An Integrated Application of Control Performance Assessment and Root Cause Analysis in Refinery Control Loops

Mehmet Yağcı * Yaman Arkun **

* Turkish Petroleum Refineries Co. (Tüpraş) R&D Center, Güzelyurt Mah. Petrol Cad. 41790 Korfez, Kocaeli, TURKEY (e-mail: mehmet.yagci@tpuras.com.tr).

** Department of Chemical and Biological Engineering, Koc University, Rumeli Feneri Yolu, 34450 Sarıyer, İstanbul, TURKEY (e-mail: yarkun@ku.edu.tr)

Abstract: Assessing the performance of control loops is an important component of Control Performance Monitoring (CPM) systems. Most of the industrial chemical processes have a large number of control loops interacting with each other in a complex way due to material and energy integration in the plant. A problem occurring in a certain control loop can easily upset the performance of the other control loops. Therefore, identification of the “bad” control loops causing a plant-wide disturbances is a crucial task. In this work, an integrated approach covering performance assessment and interaction analysis is proposed to detect the “bad” loops based on their performances. First, Minimum Variance Control (MVC) benchmark is used to screen-out the poor performing loops. Then, the spectral envelope method utilizing frequency analysis is used to identify the common oscillation periods among the loops under study. Finally, Granger causality is used to plot the interaction map between the loops. Even though these methods are well developed and used for several purposes separately, we present an integrated approach which focuses and analyzes the “bad loops”. The developed approach has been tested in a refinery plant having 18 control loops. The results show that the proposed method is clearly able to identify and isolate the root-cause control loops. The validation of results and further improvements in the control loops under study have been given.

Keywords: control loop monitoring, performance assessment, diagnosis, root-cause analysis

1. INTRODUCTION

With the advances in computer sciences and electronics, the aim of process control systems has evolved to full automatic systems with as little user intervention as possible. Although in the past, the objective was to control only the operational parameters, today advanced optimization and control systems such as real-time optimizers (RTO), model predictive controls (MPC) have been developed to control the processes in the most economic and safest way. Fig. 1 shows these components under state-of-the-art control systems hierarchy. Considering the intense information flow between layers, the health of each layer is crucial for consecutive control systems. Control Performance Monitoring (CPM) concept is the one generally used to do a health-check for base layer regulatory control.

It is well known fact that performance of controllers degrades over time due to raw material changes, operating point/range changes, decrease of lifetime of equipment, process modifications, hardware/software failures, delays and non-linear process dynamics (Jelali, 2013). A survey conducted by Paulonis and Cox (2003) covering 14000 PID controllers in 40 plants at 9 sites worldwide shows that 41% of the loops were below good performance. Another survey done by Torres et al. (2006) covering 700 control loops belonging to 12 different companies (petrochemical, pulp and paper, cement, chemical, steel and mining) demonstrates that 41% of the loops had oscillations due to tuning problems, coupling, disturbances and actuator problems. In addition, 24% of the loops had saturation problems. In order to address these problems, CPM covers a wide range of sub-components such as performance assessment, oscillation detection, non-linearity analysis, stiction detection and plant-wide disturbance analysis (or also known as root cause analysis).

Among all the problems, oscillations are one of the most important symptoms of inappropriate controller tuning, stiction, external disturbances and interactions (Jelali, 2013). Oscillations are defined as the signals having a period and amplitude. They can be detected in time series process data as periodic signals. However, usually these signals cannot be detected visually and require some signal processing techniques such as power spectrum or autocorrelation function analysis. In some cases, the amplitudes of oscillations are quite high and can be detected visually by process operators in plants. In these cases, the loops are even put in manual mode and taken out of the operation (Ordy et al., 2006). However, these oscillations tend to propagate into the plant and affect the other process variables resulting in performance degradation of control
to get the interactions. However, if the number of loops to be analyzed is quite large, Granger causality will try to analyze all possible pair-wise interactions ignoring their performances. Considering that, the proposed approach in this paper gives a concrete and efficient way to do it. The following sections 2 and 3 give an overview of the components used. Section 4 describes the integrated approach and a use-case in a refinery unit and finally section 5 gives the conclusion of the use-case.

2. PERFORMANCE ASSESSMENT

Performance assessment can be considered as the heart of CPM, because identification of "bad loops" is a crucial need since hundreds or even thousands of control loops are involved in industrial operations. In the past, researchers have tried to develop different benchmarks for performance assessment. Hugo (2001) summarized the possible standards with respect to output variances resulting from various controllers such as perfect control, best possible nonlinear control, minimum variance control, best possible MPC control, best possible PID control and open loop control. Among all the benchmarks, minimum variance control (MVC) is most widely accepted and used. MVC was first developed by Åström (1970) and later improved by Harris (1989) using closed-loop data. There are several approaches to utilize MVC as a benchmark. These are Auto-Regressive Least Square (AR-LS) model by Desborough and Harris (1992) and Filtering and Correlation (FCOR) analysis by Huang and Shah (1999). In this work, AR-LS method has been used.

Theoretically, an MVC loop does not show any variance on the control error after dead-time, $k$. Under MVC, one has the following output ($y$)-disturbance ($\epsilon$) relationship:

$$y(t + k) = F\epsilon(t + k)$$  \hspace{1cm} (1)

where $F$ stands for the moving average component of closed loop process model. Desborough and Harris (1992) introduced a normalized performance index (CPI) between 0 and 1:

$$CPI = \frac{\sigma_{MVC}^2}{\sigma_{y}^2}$$  \hspace{1cm} (2)

where lower CPI values indicate poorer performance. $\sigma_{MVC}^2$ and $\sigma_{y}^2$ shows the theoretical minimum variance and actual variance, respectively. Since MVC is not exactly achieved in real control systems, Eq. 1 is rewritten as

$$y(t + k) = F\epsilon(t + k) + \hat{y}(t)$$  \hspace{1cm} (3)

Desborough and Harris (1992) suggested an AR model for $\hat{y}$; thus, Eq. 3 becomes:

$$y(t) = F\epsilon(t) + \sum_{i=1}^{m} \alpha_i y(t - k - i + 1)$$  \hspace{1cm} (4)

where $\alpha_i$ and $m$ stand for AR model coefficients and model order, respectively. Then, the minimum achievable variance ($\hat{\sigma}_{MVC}^2$) becomes the residuals of AR model. Finally, the performance of a controller (CPI) can be estimated such that:

$$CPI = \frac{\hat{\sigma}_{MVC}^2}{\sigma_{y}^2}$$  \hspace{1cm} (5)
3. PLANT-WIDE DISTURBANCE DETECTION

3.1 Spectral Envelope

Spectral envelope is based on finding similarities in different time series variables. The method first used in categorical time series datasets by Stoffer et al. (1993). In this approach, using Fourier transform, common frequencies of a bulk data have been extracted. Later, McDougall et al. (1997) extended the proposed approach to real-valued time series. Spectral envelope derives optimal coefficients to extract common frequencies in a bulk dataset. Having that feature, one can easily analyze different signals having different order of magnitudes. Jiang et al. (2007) used that approach and developed a method to find oscillating variables and the root cause of the oscillations. Later, this method has been used by several researchers to detect plant-wide disturbances (Duan et al., 2014; Xu et al., 2016). Spectral envelope is defined as:

$$\lambda(\omega) = \sup \left\{ \frac{\beta^* P_X(\omega) \beta}{\beta^* \beta} \right\}$$

(6)

where \(\lambda(\omega)\) represents the largest portion of signal at frequency \(\omega\). \(\beta\) represents the optimum scaling vector, \(V_X\) represents the covariance matrix for \(X\) and \(P_X\) represents the power spectral density matrix for \(X\). There, \(\sup\{\cdot\}\) stands for the supremum operator. Given these, Eq. 6 represents the largest portion of signal in the power spectrum (PSD). Stoffer (1999) states that the term "envelope" represents the enclosure placed onto a specific frequency at spectrum. Due to that, first PSD must be estimated with input vector defined as:

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_z(t) \end{bmatrix}, \quad X = [x(0) \ x(1) \ldots \ x(N-1)]$$

(7)

where each row in Eq. 7 represent a set of normalized time domain data belonging to each variable (1 \(\rightarrow\) z). Normalization is done by subtracting mean of each variable and dividing by the standard deviation. It is an important step because normalization makes comparison between variables and estimation of envelope easier. After having normalized input data, the periodogram is estimated as:

$$\hat{I}_X(\omega_f) = \frac{1}{N} \left[ \sum_{t=0}^{N-1} x(t) \exp(-2\pi it\omega_f) \right]^* \left[ \sum_{t=0}^{N-1} x(t) \exp(-2\pi it\omega_f) \right]$$

(8)

In Eq. 8, * represents conjugate and transpose operation. In order to smooth the periodogram estimated, moving average filter is applied such that:

$$\hat{P}_X(\omega_f) = \sum_{j=-r}^{r} h_j \hat{I}_X(\omega_f+j)$$

(9)

PSD of continuous signal is estimated by discrete Fourier transform (DFT). However, maximum frequency that DFT can resolve is the half of Nyquist frequency (Salkind, 2006). Due to that, the maximum frequency is limited to total time period constructed from half of the all data points. For example, for \(N = 500\) data points sampled at \(Ts = 5\) second, maximum frequency is \(500/2 \times 5 = 1250\) second. Then, DFT can resolve frequencies \([0, Ts/2, Ts, 3Ts/2, \ldots, N \times Ts/2]\), \(\omega_f\) in Eq. 9 represents those discrete frequencies. After that, to smooth PSD, the coefficients \(h_j\) are selected such that:

$$\sum_{j=-r}^{r} h_j = 1$$

(10)

Satisfying Eq. 10, coefficients can be selected as \(h_j = (r - |j| + 1)/(r + 1)^2\), for \(|j| = 1, 2, \ldots, r\). Number of smoothing coefficients \(r\) is chosen depending on the desired smoothness. Selecting \(r = 2\) results in \(h_0 = 3/9, h_1 = 2/9, h_2 = 1/9\), which is enough in most of the cases. One can also increase the degree to higher numbers but should be careful about smoothing the actual frequency peaks (Stoffer et al., 2000). Then one can estimate \(\lambda(\omega)\) using \(\hat{P}_X(\omega_f)\) and \(\hat{V} = \text{diag}(\hat{V}_X)\) by rewriting Eq. 6 as:

$$\hat{\lambda}(\omega_f) = \sup \left\{ \frac{\beta^* \hat{P}_X(\omega_f) \beta}{\beta^* \beta} \right\}$$

(11)

From Eq. 11, it is clear that \(\hat{\lambda}(\omega_f)\) is the largest portion of signal at specific frequency \(\omega_f\). Then, the eigenvalue is estimated by determinant equation of:

$$|P_X(\omega) - \lambda(\omega)V_X| = 0$$

(12)

Then, optimal scaling vector \(\beta(\omega)\) can be estimated as:

$$\lambda(\omega)V_X\beta(\omega) = P_X\beta(\omega)$$

(13)

Since all the variables are normalized before, \(\hat{\beta} = \text{diag}(1, 1, \ldots, 1)\). Thus, Eq. 13 reduces to \(\beta^* \hat{\beta} = 1\). At the end, one can easily state that \(\hat{\lambda}(\omega_f)\) is the largest eigenvalue of \(\hat{P}_X(\omega_f)\), and \(\hat{\beta}(\omega_f)\) is the corresponding eigenvector. Finally, one can estimate the spectral envelope for each frequency. Fig. 2 shows a bank of simulated signals having 12 variables oscillating at frequencies 0.1 Hz and 0.3 Hz, adopted from Jiang et al. (2007).

After detection of common frequencies, one can determine which variables are oscillating at those frequencies. To that end, Stoffer (1999) proposed applying statistical hypothesis testing on \(\hat{\beta}(\omega_f)\). He states that if \(\hat{\lambda}(\omega_f)\) has a distinct root, for large number of observations \(v_N [\hat{\beta}(\omega) - \beta(\omega)]\) converges to multivariate normal distribution and
\[ v_N \left( \lambda(\omega) - \lambda(\omega) \right) / \lambda(\omega) \text{ converges to standard normal distribution. If proposed moving average smoothing in Eq. } 10 \text{ is used, } v_N \text{ becomes:} \]

\[ v_N = \left( \sum_{j=-r}^{r} h_j^2 \right)^{1/2} \quad (14) \]

Then, one can write the asymptotic covariance matrix of the optimum scaling vector \( \hat{\beta}(\omega) \) as:

\[ \mathbf{V}_{\beta} = v_N^{-2} \lambda_1(\omega) \sum_{i=2}^{z} \lambda_i(\omega) [\lambda_1(\omega) - \lambda_i(\omega)]^{-2} \beta_i(\omega) \beta^*_i(\omega) \quad (15) \]

In Eq. 15, \( \lambda_1(\omega) = [\lambda(\omega), \lambda_2(\omega), \ldots, \lambda_z(\omega)] \) are the eigenvalues of \( \mathbf{P}_X(\omega) \) in decreasing order and \( \beta_i(\omega) = [\beta(\omega), \beta_2(\omega), \ldots, \beta_z(\omega)] \) are the corresponding eigenvectors.

Then, the statistical test on \( \hat{\beta}(\omega) \) can be written as:

\[ \text{Test Stat}_i(\omega) = \frac{2 |\beta_{1,i}(\omega) - \beta_{i,i}(\omega)|^2}{\sigma_i(\omega)} \quad (16) \]

where \( i = 1, 2, \ldots, z \) and \( \sigma_i(\omega) = \text{diag}(\mathbf{V}_{\beta}) \).

Stoffer (1999) states that defined \( \text{Test Stat}_i(\omega) \) in Eq. 16 converges to a Chi-squared (\( \chi^2 \)) distribution with 2 degrees of freedom. Then, for a certain confidence level \( (cL) \), if \( \text{Test Stat}_i(\omega) > \chi^2_{2,cL} \), the variable \( i \) is said to be oscillating at frequency \( \omega \). One can select \( \chi^2_{0.01,2} = 13.816 \) for \( cL = 99.9 \%) or \( \chi^2_{0.01,2} = 9.210 \) for \( cL = 99 \% \).

Detection of common frequencies and determining the oscillating signals at those frequencies give valuable information about the analyzed multivariate data. By using spectral envelope, one can detect plant-wide oscillations. However, locating the source of plant-wide disturbance is crucial since the units in the plants are connected each other by recycling streams, heat integrations and even simple upstream and downstream connections, problems arising in a specific location propagates and affects other loops. With that goal, based on spectral envelope, Jiang et al. (2007) proposed OCI (oscillation contribution index) which gives an indication about the root cause of those oscillations. OCI represents the relative contribution of a variable to that specific frequency and can be calculated as:

\[ \text{OCI}_i(\omega) = \frac{\hat{\beta}_{1,i}(\omega)}{2\sigma_{\beta}(\omega)} \quad (17) \]

The term \( \sigma_{\beta}(\omega) \) in Eq. 17 represents the standard deviation of all optimal scaling coefficients at a certain frequency \( \omega \) for all variables. Jiang et al. (2007) state that variables having \( \text{OCI}_i(\omega) > 1 \) are the most probable causes of plant-wide oscillations at frequency \( \omega \). They also state that in case of no root cause, OCI values will be less than one. However, in such cases, OCI can be beneficial to sort variables according with respect to their contributions at frequencies identified in the spectral envelope.

### 3.2 Granger Causality

Granger causality was first introduced by Granger (1969) in order to assess the causality between different stock indices. Basically, Granger causality states that if including the past observations of variable \( x_1 \) into linear prediction of variable \( x_2 \) yields better model for \( x_2 \), then \( x_1 \) causes \( x_2 \). Granger causality uses bivariate AR models such that:

\[ x_1(t) = \sum_{j=1}^{k} A_{11,j} x_1(t-j) + \sum_{j=1}^{k} A_{12,j} x_2(t-j) + \epsilon_{1,2}(t) \quad (18) \]

\[ x_2(t) = \sum_{j=1}^{k} A_{21,j} x_1(t-j) + \sum_{j=1}^{k} A_{22,j} x_2(t-j) + \epsilon_{2,1}(t) \quad (19) \]

where \( k \) is the model order, \( A \)'s are the AR model coefficients and \( \epsilon \) represents the model residuals. The models represented by Eq. 18 and 19 are also called as unrestricted model. Then, the restricted models are obtained such that:

\[ x_1(t) = \sum_{j=1}^{k} B_{1,j} x_1(t-j) + \epsilon_1(t) \quad (20) \]

\[ x_2(t) = \sum_{j=1}^{k} B_{2,j} x_2(t-j) + \epsilon_2(t) \quad (21) \]

Then one can say, if \( \epsilon_{1,2} < \epsilon_2 \), \( x_1 \) causes \( x_2 \). The significance of interaction can be measured by:

\[ F_{i \rightarrow j} = \frac{\text{var}(\epsilon_{j})}{\text{var}(\epsilon_{j,1})} \quad (22) \]

The model order can be determined by using Akaike information criteria (AIC) or Bayesian information criteria (BIC) over a set of possible model orders. Using that procedure, all pairs of variables are analyzed and reported.

### 4. APPLICATION OF METHODS TO REFINERY CONTROL LOOPS

Two components, performance assessment and root cause analysis, have been applied to a plant in Izmit Refinery of Turkish Petroleum Refineries Corporation. The plant has 18 loops controlling mainly a reactor and a distillation column connected to each other. FC, FFC, PC, LC and TC controllers, respectively. The loops are sampled for 15 seconds which results in 5760 sample points for each day. The suggested application procedure is as follows:

1. The performance of each control loop is calculated by AR-LS method. The dead-times of the loops are determined with operators and/or process control engineers. For the model order, 30 can be selected as suggested by Thornhill et al. (1999).
2. The control loops having CPI less than 0.75 are identified as poor performing loops and filtered for root-cause analysis.
3. Spectral envelope method is applied to poor performing loops. PSD smoothing coefficient \( r \) can be chosen as 2. Statistical tests can be assessed for 99.9\% confidence level, having \( \chi^2_{0.01,2} = 13.816 \). This makes sure that most of the oscillating loops are isolated. The maximum common period is obtained from the envelope plot.
4. Granger causality is applied to same list of poor performing loops. The maximum lags is selected as the multiple of sampling period which equals to maximum common period obtained in step 3.
(5) The interactions are plotted.
(6) The results are validated with operators and/or process control engineers.

Step 1 & 2: Table 1 shows the results of performance assessment. The loops having CPI less than 0.75 are given on the left side.

Table 1. Results of Performance Assessment

<table>
<thead>
<tr>
<th>Loop Name</th>
<th>CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC3</td>
<td>0.3100</td>
</tr>
<tr>
<td>PC3</td>
<td>0.3647</td>
</tr>
<tr>
<td>FC9</td>
<td>0.4258</td>
</tr>
<tr>
<td>PC1</td>
<td>0.4719</td>
</tr>
<tr>
<td>FFC1</td>
<td>0.5106</td>
</tr>
<tr>
<td>TC1</td>
<td>0.5338</td>
</tr>
<tr>
<td>FFC2</td>
<td>0.5464</td>
</tr>
<tr>
<td>TC3</td>
<td>0.6929</td>
</tr>
<tr>
<td>LC1</td>
<td>0.7217</td>
</tr>
<tr>
<td>FC1</td>
<td>0.8552</td>
</tr>
<tr>
<td>PC5</td>
<td>0.8594</td>
</tr>
<tr>
<td>FC3</td>
<td>0.8651</td>
</tr>
<tr>
<td>PC2</td>
<td>0.8973</td>
</tr>
<tr>
<td>PC2</td>
<td>0.8999</td>
</tr>
<tr>
<td>FC8</td>
<td>0.902</td>
</tr>
<tr>
<td>FC7</td>
<td>0.9333</td>
</tr>
<tr>
<td>FC4</td>
<td>0.9637</td>
</tr>
<tr>
<td>FC6</td>
<td>0.9831</td>
</tr>
</tbody>
</table>

Step 3: Fig. 3 shows the envelope for poor performing loops. It is seen that the maximum oscillation period is 877 second, approximately 60 samples for 15 second sampling interval. Similarly, Table 2 shows OCI values for spectral envelope analysis. The table indicates that the main causes of plant-wide oscillations are TC1 and LC1.

Step 4 & 5: Finally, Fig. 4 shows the interaction plots with F scores between the poor performing loops. As expected from spectral envelope analysis, loops TC1 and LC1 are the main actors of plant-wide oscillations. The most affected loop is seen as PC1.

After reporting the root-causes and interaction plots between loops, the worst loops TC1 and LC1 were examined.

Table 2. Spectral Envelope Contribution Index

<table>
<thead>
<tr>
<th>Loop Name</th>
<th>CPI</th>
<th>OCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC1</td>
<td>0.5338</td>
<td>2.05</td>
</tr>
<tr>
<td>LC1</td>
<td>0.7217</td>
<td>1.79</td>
</tr>
<tr>
<td>PC1</td>
<td>0.4719</td>
<td>0.95</td>
</tr>
<tr>
<td>TC3</td>
<td>0.6929</td>
<td>0.90</td>
</tr>
<tr>
<td>FC9</td>
<td>0.4258</td>
<td>0.45</td>
</tr>
<tr>
<td>FFC1</td>
<td>0.5106</td>
<td>0.43</td>
</tr>
<tr>
<td>LC3</td>
<td>0.3100</td>
<td>0.21</td>
</tr>
<tr>
<td>PC3</td>
<td>0.3647</td>
<td>0.18</td>
</tr>
<tr>
<td>FFC2</td>
<td>0.5464</td>
<td>0.01</td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

Oscillations in control loops are the most common reasons of performance degradations. Often these oscillations propagate through the other loops due to upstream-downstream and recycle connections between control loops. In complex industries like refineries, detection and isolation of the main bad loops is very important to avoid oscillation problems. There are several methods in the literature to detect these oscillations. However, most of the methods suffer from automatic detection of root causes. Increasing computing power enables different methods to emerge and analyze plant-wide disturbances in efficient ways. For example, in this study, spectral envelope method utilizes 5760×18 matrix to estimate TestStat for 2880 different frequencies; whereas, Granger causality builds 10260 bivariate and univariate AR models for 18 variables and 60 different discrete dead-time values. Therefore, both methods require relatively good computing power.

In general, although spectral envelope and Granger causality methods seem to be separate approaches to detect plant-wide disturbances; here they were used as complementary methods to deeply characterize the plant-wide disturbances. Spectral envelope is a powerful method to isolate common frequencies for oscillating loops, however it suffers from detecting interactions between loops. In this work, the common oscillating loops have been identified by spectral envelope method and interactions have been captured by Granger causality. The suggested procedure that is integrated with performance assessment techniques
enables efficient calculation effort and reliable detection of plant-wide disturbances.

Plotting the interaction map between poor performing loops is not the final step in this process. Next, the identified loops should be examined with plant or process control engineers. The control loops might suffer from different causes of oscillations such as control valve stiction or excessive PID gains. Today, most of CPM software have stiction detection/characterization and PID tuning analysis covering from aggressiveness/sluggishness tests to optimal parameter estimation. An extensive discussion for CPM components and their applications to a refinery unit has been given in Yağcıl (2016).

ACKNOWLEDGEMENTS

The authors acknowledge the support of The Scientific and Technological Research Council of Turkey (TÜBİTAK) within the frame of project number 3150984. The authors thank to Tüpraş Izmit Refinery Advanced Process Control Team for their valuable comments and discussions.

REFERENCES


