

Loop Status Monitoring and Fault Localisation

CHUNMING XIA AND JOHN HOWELL

*Department of Mechanical Engineering
University of Glasgow, Glasgow G12 8QQ, Scotland, UK
Email: j.howell and chunming @mech.gla.ac.uk*

Abstract

Loop status monitoring involves the declaration of deterministic trends, such as oscillations and drifting, in loops that are in multi-loop plant configurations. By analysing various time domain statistics pertaining to PI/PID control loops, a trend can be recognised as one of seven categories. The scientific basis for working with the particular statistics is given and the categorisation process is described. These statistics can be combined to produce an *Overall Loop Performance Index* for each loop, which can be compared to localise a single fault in a multi-loop arrangement. Estimation methods for these time domain statistics are outlined and the performance of the approach is demonstrated on both simulated and real plant data sets.

Keywords: closed loop assessment, process control, detection and isolation

1. INTRODUCTION

Although loop performance monitoring has been well-researched, especially minimum variance benchmark based techniques [1-3], the emphasis has been on audit, in a sense that the underlying performance is of issue, rather than on what the loop is actually doing at any particular time. When the focus has been on the actual trend itself, workers have looked at the shape of the trend [4-8] with the aim of diagnosing faults or disturbances rather than of determining loop performance. Control loop status monitoring can be defined as the near-real-time declaration of what a loop is actually doing at that time [9,10]. The status might indicate that the loop is (1) well-behaved & in steady state, or (2) well-behaved but with controller compensation, or (3) undergoing a short-term transient, or (4) undergoing a trend that is disturbed in some non-stationary manner, or (5) cycling at a relatively low fundamental frequency, or (6) cycling at a fundamental frequency similar to the natural frequency of the loop, or (7) out of control (Critical). For instance category (4) might arise because of feedstock variability whilst category (5) might arise because of a cyclic disturbance, or of a valve problem involving a phase lag. Transients can be classed as either short-term or long-term depending on whether or not the transient lasts longer than the loop's settling time. This is an indicator of whether the response is caused solely by a short-term event, like a change in operating point, or by something more sustained. From a performance point of view the long-term is more important. Some of the proposed loop statuses might be of interest to the operator, whilst others might be of more interest to the control/maintenance engineer. Some might be of use when analysing data off-line. For instance the maintenance engineer would like to know when a loop is limit cycling, whereas someone analysing off-line data would like to know if the plant was operating in an unsteady manner. The benefit might be as much to do with providing information about plant operation, as with understanding the performance of an individual control loop. This would particularly be so, for instance, if the controllers were compensating for a leak.

The statuses of individual loops cannot be viewed in isolation, however, because in certain circumstances a number of loops might be affected by the degradation of one loop or by a plant-wide disturbance. A single loop might cause an entire plant to oscillate, but only whilst operating in one particular regime. Both the operator and control/maintenance engineer would be interested in not only locating the problem

loop, but also in doing so whilst the situation still exists on the plant, since this would enable an engineer to experiment with controller settings. Although significant work has been presented [1, 2] on the extension of closed loop performance assessment methods to multivariable systems, the diagnosis of multivariable systems has not been thoroughly investigated [1].

This paper proposes various statistics to facilitate the status monitoring of PI/PID loops and to isolate a problem loop. These statistics can be complemented by the control loop performance monitor of Hägglund [11], which focuses on Statuses (5) & (6), and by the automatic detection of sluggish control loops [12], which can be performed during Status (3). Although loop status is a qualitative description, a quantitative version can be formulated by assigning different real number values to the various categories. The quantitative statistic can then be combined with the other statistics to form an index to isolate a problem loop in a number of interacting loops. A particular combination has been formulated, the *overall loop performance index (OLPI)*, which has been designed to increase as loop performance deteriorates. A problem loop can then be isolated by finding that loop with the highest OLPI. The identification of the root cause of a plant-wide disturbance, such as plant-wide oscillations, is of importance to both auditors and operators. Although there has been a great deal written about the detection/diagnosis of a single loop and about specific oscillation problems [11, 13-15], few have attempted to deal with plant-wide detection and diagnosis. Thornhill *et al.* have applied principal component analysis to power spectra with the aim of detecting groups of oscillatory process measurements [16], and have also examined the detection of multiple oscillations in control loops [17].

By focusing on ‘current’ loop status rather than on an audit of loop performance, this paper can be viewed as a step towards Kozub’s vision. Kozub [18] has compared experiences with time series based approaches with those with single statistic based approaches. He observed that although the use of a single number statistic can hardly be as effective as analysis tools that offer far more detail, the latter are very resource intensive whilst a group of single number statistics would be useful to provide a first pass indication that a problem exists. Kozub sees the potential for both detecting and determining the extent of under-damped, or cyclical, response trend characteristics and

views the automation of controller performance monitoring and diagnosis as an important new challenge.

This paper has three main sections. Section 2 gives the formulation and scientific basis, loop status monitoring is then developed in Section 3 and the extension to fault localisation is considered in Section 4. Detailed derivations of equations have been placed in appendices. The method has been tested on a number of applications, both real and simulated. Those pertaining to Eastman plants are discussed in Section 4.

2. FUNDAMENTALS

This section describes a set of statistics, which, when examined together, provide a means of deciding the Loop Status of either a single PI/PID loop, or the master loop of a cascade PI/PID-based control system. An explanation is then given of how these statistics might be estimated from plant data.

2.1 The Model

Fig. 1 represents a control loop by means of a linear block diagram. Blocks P & H represent the process and controller transfer functions, u_t is the controller output, y_t is the controlled variable and d_t is a deterministic disturbance, which is modulated with additive Gaussian white noise, e_t , that is independent of d_t . Non-linear effects caused by, for instance, hysteresis, are accommodated by representing them as 'pseudo' external deterministic disturbances. Note that the set-point is omitted i.e. it is assumed that the loop is a regulator. The approach does accommodate infrequent changes to the set-point, and such changes are seen as additional disturbances to d_t .

The status of a loop can be described by the deterministic, or underlying, trend that is observed. The trend of interest is therefore the deterministic component of time series y_t . If \hat{y} is defined as the deterministic component of y_t , and e_y is defined as the direct effect of the noise on y_t , then by definition

$$y_t = \hat{y} + e_y = \frac{1}{1+HP} [d_t + e_t], \quad (1)$$

which can be separated into deterministic and stochastic components:

$$\hat{y} = \frac{1}{1+HP} d_t, \quad e_y = \frac{1}{1+HP} e_t \quad (2)$$

Similarly the controller output time series can be represented by

$$u_t = \underbrace{(-H\hat{y})}_{\hat{u}} + \underbrace{(-He_y)}_{e_u} \equiv \hat{u} + e_u \quad (3)$$

where \hat{u} is the deterministic time series observed in the controller output and e_u is the noise.

Note that e_y is filtered white noise. Fig. 2 shows the frequency response of this filter for a typical PI controlled, first order process plant with time delay: it has a high-pass filtering feature, with a cut-off frequency B_0 , so that e_y will nearly behave as banded

white noise with lower cut-off B_0 and upper cut-off at the *Nyquist* frequency

$B : B = \left(\frac{\pi}{T_s} \right) \text{rads}^{-1}$, where T_s is the sample interval. The precise location of B_0 is not important.

2.2 The Basis

Sinusoidal trends are common in control loops and can be categorised as either Loop Status (2) compensating (if the frequency is very low) or (5) - (6) cycling (if the frequency is either relatively low or near the natural frequency of the loop). An examination of the behaviour of a PI control loop, when subjected to a pure sinusoidal disturbance of frequency ω , therefore provides a basis for understanding the approach. This behaviour is assessed by analysing three loop statistics, η_y , η_u and their ratio R . The behaviour of a normalised version of R , called R_n , is also examined, particularly in relation to its ability to contribute to the categorisation of trends with other Loop Statuses. In particular, its standard deviation, σ_{R_n} , is found to provide discrimination between Statuses (4) and (5). It will be shown that the analysis of statistics η_y , η_u , R_n and σ_{R_n} combined, provide the necessary discrimination between the various statuses of a PI loop. Differences in their behaviour for PID control is also examined.

Definition 1: Signal-to-noise ratio indices η_y and η_u are defined as:

$$\eta_y = \frac{\sigma_{\hat{y}}^2}{\sigma_{e_y}^2}, \quad (4)$$

$$\eta_u = \frac{\sigma_{\hat{u}}^2}{\sigma_{e_u}^2} \quad (5)$$

where $\sigma_{\hat{y}}^2$, $\sigma_{e_y}^2$, $\sigma_{\hat{u}}^2$ and $\sigma_{e_u}^2$ are the variances of \hat{y} , e_y , \hat{u} and e_u respectively.

Statistic η_y will be the same as the minimum variance performance index [1,2], when the predicted, deterministic component (called the *predictable* component) is estimated by a d -step ahead AR model, where d is the time delay. Although in theory a perfect control loop operating in steady state will have an η_y of zero, industry practice is to recognise that small disturbances do arise and to deem a loop to be performing well if

$\eta_y < \gamma$, where γ is typically 0.5 for a temperature loop, but might be smaller for other loops, such as flow loops. A customised value could be obtained by analysing the normal steady state operating data.

Definition 2: R is defined as the ratio of the indices:

$$R = \frac{\eta_y}{\eta_u} \quad (6)$$

2.2.1 PI Control

Theorem 1: If a control loop is under PI control of form $K_p \left[1 + \frac{1}{T_i s} \right]$, where K_p is the gain and T_i is the integral time constant, and the process has a sine wave oscillatory trend of frequency ω and magnitude A , which is modulated by banded noise e_y with uniform power spectrum of amplitude G over the bandwidth $[B_0, B]$, then

$$\eta_y = \frac{\frac{1}{2} A^2}{G(B - B_0)}, \quad (7)$$

$$\eta_u = \underbrace{\left(\frac{1 + \frac{1}{(T_i \omega)^2}}{Q} \right)}_{J(\omega)} \eta_y = J(\omega) \eta_y, \quad (8)$$

$$R(\omega) = \frac{1}{J(\omega)} = \frac{Q}{1 + \frac{1}{(T_i \omega)^2}} \quad (9)$$

where $Q = 1 + \frac{K_B}{\pi^2} \left[\frac{T_s}{T_i} \right]^2$, T_s is the sampling time and $K_B = \frac{B}{B_0}$.

Proof: See Appendix A.1.

Thus there is a simple relationship between η_y and η_u , which can be represented by a meaningful frequency diagram. Also η_y is independent of frequency and constant for a sinusoid of fixed magnitude and stationary noise. The assumption of a uniform spectrum is justified, firstly because G doesn't feature in the final results (Equations 8

& 9), and secondly because $G(B - B_0)$ is the variance of e_y . Fig. 3 shows a typical frequency plot of the three key statistics η_u , η_y & R for a well-performing control loop with $Q=1$, $\eta_y=0.5$ and $T_i=20$. Although parameter Q is dependent on controller settings, for most process control loops $Q \approx 1$, because $\frac{T_s}{T_i} \ll 1$ and K_B is relatively small. The following normalised version of R is preferred, because this will make the formulation of a test more straightforward.

Remark 1: A normalised value of R is defined as:

$$R_n = \frac{R}{Q} = \frac{1}{1 + \frac{1}{(T_i\omega)^2}} \quad (10)$$

Fig. 4(a) shows a plot of R_n versus $T_i\omega$. It is significant that $R_n \geq 0.8$ when $T_i\omega \geq 2$ (or $\omega \geq \frac{2}{T_i}$) because this can be used to discriminate between low and ultimate frequency trends.

Remark 2: The $R_n - \omega$ relationship of Eq. 10 can be extended to more general deterministic oscillatory trends such as the non-sinusoidal periodic oscillations typical of limit cycles, where ω is now the dominant, fundamental angular frequency of the trend, and the result of $R_n \geq 0.8$, when $\omega \geq \frac{2}{T_i}$, still holds (See Appendix A.2).

Fig. 3 shows that for a well performing loop with a constant η_y , ratio η_u behaves quite differently at low frequencies. Consider the case when η_u exceeds γ , whilst at the same time η_y does not (i.e. $\eta_u > \gamma$ & $\eta_y < \gamma$). Here there is a significant predictable signal in the controller out and very little in the controlled variable. This implies that the controller is having to compensate to regulate the controlled variable to setpoint. Loop Status (2) denotes this situation. Incidentally this result confirms ones intuitive belief that a low frequency trend will be more visible in the controller output than in the controlled variable because the controller output will have a higher signal to noise ratio.

Second, the results show that, for any specific controller, the ratio, R_n , will be solely frequency dependent. Equation 10 shows that it is independent of the value of η_y , even if controller performance is poor. This means that R_n can be used to distinguish between certain oscillatory trends because it obtains a value $\ll 1$ when the trend is of low frequency. Hägglund [11] has pointed out that, if a PI controller is properly tuned, then usually the integral time constant approximates to the ultimate oscillation period of the closed loop system i.e. the period of oscillation that occurs when its gain is increased. Substituting $\omega = \frac{2\pi}{T_i}$ into Equation 10 gives $R_n \approx 1$ i.e. R_n will be on the plateau of the R_n versus ω plot if a loop is oscillating at its ultimate frequency. This property can be used to discriminate between an ultimate cyclic oscillation and an oscillatory disturbance of a lower frequency. A conservative test can be formulated based on $R_n > R_n\left(\omega = \frac{2}{T_i}\right) = 0.8$, because an ultimate oscillation will have a frequency of about $\frac{2\pi}{T_i} > \frac{2}{T_i}$.

Any variation of R_n with time will suggest that there might be some form of low frequency, non-stationary behaviour, which has a continually shifting dominant frequency. Any non-stationary behaviour at higher frequencies would not have any affect on R_n whereas, as can be seen from Fig. 3, this behaviour at lower frequencies will result in a significant variation in R_n . Such variation can be parameterised by estimating the moving standard deviation, σ_{R_n} , with time: a large σ_{R_n} will indicate low frequency, non-stationary behaviour.

2.2.2 PID

This sub-section examines the extension to PID control. Most of the results above also hold for PID control, but with the exception of R .

Extension 1: If a control loop is under PID control of form $K_P \left[1 + \frac{1}{T_i s} + T_d s \right]$ and the process has an oscillatory trend of frequency ω , then

$$R = \frac{Q_D}{\left(1 - \frac{2}{K_d}\right) + \frac{(T_i \omega)^2}{K_d^2} + \frac{1}{(T_i \omega)^2}} \quad (11)$$

where K_d is the ratio of the controller integral time constant T_i to the derivative time constant T_d , and Q_D is a constant as given by equation 44 in the Appendix A.3.

Proof: See Appendix A.3.

Remark 3: A normalised value of R can be then be defined as:

$$R_n = \frac{R}{Q_D} = \frac{1}{\left(1 - \frac{2}{K_d}\right) + \frac{(T_i \omega)^2}{K_d^2} + \frac{1}{(T_i \omega)^2}} \quad (12)$$

The frequency response of R_n for PID action is shown in Fig. 4(b), which differs from Fig. 4(a) in that there is no high frequency plateau. This derivation is for a particular form of PID controller. The frequency responses of R_n for other forms of PID controller should have a similar shape, but will have different equations for Q_D . Equation 43 in the Appendix A.3 gives a generic way to calculate the value of Q_D .

2.3 Estimation of η_y , η_u , R_n and standard deviation of R_n

A moving data window with N latest sampled data, $\{Y(i), i=1 \cdots N\}$ and $\{U(i), i=1 \cdots N\}$, is used to estimate both η_y and η_u after every sample interval. Let the estimation results be the time series $\hat{\eta}_y(k)$ and $\hat{\eta}_u(k)$, where k stands for the discrete time sequence.

Step 1: Normalisation.

The data is normalised as follows:

$$y(i) = \frac{Y(i) - \mu_Y}{\sigma_Y}, \quad (13)$$

$$u(i) = \frac{U(i) - \mu_U}{\sigma_U}, \quad i=1 \dots N \quad (14)$$

where μ_Y (or μ_U) and σ_Y (or σ_U) are the sample mean and sample standard deviation of Y (or U) over N samples. For computational simplicity, the controller set-point can be used instead of μ_Y when the set-point does not change, because there is little difference between them when a loop is well tuned.

Step 2: Extraction of deterministic trends $\hat{y}(i)$ and $\hat{u}(i)$ from $y(i)$ and $u(i)$.

Many de-noising techniques are available to estimate predictable components from a time series. A 30th order d -step ahead AR model with least squares estimation was adopted in the applications described here, with a data window of length $N=1000$. The choice of prediction horizon d , model order and data window length, was based on the recommendations of Thornhill *et. al.* [19], who also discuss how these parameters affect the models that are obtained. Note that e_y is the same as residual, r , in that paper.

Step 3: Estimation of $\hat{\eta}_y(k)$ and $\hat{\eta}_u(k)$

Estimates of the signal-to-noise ratio indices η_y and η_u can then be derived as:

$$\hat{\eta}_y(k) = \frac{\sigma_{\hat{y}}^2}{\sigma_{e_y}^2}, \quad (15)$$

$$\hat{\eta}_u(k) = \frac{\sigma_{\hat{u}}^2}{\sigma_{e_u}^2}, \quad (16)$$

where $\sigma_{\hat{y}}^2$ and $\sigma_{e_y}^2$ are the sample variances of \hat{y} and the residual noise, and $\sigma_{\hat{u}}^2$ and $\sigma_{e_u}^2$ are sample variances of \hat{u} and the residual noise.

An estimate of normalised R_n can then be obtained:

$$\text{for PI control: } \hat{R}_n(k) = \frac{\hat{\eta}_y(k)}{Q\hat{\eta}_u(k)} \quad (17)$$

for PID control:
$$\hat{R}_n(k) = \frac{\hat{\eta}_y(k)}{Q_D \hat{\eta}_u(k)} \quad (18)$$

The exponential weighted approach provides an estimate of the sample standard deviation of R_n , i.e., $\sigma_{\hat{R}_n}(k)$. The EWMA of R_n , i.e. $\mu_{\hat{R}_n}(k)$, is first calculated:

$$\mu_{\hat{R}_n}(k) = (1 - \lambda_1) \mu_{\hat{R}_n}(k-1) + \lambda_1 \hat{R}_n(k), \quad (19)$$

and then

$$\sigma_{\hat{R}_n}^2(k) = (1 - \lambda_2) \sigma_{\hat{R}_n}^2(k-1) + \lambda_2 (\hat{R}_n(k) - \mu_{\hat{R}_n}(k))^2, \quad (20)$$

where $\lambda_1 \in [0,1]$ and $\lambda_2 \in [0,1]$ are constant factors ($\lambda_1 = \lambda_2 = 0.1$ were used as defaults).

3. LOOP STATUS MONITORING

This Section describes how these statistics can be interpreted to derive the status of a loop. A quantitative interpretation is introduced, which will be used to compare the performances of different loops.

3.1 The Qualitative Loop Status Statistic

Table 1 gives test criteria for PI control. Note that the criteria involve the loop settling time, T_{st} , since this provides discrimination between short-term and longer term transients. A short-term transient is deemed to occur if the loop departs from either the steady state, or compensating condition, for a time duration that is less than the settling time i.e. $\eta_y > \gamma$ & $\eta_u > \gamma$ over time period $\Delta T : \Delta T < T_{st}$. If a value for T_{st} cannot be obtained, then obviously the relevant transients cannot be sub-categorised as short-term or long-term. The standard deviation of \hat{R}_n , i.e. $\sigma_{\hat{R}_n}$, is tested against a tolerance ξ . A suitable value for ξ is determined on the basis of simulations of various periodic and non-stationary trends. Fig. 5 shows the probability density functions (pdfs) of \hat{R}_n for these trends. The shape of pdf($\hat{R}_n(k)$) is very narrow ($\sigma_{\hat{R}_n} < 0.02$) for a cyclic type of trend, whilst the distribution is much wider ($\sigma_{\hat{R}_n} > 0.1$) for a long term non-stationary disturbance. A default value of 0.05 is found to distinguish between cyclic and non-stationary trends. As explained in Section 2.2, an *ultimate cyclic* trend occurs in a PI loop when $\hat{R}_n \approx 1$. Finally the category '*critical*' is created to accommodate those test outcomes, which cannot be attributed to the other statuses. These test outcomes can arise when the controller output proves to be ineffective, for instance because the loop is diverging and/or the controller output saturates. In all these cases $\eta_y \geq \gamma$. The value of ratio η_u will depend on the situation at that time. For instance if the controlled variable is diverging whilst the controller attempts to compensate, $\eta_u > \gamma$ but $\eta_y \geq 1.2\eta_u$, i.e., $R_n \geq 1.2$ (this tolerance is sufficiently large to accommodate uncertainty in the value of Q). Alternatively, the control output might be saturated, and hence constant, in which case η_u is contrived to be zero.

The table for PID control is almost identical with the exception that there is no test for Loop Status (6) and test $\hat{R}_n \geq 0.8$ is eliminated throughout because, referring to Fig. 4, there is no one-to-one mapping between R_n and the frequency of the trend. The Loop Status (6) situation is then recognised as Loop Status (5), which is renamed *long-term cyclic*.

A slightly different situation arises if the setpoint is changed whilst data is collected. As explained in Step 1 of Section 2.3, the controlled variable data set is normalised with respect to the setpoint unless the setpoint is changed, when mean value μ_y centred values are used instead. In this case, the effect of a setpoint change can be likened to that of an external disturbance, which would result in a particular Loop Status that reflects the trend style of the setpoint change, for example, a step change of the setpoint will lead to a *short-term transient* status. It would be difficult to distinguish between a frequently changing setpoint and a long term disturbance.

3.2 The Quantitative Loop Status Statistic

A quantitative version of Loop Status, the *LS* statistic, can be found by assigning the scores given in Table 2 to each category. The scores are merely chosen to produce a statistic that increases as the loop deteriorates. The quantitative version can then be filtered, for instance by applying a simple EWMA filter:

$$LL(k) = (1 - \lambda)LL(k-1) + \lambda LS(k), \quad (21)$$

where $LL(k)$ is the filtered statistic, $LS(k)$ is the score, λ is the filter factor, which is chosen so that it solely suppresses spikes, and k is the time sequence. The benefit of this formulation is that it converges to a value that represents the loop status.

A direct use of this quantitative statistic would be in the provision of an operator aid. A controller icon on a plant schematic could be made to change its colour or pattern in response to a change in loop status. Thus for instance the icon might be changed from green to yellow, when the controller has to compensate. Before a controller icon could be displayed properly, the smoothed/filtered *LL* statistic would have to be converted back to its qualitative description. The conversion rules are given in Table 3.

3.3 Examples of Loop Status Monitoring

Three examples are shown in Fig. 6. The upper panels show what happened when a slow sensor drift was introduced into a PID loop, the middle panels show the effect of a sticking control valve in a PI loop, and the lower panels show the effect of a load change in a PI loop. The plots show deviations in controlled variables from their setpoints, controller output deviations from their initial values and the filtered quantitative statistics LL , all versus time.

A simulated slow sensor drift: the LL value changes to status 1 to indicate that there is some kind of small disturbance or fault, which has been compensated for by the controller.

Control valve stick obtained from a real plant: the LL value is at 2.5 and $\sigma_{\hat{R}_n} = 0.016$, which indicates a low frequency, cyclic trend, which may be caused by a valve problem (sticking) or some external disturbance. Plant operators have confirmed that the valve was sticking.

A simulated change in load: the load change results in a step change in controller output causing the status to increase transiently. This indicates a short-term step-style load disturbance or, perhaps, a sudden change in sensor bias. In practice the controller parameters are no longer optimised.

Fig. 7 gives an example of the difference between the long-term sub-categories, non-stationary and cyclic disturbance. Focusing on the R_n plots, there is a more pronounced variation in the non-stationary case, which results in a larger $\sigma_{\hat{R}_n}$. A value of 0.0126 is obtained for the low frequency cyclical disturbance, whilst a value of 0.1158 is obtained when a long-term, non-stationary disturbance is introduced instead. The default ξ of 0.05 would discriminate between these two.

4. FAULT LOCALISATION IN A NUMBER OF INTERACTING LOOPS

4.1 OLPI Derivations & Fault localisation

The aim is to devise a loop index, the overall loop performance index or *OLPI*, whose value is largest for the problem loop. Thus an index is required that is suitably scaled (for inter loop comparison) and increases with loop degradation. Observe that the larger the value of *LL*, the worse the loop's performance, but this is too coarse a measure for purposes of comparison. On the other hand either η_y or η_u or both will be large for $LS > 0$, so one possibility might be

$$OLPI(k) = \max(\hat{\eta}_y(k), \hat{\eta}_u(k)) / \gamma \quad (22)$$

where γ is the same tolerance as before. Intuitively speaking, an indicative trend should be strongest in the loop from whence it came, so the signal-to-noise ratio indexes of the problem loop should be larger than those of affected loops. However it is important that the *OLPI* index should also take the extent of loop degradation into account, and therefore index *LL(k)* is incorporated into Equation 22 to give:

$$OLPI(k) = LL(k) \max(\hat{\eta}_y(k), \hat{\eta}_u(k)) / \gamma \quad (23)$$

This leads to the following localisation and partial diagnosis procedure:

- check *OLPI* values for each loop within a group of interacting controllers, locate the loop with the largest *OLPI* value, and mark this loop as most probably containing the root cause;
- check the Loop Status of this loop to narrow down the possible causes;
- perform further analyses, such as a non-linearity test [20] and others [15,19], to diagnosis the fault or disturbance in that loop.

4.2 Examples of OLPI-Based Fault Localisation

Example 1: Real data from an Eastman Chemical plant

Fig. 8 shows steady state data obtained for an Eastman Chemical plant that contains 9 master loops [20]. The Overall Loop Performance Index (*OLPI*) was evaluated for each loop by applying a moving data window of 1000 sample points at each time step assuming that all the setpoints remained unchanged during this period. All default thresholds were used. The *OLPI* values were then averaged over the two days. The results (Table 4) suggest that loop LC2 (tag=22) contains the root-cause, because of its

largest *OLPI* value. Its loop status then indicates that it is operating in a long-term, cyclic mode. A non-linearity test, applied to this data, has shown that this loop is operating badly, and this has also been confirmed by plant engineers [20].

Example 2: Case studies from Tennessee Eastman (TE) process benchmark

The benchmark was combined with the plant-wide decentralised control scheme developed by McAvoy *et. al.* [21], to which the reader can refer for typical plots. To save space, only the plots pertaining to the most important loops are given here. The process plant contains multiple control loops as outlined in Table 5: many of these are standard cascade control systems. Only the master loops are analysed, because a failure in an inner loop is likely to be seen as a degradation in the performance of the outer loop and the setpoints on their inner loops change too frequently. The setpoints of the master loops remained unchanged. Table 5 also contains values for the threshold parameters γ , two of which were customised by analysing normal steady state data.

Two case studies are presented: the first demonstrates that the *OLPI* should be a maximum for a loop in which a composition has a step change; the second looks at the effect of a gradual deterioration in the process. In both cases it was difficult to decide whether loops 3 and 9 were badly tuned or were continually subjected to excessive noisy disturbances.

Fig. 9: step change disturbance to Loop 5 (B composition). Loop 5 regulates the composition of B by diverting a proportion of the recycle stream out of the process via a purge valve. The perturbed recycle stream then affects the composition into the reactor and this disturbs Loop 1 (a cascade system that regulates the A/C component into the reactor), and the reactor itself. Thus reactor pressure (Loop 3), condenser cooling water temperature (Loop 4) and eventually the product flow rate (Loop 8) are affected. It is easily seen that Loop 5 (Purge composition B—purge flow rate cascade control) has a significantly larger *OLPI* value than others. The LL trend of loop 5 confirms that the behaviour is associated with a short-term transient disturbance. After a period of time, Loops 1, 4, 5 all return to good conditions ($LL=0$), whilst Loops 3 and 8 continue to compensate for the change in operating point.

Fig. 10: slow drift in reaction kinetics. The reactor is central in the T-E process plant. Its performance affects the rest of the plant and the plant recycle contributes to the generation of complicated transients. Although Loop 3 (the reactor pressure cascade control loop) is central to its performance, other loops are coupled tightly. Initially the change in reactor kinetics causes a small, slow drift in the OLPI's of at least 5 loops. However the plant then experiences larger, short-term transients and during this period the OPLI of Loop 3 rises, significantly, to about a factor of 10 greater than any other loop. Application of the localisation procedure thus suggests that Loop 3 contains the root cause, because the OLPI is considerably greater than others.

5. CONCLUSIONS

Loop status has been described as a qualitative measure of current loop performance. The basis for defining seven qualitative descriptions has been given and quantitative versions of loop status statistics tests have been described that enable a trend to be categorised. Loop Status monitoring and fault localisation based on a comparison of overall loop performance indices has proved to be successful and robust by both simulation and real industrial data analysis. This comparison can point out the problem loop, and Loop Status information helps to narrow down the possible root cause. Although the approach has been developed for PI/PID controllers, it should be equally applicable for those controllers that have a frequency dependent statistic R , which is the ratio of the signal to noise ratios of the controlled variable and controller output. Thus the approach is not suitable for a P control loop, because R will always be unity. Equally it is also not feasible for open-loop control, because R is intended to reflect the relationship between the control error and the controller output, and this does not exist for an open-loop control system.

This approach is successful when analysing a plant with interacting loops that is perturbed by ONE dominant deterministic disturbance at a time. It is unlikely to work for multiple faults that result in multiple dominant disturbances. An extension is under research to accommodate the multiple-faults situation. Finally the examples have focused on data obtained from plants in steady state; although more difficult to interpret, an explanation has been given of how data involving setpoint changes could be analysed.

Acknowledgements

Chunming Xia & John Howell gratefully acknowledge Nina Thornhill from UCL, and John Cox & Michael Paulonis from *Eastman Chemical Company, Kingsport, TN, USA*, for the supply of real plant data. Mrs Thornhill also edited the final draft for us.

References

1. Harris, T.J., C.T. Seppala and L.D. Desborough, A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *Journal of Process Control* 9 (1999), pp. 1-17.
2. Huang, B. and S.L. Shah, *Performance assessment of control loops*. Springer, 1999.

3. Harris, T.J. and C.T. Seppala, Recent Developments In Controller Performance Monitoring and Assessment Techniques, Chemical Process Control VI, Tucson, AZ, (Jan. 2001).
4. Janusz, M.E. and V. Venkatasubramanian, Automatic generation of qualitative descriptions of process trends for fault detection and diagnosis, Engineering Applications of Artificial Intelligence, 4 (1991), pp.329-339.
5. Bakshi, B.R. and G. Stephanopoulos, Representation of Process trends. III. Multiscale extraction of trends from process data, Computers & Chemical Engineering, 18 (1994a), pp.267-302
6. Bakshi, B.R. and G. Stephanopoulos, Representation of Process trends. IV. Induction of real-time patterns from operating data for diagnosis and supervisory control, Computers & Chemical Engineering, 18 (1994b), pp.303-332
7. Wong, J.C., K.A. McDonald and A. Palazoglu, Classification of abnormal plant operation using multiple process trends, Journal Process Control, 11 (2001), pp. 409-418.
8. Rengaswamy, R., T. Hägglund and V. Venkatasubramanian, A qualitative shape analysis formalism for monitoring control loop performance, Engineering Applications of Artificial Intelligence, 14 (2001) pp.23-33.
9. Xia, C. and J. Howell, Controller output based, single number statistics for control loop status monitoring. 4TH IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries, Korea, June, (2001a).
10. Xia, C. and J. Howell, Loop status statistics. 4TH IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries, Korea, June, (2001b).
11. T. Hägglund, A control-loop performance monitor, Control Eng. Practice 3 (1995), pp. 1543-1551.
12. T. Hägglund, Automatic detection of sluggish control loops. Control Eng. Practice 7 (12) (1999), pp. 1505-1511.
13. Miao, T., D. Seborg, Automatic detection of excessively oscillatory feedback control loops, *IEEE Int. Conf. On Control Applications*, Hawaii, USA., 1999.
14. Horch, A simple method for the detection of stiction in control valves, Control Eng. Practice, 7 (1999), pp. 1221-1231.
15. Forsman, K., A. Stattin, A new criterion for detecting oscillations in control loops, *European Control Conference*, Karlsruhe, Germany, CP8-3, 1999.

16. Thornhill, N. F., Shah, S. L., Huang, B., & Vishnubhotla, A. 2002, Spectral principal component analysis of dynamic process data, *Control Engineering Practice*, vol. 10, no. 8, pp. 833-846.
17. Thornhill, N. F., Huang, B., & Zhang, H. 2003, Detection of multiple oscillations in control loops, *Journal of Process Control*, vol. 13, no. 1, pp. 91-100.
18. D.J. Kozub, Controller performance monitoring and diagnosis, Proceedings of Fifth International Conference on Chemical Process Control, CACHE Corp, 1997, pp. 83-96.
19. Thornhill, N.F., M. Oettinger and P. Fedenczuk, Refinery-wide control loop performance assessment. *Journal of Process Control* 9 (1999), pp. 109-124.
20. Thornhill, N.F., C. Xia, J. Howell, J. Cox, M. Paulonis, Analysis of plant-wide disturbances through data-driven techniques and process understanding, 15TH IFAC World Congress, Barcelona, 2002, Session 38.
21. McAvoy, T.J., N. Ye, G. Chen, An improved base control for the Tennessee Eastman problem, Proceedings of American Control Conference, WA9-10:15, 1995, pp.240-244.

Appendices

A.1 Proof Of Theorem 1

Assume that a closed loop has a sine wave oscillatory trend with frequency ω , then:

$$\hat{y}(t) = A \sin(\omega t), \quad (24)$$

With reference to Section 2.1, assume that the power spectrum, $G_{e_y}(\omega)$, of e_y is uniform over the bandwidth, that is,

$$G_{e_y}(\omega) = \begin{cases} G & B_0 \leq \omega \leq B \\ 0 & \begin{cases} \omega > B \\ \omega < B_0 \end{cases} \end{cases} \quad (25)$$

where *Nyquist* angular frequency $B = \pi/T_s$ (T_s is sampling time interval) and B_0 is low-band limit. Define K_B : $B_0 = B/K_B$.

Variance of e_y : $\sigma_{e_y}^2$

$$\sigma_{e_y}^2 = R_{e_y}(0) = \int_{B_0}^B G_{e_y}(\omega) d\omega = G(B - B_0) \quad (26)$$

Variance of e_u : $\sigma_{e_u}^2$

According to the input/output power spectrum relationship, the auto-spectrum of e_u can be given by:

$$G_{e_u}(\omega) = |H(j\omega)|^2 G_{e_y}(\omega) = K_p^2 \left(1 + \frac{1}{T_i^2 \omega^2} \right) G_{e_y}(\omega) \quad (27)$$

where K_p and T_i are PI controller parameters, and the variance of e_u is:

$$\sigma_{e_u}^2 = \int_{B_0}^B G_{e_u}(\omega) d\omega = G(B - B_0) K_p^2 Q \quad (28)$$

where

$$Q = 1 + \frac{K_B}{\pi^2} \left[\frac{T_s}{T_i} \right]^2 \quad (29)$$

Variances of \hat{y} and \hat{u} : $\sigma_{\hat{y}}^2, \sigma_{\hat{u}}^2$

The s -domain expression of $\hat{y}(t)$ is:

$$\hat{Y}(s) = \frac{A\omega}{s^2 + \omega^2} \quad (30)$$

$$\text{and } \hat{U}(s) = H(s)\hat{Y}(s) = K_p \left[1 + \frac{1}{T_i s} \right] \left(\frac{A\omega}{s^2 + \omega^2} \right) \quad (31)$$

The controller output can be obtained by Laplace inverse transformation, as

$$K_p \left[A \sin \omega t + \frac{A}{T_i \omega} (1 - \cos \omega t) \right] \quad (32)$$

Since the data analysis is based on deviation variables, form $\hat{u}(t)$ by removing the constant term $\frac{K_p A}{T_i \omega}$ to give

$$\hat{u}(t) = K_p A \sqrt{M} \sin(\omega t + \beta) \quad (33)$$

where $M = 1 + \frac{1}{(T_i \omega)^2}$, and $\beta = -\tan^{-1}\left(\frac{1}{T_i \omega}\right)$,

The variance of deviations $\hat{u}(t)$ and $\hat{y}(t)$ can be given by:

$$\sigma_{\hat{u}}^2 = \frac{1}{2} K_p^2 A^2 M \quad (34)$$

$$\sigma_{\hat{y}}^2 = \frac{1}{2} A^2 \quad (35)$$

Signal-to-noise ratio indexes: η_y and η_u

Substituting the above equations for $\sigma_{\hat{y}}^2, \sigma_{e_y}^2, \sigma_{\hat{u}}^2$ and $\sigma_{e_u}^2$ into Equations 4,5 (Definition 1) yields Equations 7 & 8.

Ratio of indexes: R

Substituting the above equations for η_y and η_u into Equation 6 yields Equation 9.

A.2 General deterministic oscillatory trends (PI)

A more general expression for the deterministic oscillatory trend $\hat{y}(t)$ is a Fourier series representation of sine waves with n components,

$$\hat{y}(t) = \left(\sum_{k=1}^n A_k \sin(k\omega t) \right), \quad (36)$$

where component $A_1 \sin(\omega t)$ is the dominant fundamental component, i.e., ω is the fundamental frequency and $A_1 \gg A_k, k = 2 \dots n$.

Applying the Principle of Linear Superposition, the controller output deviation can be derived as follows:

$$\hat{u}(t) = K_p \sum_{k=1}^n A_k \sqrt{M_k} \sin(k\omega t + \beta_k) \quad (37)$$

where $M_k = 1 + \frac{1}{(T_i k \omega)^2}$ and $\beta_k = -\tan^{-1}(\frac{1}{T_i k \omega})$

The variance of $\hat{u}(t)$ is:

$$\sigma_{\hat{u}}^2 = E(\hat{u}^2) = K_p^2 E \left[\left(\sum_{k=1}^n A_k \sqrt{M_k} \sin(k\omega t + \beta_k) \right)^2 \right] = K_p^2 \left(\frac{1}{2} \sum_{k=1}^n A_k^2 M_k \right) \quad (38)$$

Note that the mathematical expectations of the terms involving the cross-product of sinusoids are zero because of orthogonality.

The variance of $\hat{y}(t)$ is:

$$\sigma_{\hat{y}}^2 = \frac{1}{2} \sum_{k=1}^n A_k^2 \quad (39)$$

The ratio R can be given by

$$R = \frac{\eta_y}{\eta_u} = \frac{\frac{1}{2} \sum_{k=1}^n A_k^2}{K_p^2 \left(\frac{1}{2} \sum_{k=1}^n A_k^2 M_k \right)} = \frac{Q \sum_{k=1}^n A_k^2}{\sum_{k=1}^n A_k^2 M_k} = \frac{Q}{G(B-B_0) K_p^2 Q} \quad (40)$$

If $\omega \geq \frac{2}{T_i}$, then $M_k \leq 1 + \frac{1}{4k^2}$, and

$$R_n \geq \frac{\sum_{k=1}^n A_k^2}{\sum_{k=1}^n A_k^2 \left(1 + \frac{1}{4k^2} \right)} = \frac{1}{1 + \frac{\sum_{k=1}^n (A_k^2 \frac{1}{4k^2})}{\sum_{k=1}^n A_k^2}} \cong \frac{1}{1 + \frac{\frac{1}{4} A_1^2}{A_1^2}} = 0.8 \quad (41)$$

which is the same as for the single sine wave.

A.3 Oscillatory trend with frequency ω (PID)

For a generalised PID controller the ratios R and R_n can be represented as:

$$R = \frac{Q_D}{M_D} ; R_n = \frac{R}{Q_D} = \frac{1}{M_D} \quad (42)$$

where Q_D is a constant and M_D is a frequency dependent function. If the controller has a transfer function $H(s)$, the analysis of Appendix A.1 can be repeated to produce the following expression for Q_D :

$$Q_D = \frac{1}{K_p^2} \frac{\int_{B_0}^B |H(j\omega)|^2 d\omega}{B - B_0} \quad (43)$$

where K_p is the controller gain, $B = \frac{\pi}{T_s}$, and B_0 can be chosen as the notch frequency

$\frac{1}{\sqrt{T_i T_d}}$. Thus for a PID controller with transfer function

$$K_p \left(1 + \frac{1}{T_i s} + T_d s \right)$$

where T_d is the derivative time constant,

$$Q_D = \left(1 - \frac{2}{K_d} \right) + \frac{\pi^2}{3K_d^2} \left(\frac{T_i}{T_s} \right)^2 + \frac{K_B}{\pi^2} \left(\frac{T_s}{T_i} \right)^2 \quad (44)$$

$$M_D = \left(1 - \frac{2}{K_d} \right) + \frac{(T_i \omega)^2}{K_d^2} + \frac{1}{(T_i \omega)^2} \quad (45)$$

and K_d is the ratio of T_i/T_d .

When $K_d \rightarrow \infty$, i.e., there is no derivative action, R will equal Q/M , i.e. the result will be the same as for PI control. For a normal value of K_d , the $R_n - \omega$ relationship will differ from that of PI control (Fig. 4). It can be seen that the curve is not a monotonically increasing function, so the value of R_n cannot be used to determine whether the trend is of high frequency.

Criteria						Loop Status
$\hat{\eta}_y \geq \gamma$	$\hat{\eta}_u \geq \gamma$	$\hat{R}_n \leq 1.2$	$\hat{R}_n \geq 0.8$	$\sigma_{\hat{R}_n} > \xi$	$\Delta T \geq T_{ST}$	
No	No	-	-	-	-	(1) Steady
No	Yes	-	-	-	-	(2) Compensated
Yes	Yes	Yes	-	-	No	(3) Short-term Decaying Transient
Yes	Yes	Yes	No	Yes	Yes	(4) Long-term Non-stationary
Yes	Yes	Yes	No	No	Yes	(5) Long Term Low frequency cyclic
Yes	Yes	Yes	Yes	-	-	(6) Ultimate cyclic
Yes	Yes	No	-	-	-	(7) Critical
Yes	No	-	-	-	-	(7) Critical

Table 1 Loop Status Criteria For PI Control

1.1.1.1.1.1 Loop Status	LS
Steady	0
Compensated	1
Short term transient	2
Long term transient	2.5
<ul style="list-style-type: none"> • Cyclic ($\sigma_{\hat{R}_n} \leq \xi_3$) • Non-stationary ($\sigma_{\hat{R}_n} > \xi_3$) 	
Ultimately cyclic	3
Critical	4

Table 2 Quantitative values of Loop Status

Status	Range
Compensated	$0.5 < LL < 1.5$
Short term transient	$1.5 \leq LL < 2.25$
Long term transient	$2.25 \leq LL \leq 2.75$ (PI)
• Cyclic ($\sigma_{\hat{R}_n} \leq \xi_3$)	$2.25 \leq LL \leq 3.5$ (PID)
• Non-stationary ($\sigma_{\hat{R}_n} > \xi_3$)	
Ultimately cyclic	$2.75 < LL \leq 3.5$ (only for PI)
Critical	> 3.5

Table 3 Rules for converting the filtered statistic LL into a qualitative version

Tag	Loop	Mean(<i>OLPI</i>)	Loop Status
1	PC1	11.6	Compensated
5	LC1	14.3	Long-term transients (cyclic)
10	FC4	48.2	Compensated
13	TC1	38.4	Long-term transients (ultimate cyclic)
16	PC2	< 5	Ultimate cyclic status detected, temporarily, early on.
17	LC3	< 5	Compensated
22	LC2	132	Long-term transients (cyclic)
25	TC2	13	Long-term transients (cyclic)
30	FC7	7	Compensated

Table 4 Loop Status & *OLPI* for real data from an Eastman Chemical Plant

Loop	Node	γ
1	A/C composition — A flow rate cascade control	Default
2	G/H composition — reactor level — D,E flow rate control	Default
3	Reactor pressure — temperature cascade control	7*
4	Condenser cooling control	Default
5	Purge composition B — purge flow rate cascade control	Default
6	Separator level — underflow rate cascade control	Default
7	Stripper level — product flow rate cascade control	Default
8	Product flow rate — C flow rate cascade control	Default
9	Product composition E — stripper temperature cascade control	1.4 *

- Loops 3 and 9 have somewhat non-steady initial behaviours, so customised γ values are used.

Table 5 The nine master loops of TE process benchmark

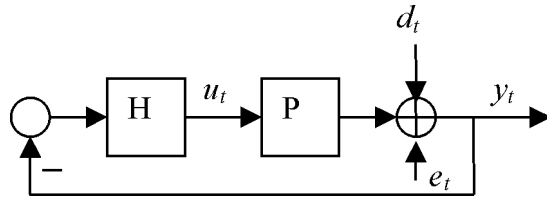


Fig. 1 Loop with deterministic disturbance & white noise excitation

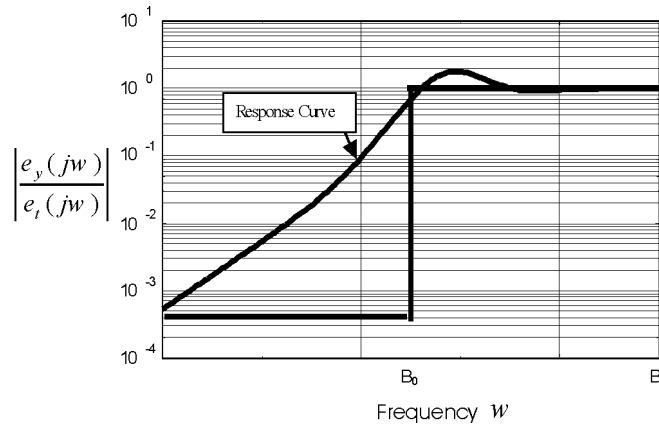


Fig. 2 Filtered White Noise Frequency Response For A Typical PI Control Loop

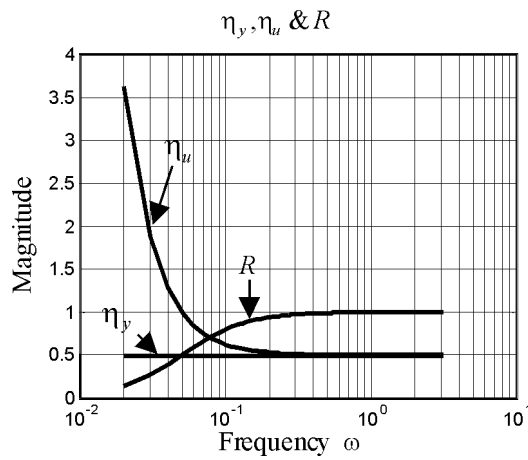


Fig. 3 η_u, η_y & R relationships when $\eta_y=0.5, Q=1$ & $T_i=20$

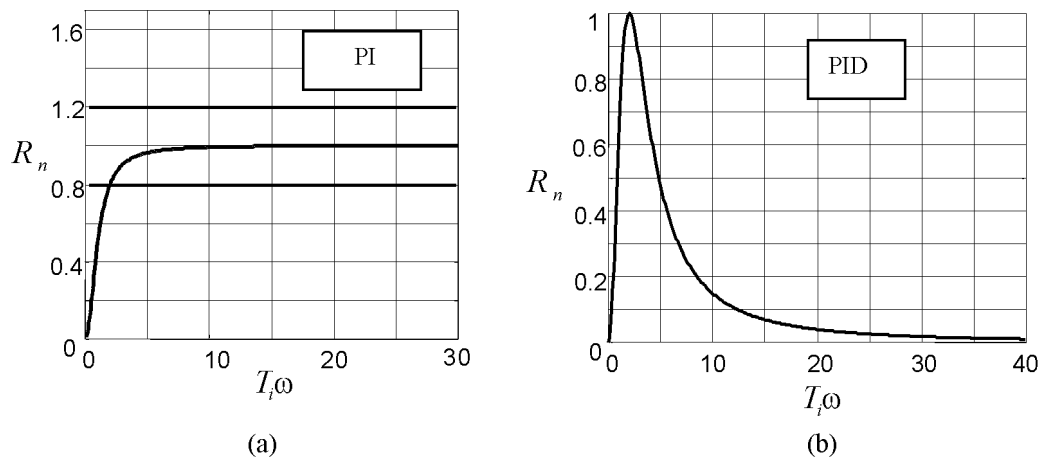


Fig. 4 Variation of R_n with $T_i\omega$ for PI & PID controllers

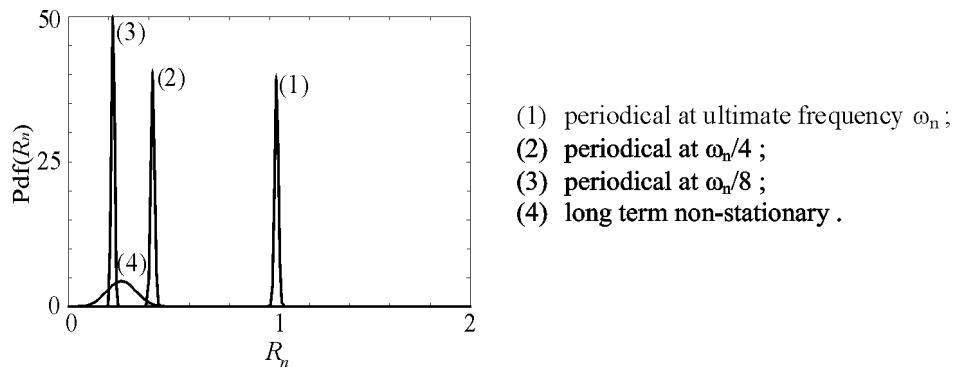


Fig. 5 PDFs of R_n for 4 trends obtained via simulation

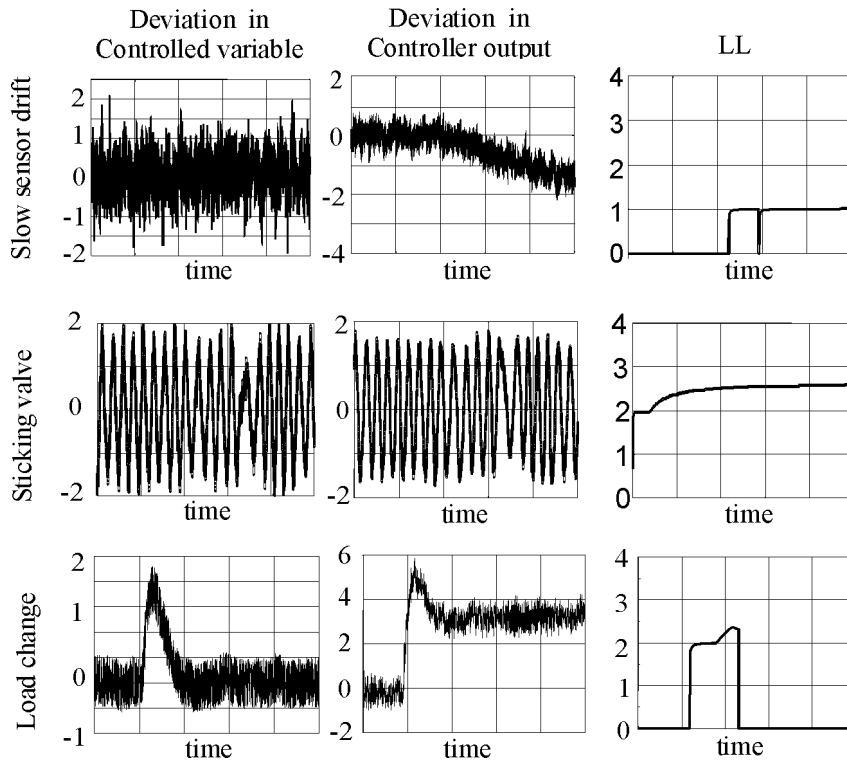


Fig. 6 Examples of filtered LL Statistics

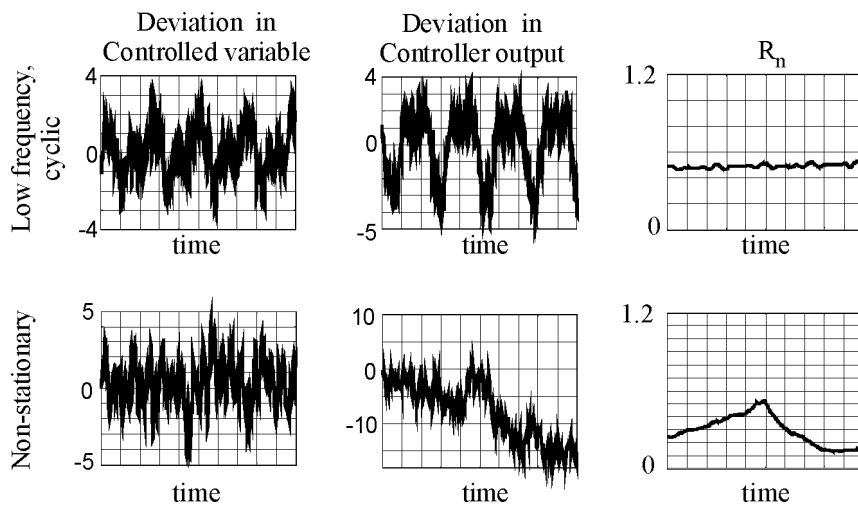


Fig. 7 Comparison between a low- f oscillation and a non-stationary disturbance

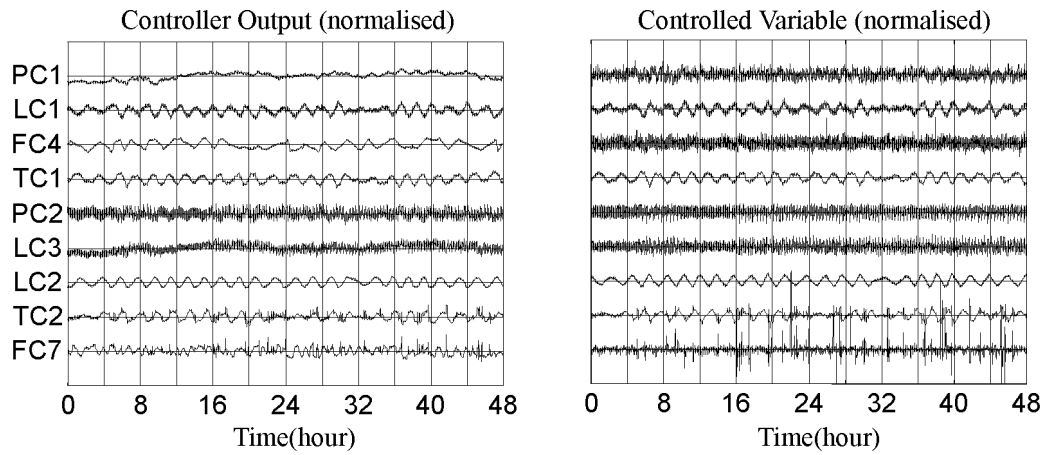


Fig. 8 Real plant data from an Eastman Chemical Plant

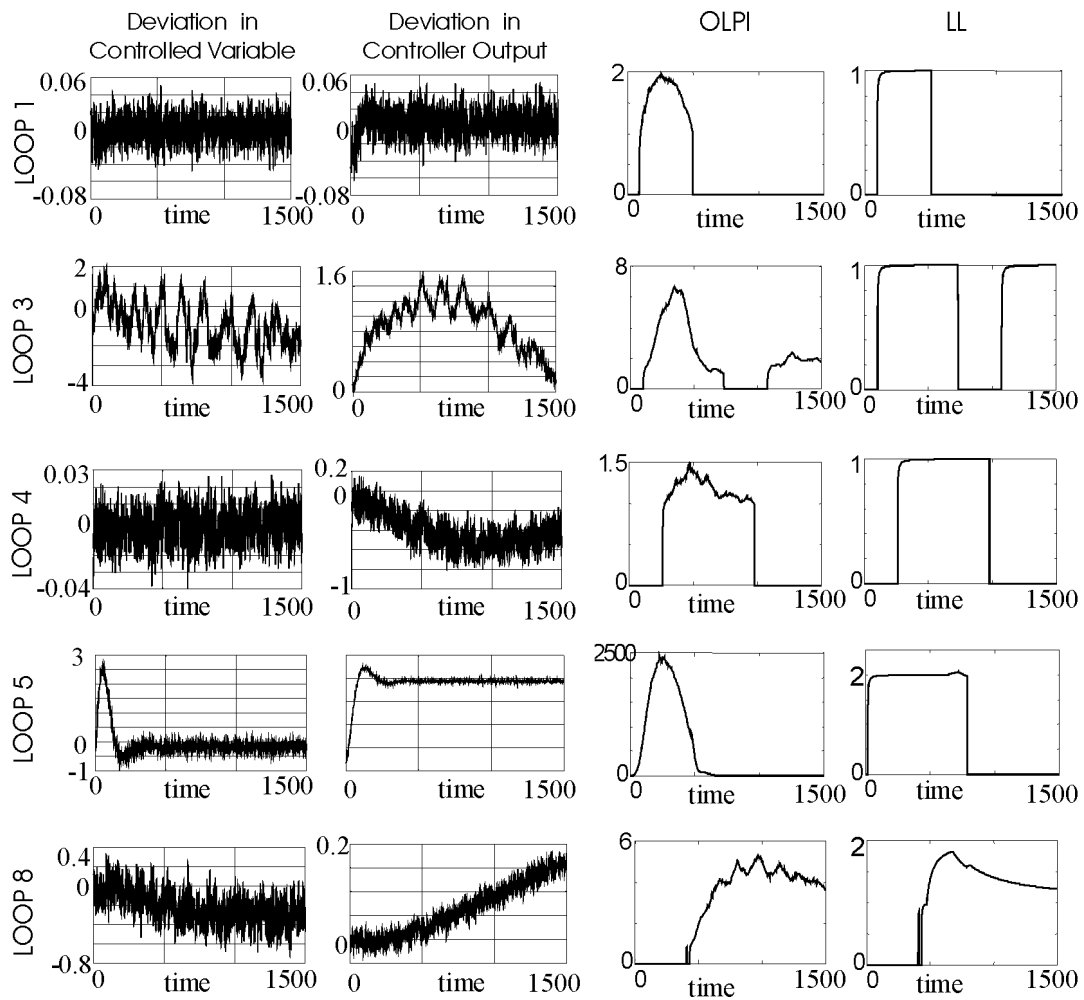


Fig. 9 Step change disturbance to B composition: Data, OLPI & LL

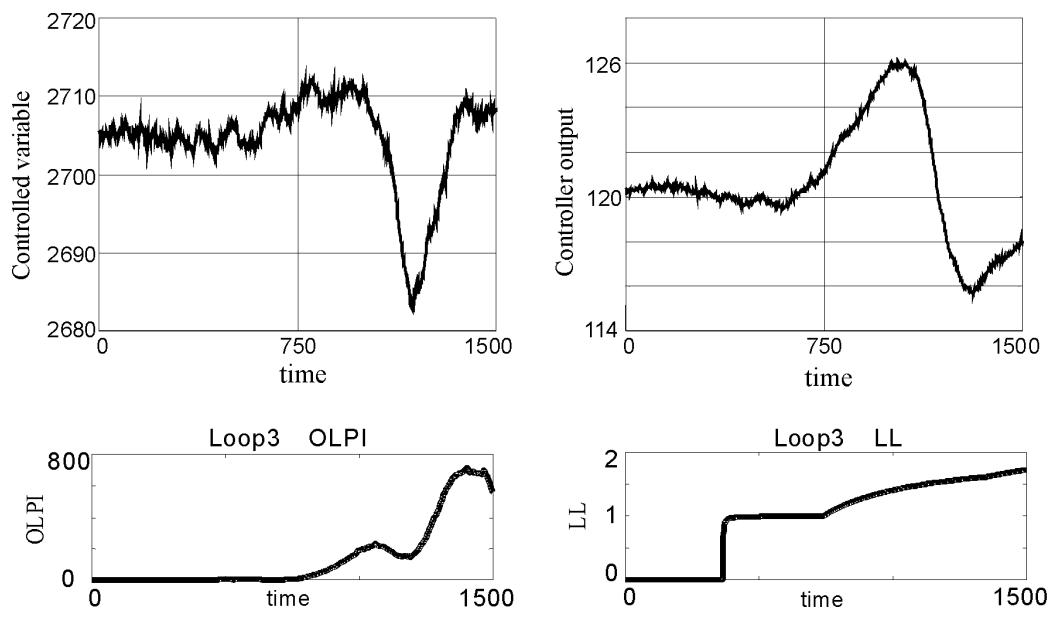


Fig. 10 Reaction kinetic slow drift: Loop 3 data, OLPI & LL