA Industrial Research Chair Program at the University of Alberta.

References

Abramowitz, M. and I. A. Stegun (1972). *Handbook of Mathematical Functions*. Dover Publications. New York.

Choudhury, M. A. A. S. and Sirish L. Shah (2001). Detection and diagnosis of system nonlinearities using higher order statistics. Internal Report IRC021030. Department of Chemical and Materials Engineering, University of Alberta, Canada.

Choudhury, M. A. A. S., Nina F. Thornhill and Sirish L. Shah (2003*a*). A data-driven model for valve stiction. In: *to appear in ADCHEM 2004*. Hong Kong.

Choudhury, M. A. A. S., Nina F. Thornhill and Sirish L. Shah (2003*b*). Modelling valve stiction. *Submitted to Control Engineering Practice, Jan. 2003.*

Choudhury, M. A. A. S., Sirish L. Shah and Nina F. Thornhill (2002). Detection and diagnosis of system nonlinearities using higher order statistics. In: *15th IFAC World Congress*. Barcelona, Spain.

Collis, W.B., P.R. White and J.K. Hammond (1998). Higher-order spectra: The bispectrum and trispectrum. *Mechanical Systems and Signal Processing* **12**, 375–394.

Fackrell, Justin W. A. (1996). Bispectral Analysis of Speech Signals. PhD thesis. The University of Edinburgh. UK.

Hinich, Melvin J. (1982). Testing for gaussianity and linearity of a stationary time series. *Journal of time series analysis* **3**, 169–176.

Horch, Alexander (1999). A simple method for detection of stiction in control valves. *Control Engineering Practice* **7**, 1221–1231.

Kim, Y. C. and E. J. Powers (1979). Digital bispectral analysis and its applications to nonlinear wave interactions. *IEEE Transactions on Plasma Science* **PS-7**, 120–131.

Nikias and A. Petropulu (1993). *Higher-Order Spectra: A nonlinear signal processing frame work*. Prentice Hall. New Jeresy.

Rao, T. S. and M. M. Gabr (1980). A test for linearity and stationarity of time series. *Journal of time series analysis* **1**(1), 145–158.

Rosenblatt, M. and J. W. Van Ness (1965). Estimation of the bispectrum. *Ann. Math. Stat.* pp. 420–436.

Stuart, Alan and J. Keith Ord (1987). *Kendall's Advanced Theory of Statistics*. Vol. 1. 5 ed.. Charles Griffin and Company Limited. London.

Yuan, J. (1999). Testing linearity for stationary time series using the sample interquartile range. *Journal of Time Series Analysis* **21**, 713–722.

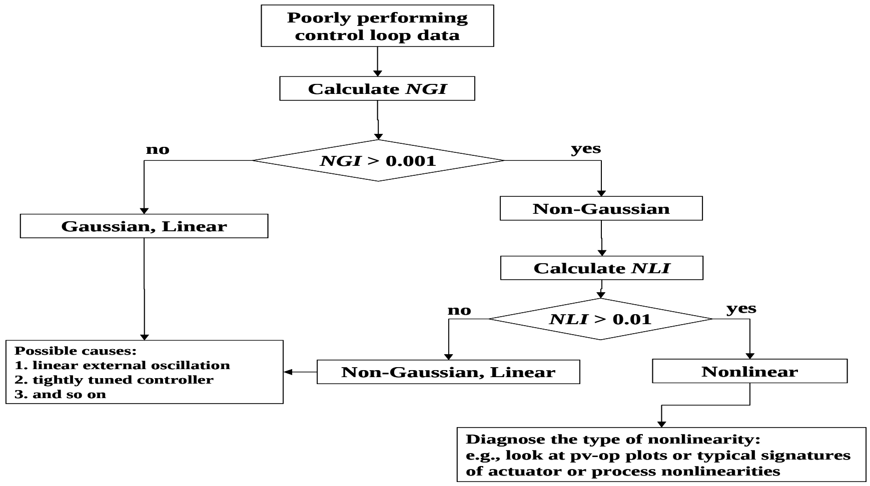


Figure 1: *Rule based decision flow diagram*

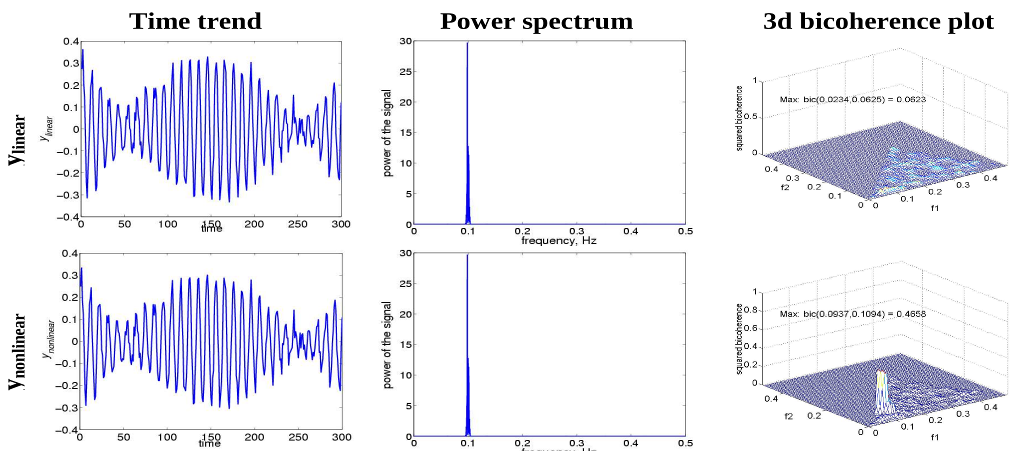


Figure 2: *Higher order statistical analysis of two signals: y_{linear} (top) and $y_{nonlinear}$ (bottom)*

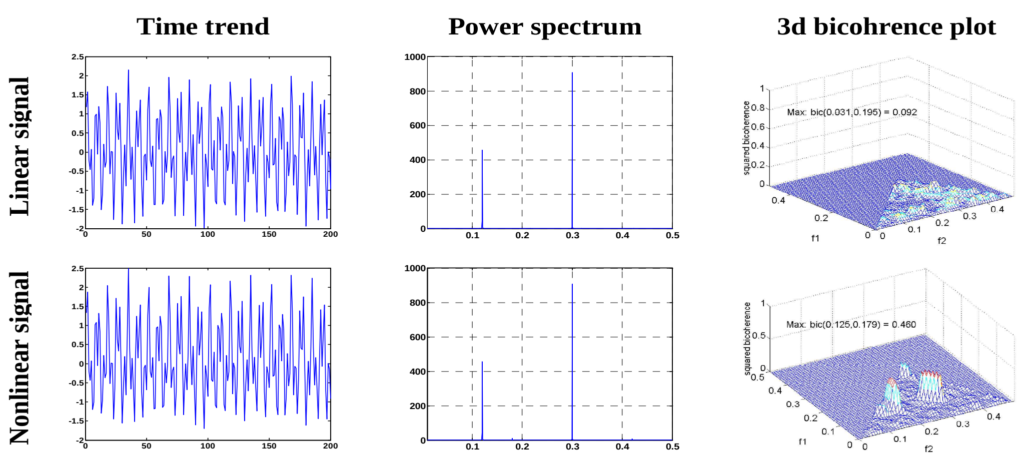


Figure 3: *Results for the linear and nonlinear sinusoid signals*

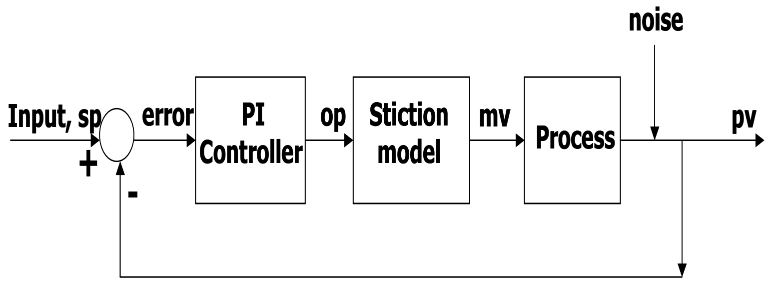


Figure 4: *Block diagram of a simple SISO process under feedback control*

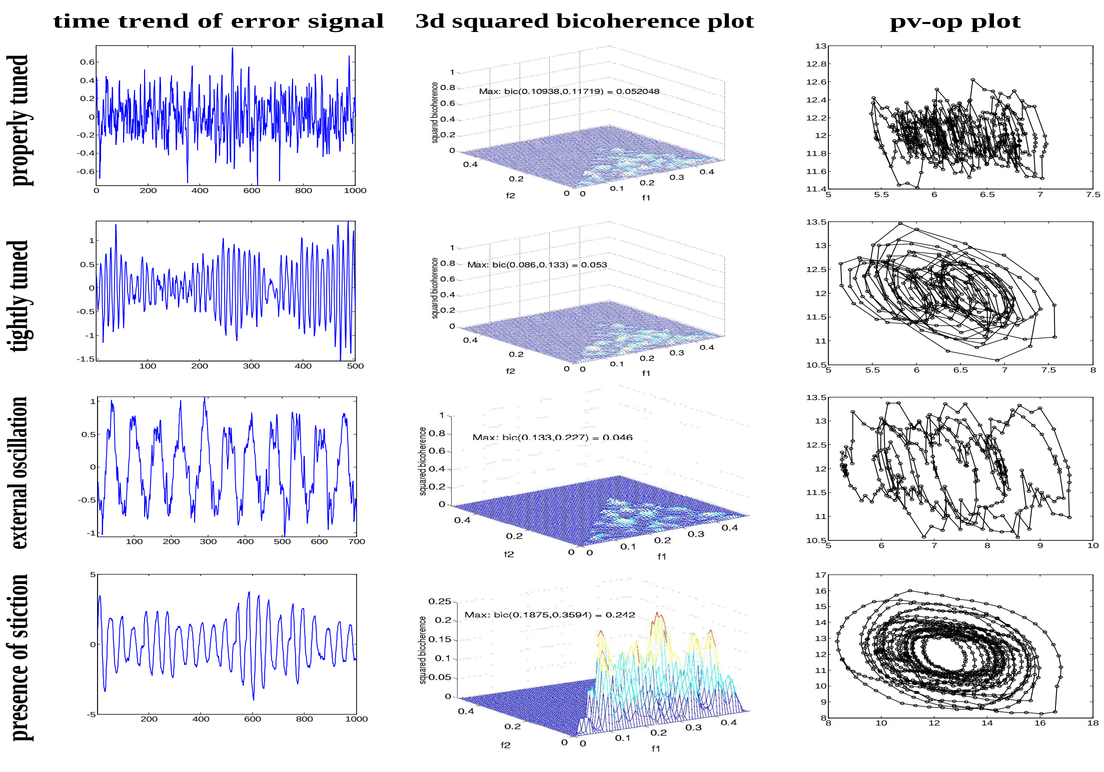


Figure 5: *Higher order statistical analysis of simulated process under feedback control*

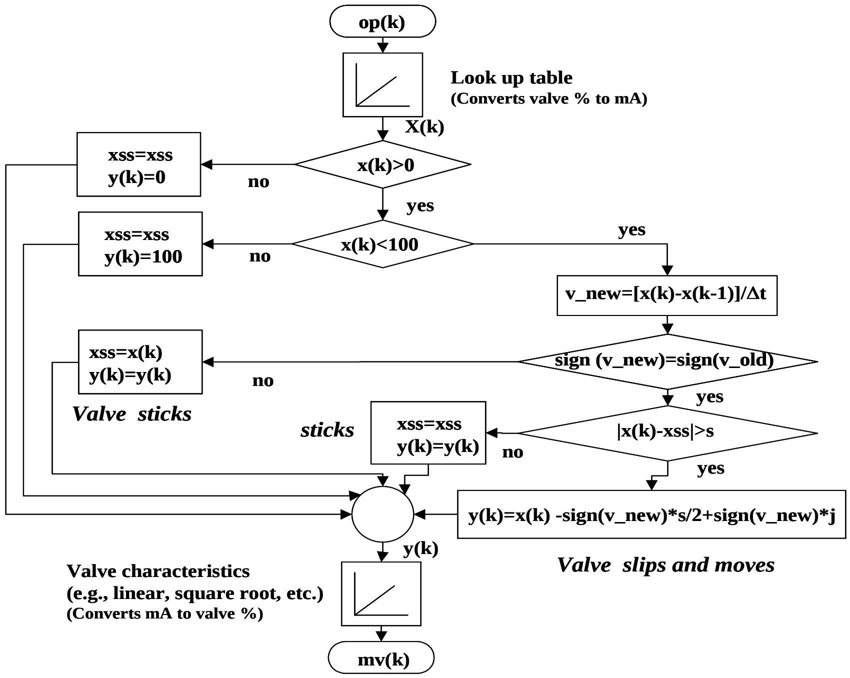


Figure 6: *Flow chart for data-driven stiction model*

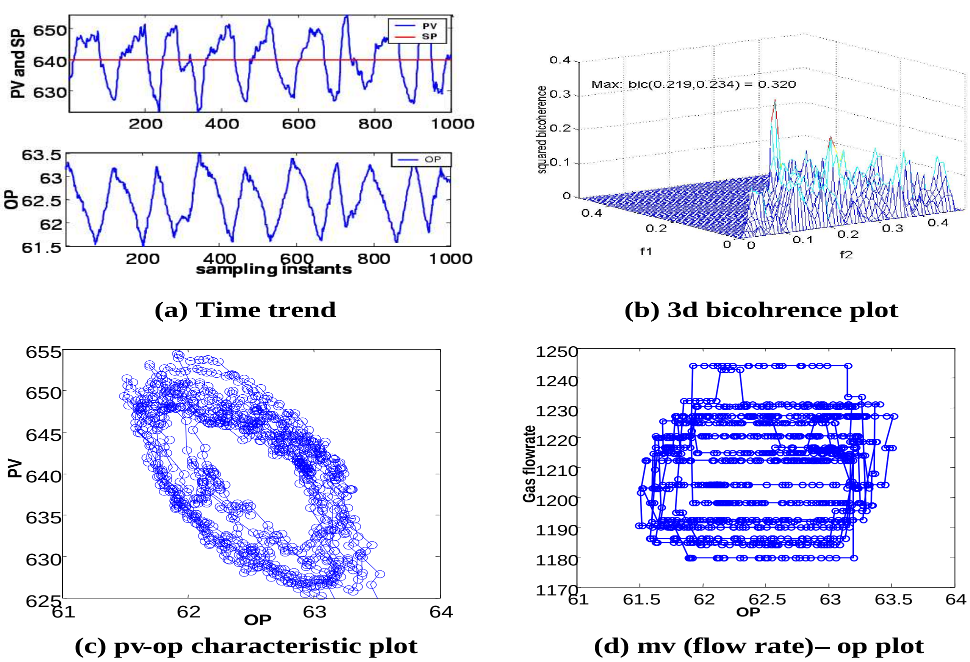
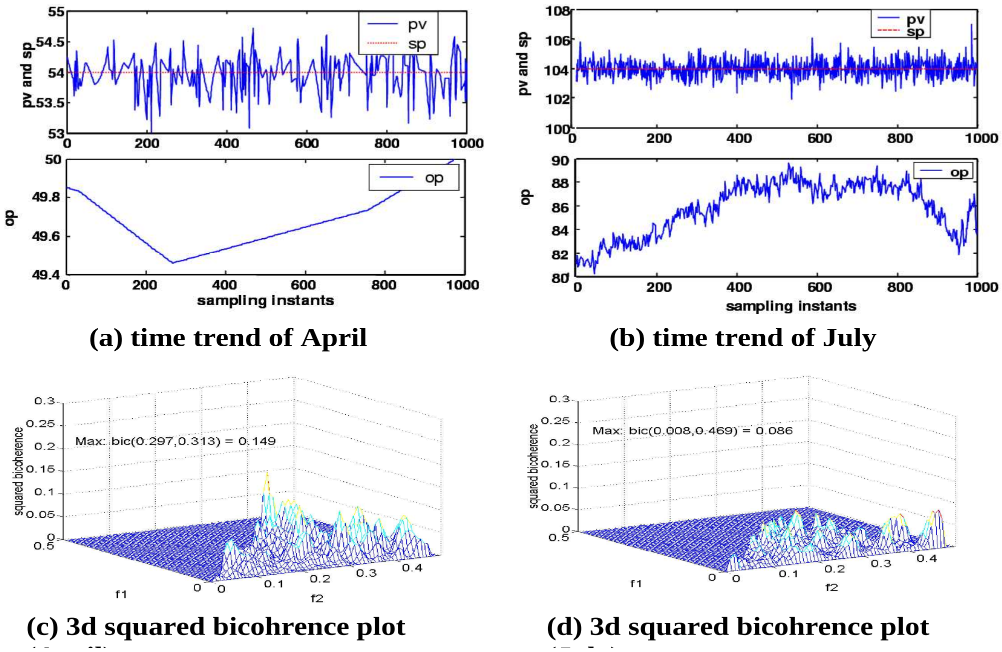


Figure 7: *Analysis of time series data from an industrial temperature control loop*



(a) time trend of April **(b) time trend of July**

(c) 3d squared bicohrence plot **(d) 3d squared bicohrence plot**

Figure 8: *Analysis of flow loop 1 data before (April 2001 - left) and after the (July 2001 - right) plant maintenance shutdown period*

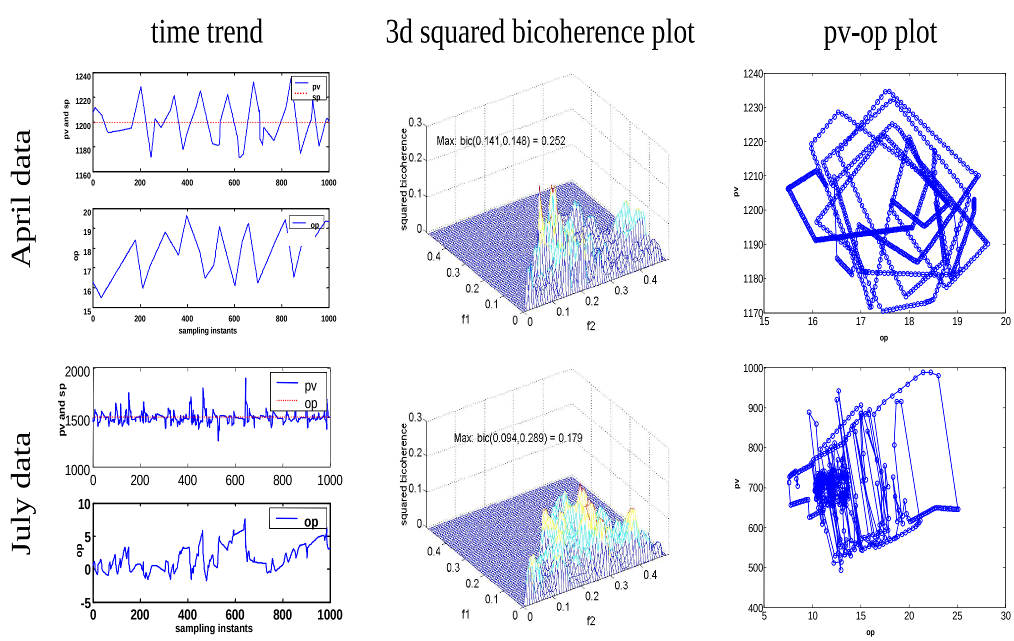


Figure 9: *Analysis of flow loop 2 data before (April 2001 - top) and after (July 2001 - bottom) the plant maintenance shutdown period*