

CAUSE AND EFFECT ANALYSIS OF A CHEMICAL PROCESS

ANALYSIS OF A PLANT-WIDE DISTURBANCE

Margret Bauer and Nina F. Thornhill, University College London, Torrington Place, London WC1E 7JE
 Adrian Meaburn, BP, Saltend, Hedon, Yorks, HU12 8DS

Abstract

In continuous chemical processes, disturbances in the process conditions can propagate widely and cause secondary upsets in remote locations. The aim of this paper is to apply some recent data-driven methods for detection and diagnosis of process disturbances using historical process data that have been proving successful in a range of applications. An industrial case study is presented in which a plant-wide control system disturbance caused by the presence of a recycle was successfully located and then verified by further plant testing.

1. INTRODUCTION

A plant-wide oscillation in a chemical process often has an impact on product quality and running costs and there is a motivation for automated diagnosis of the source of such a disturbance. The situation is made worse when the oscillation propagates throughout a dynamic system such as a chemical plant.

This article applies established methods of analyzing data from routine operation to detect the extent and severity of a plant-wide oscillation. It also presents a recently developed method called *transfer entropy* to identify the direction of propagation of the disturbance. Transfer entropy is sensitive to directionality even in the absence of an observable time delay.

Root causes of plant-wide oscillations due to sticking valves have been studied extensively and solutions now exist for that case [1-5], however other causes remain to be addressed including oscillations emerging in processes with recycle, also known as the snowball effect. The transfer entropy method looks promising for diagnosis of this type of problem.

The methods are presented through a worked example from BP Chemicals Hull in a process with two recycles. The data were 1 min samples comprising 6 automatic control loops. Further trials were conducted which were guided by the findings of the study. The methods successfully pointed out areas for investigation but site expertise is still required to get to the bottom of them. There is need for industrial tools that will help company expertise focus on problem areas early and to combine data-driven analysis with their knowledge and understanding of the process.

2. CASE STUDY

The process schematic for the case study is shown in Figure 1. The tag descriptions are given in Table 1.

3. VISUAL INSPECTION

Normalised time trends: Normalised time trends for the set points (*sp*), controlled variable (*pv*), controller error (*sp-pv*) and controller output (*op*) are plotted in Figure 2. Normalisation means that the mean values of the time trends were removed and their standard deviations were scaled to unity before plotting. There were some changes in set point during the episode shown particularly in LC3. The analysis of controller error is beneficial in the presence of set point changes, for

instance the high frequency deviations look small in the time trend for Tag 3 in the *pv* plot (Figure 2) because Tag 3 (LIC50064) also varied as it was tracking a slowly varying set point. The plot of the controller error for Tag 3 shows the oscillatory deviations more prominently.

Tag	name	description
1	LC1	buffer tank level
2	LC2	flash drum level
3	LC3	reflux drum level
4	LC4	reactor level
5	PC1	reactor pressure
6	TC1	reactor temperature
7	LC1.op	buffer tank level controller output
8	LC2.op	flash drum level controller output
9	LC3.op	reflux drum level controller output
10	LC4.op	reactor level controller output
11	PC1.op	reactor pressure controller output
12	TC1.op	reactor temperature controller output

Table 1. Tag descriptions

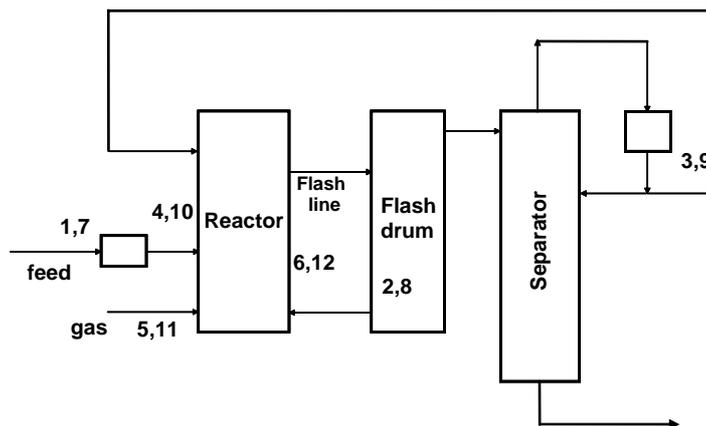


Figure 1. Process schematic

Disturbances: The controller outputs for LC1 and LC4 (Tags 7 and 10) show long term coordinated deviations which have the appearance of long excursions. Other tags show distinct oscillations. The principal aim of the analysis of the plant data is to understand the origin of the oscillation and to suggest further tests to verify the hypothesis.

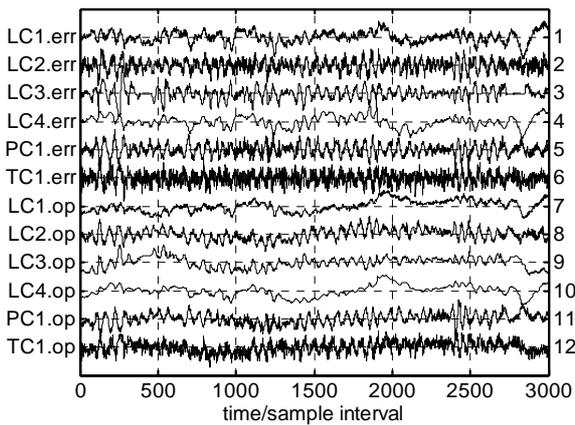
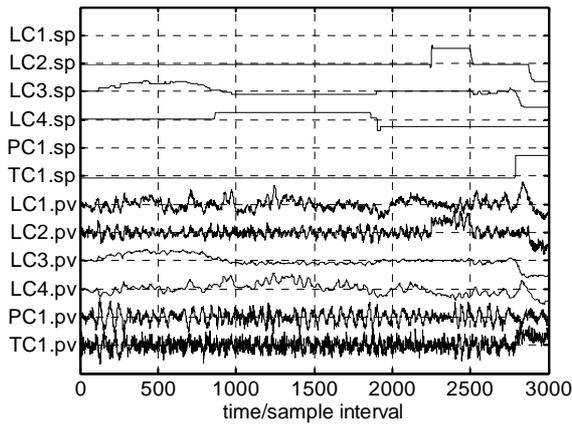


Figure 2. Upper panel: Set points (*sp*) and process measurements (*pv*). Lower panel: Controller errors (*err*) and controller outputs (*op*)

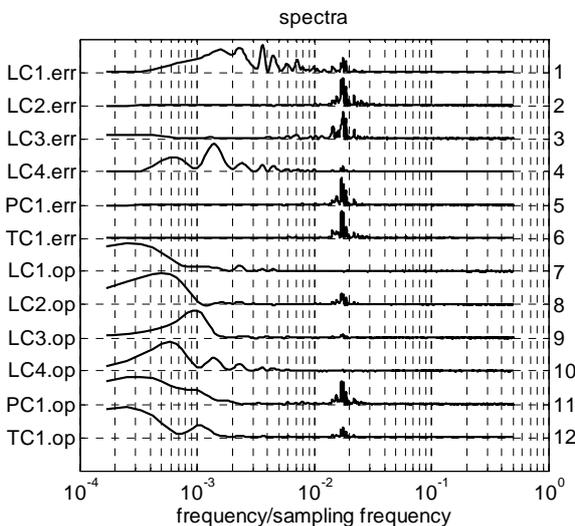


Figure 3: Spectra of controller errors and outputs

4. SPECTRAL AND OSCILLATION ANALYSIS

4.1. Spectral analysis

Spectral analysis: A plant-wide spectral analysis can detect measurements having similar spectral features. The underlying notion is that measurements whose spectra are similar are subject to the same disturbance. Finding clusters of measurements with similar disturbances is the first step in a plant-wide diagnosis.

Power spectra are invariant to phase of a signal which means they are insensitive to the time delays between one tag and another.

Figure 3 shows the power spectra. The axes for the spectra and time trends in Figure 2 in are reciprocal. If there is a spectral peak at, say, 0.018 on the frequency axis then the time trend has an oscillation with a period of $1/0.018 = 56$ samples per cycle. Examples can be seen in Tags 2 and 3.

The spectra highlight dynamic features that cannot be seen easily by inspection of the time trends. Some tags (e.g. LC2, LC3, PC1 and TC1) have a persistent oscillation which can be seen as peaks at particular frequencies in the spectra. Other have a spectral features at low frequency. Tags showing low frequency effects are principally the controller outputs, Tags 7 to 12. It means that the controller outputs are drifting with long term deviations from the mean value. This is to be expected for a controller with integrating action because slowly varying process disturbances should appear as offsets in the controller output.

Spectral classification tree: Figure 4 is a spectral classification tree that groups together tags with similar spectra and hence similar dynamic features. In the tree, each whole spectrum is represented as a spot on the horizontal axis. Spectra form a cluster if they are connected to each other by short vertical lines and are well separated from all other spectra. There are two clusters in the data. Tags 1 and 4 have similar spectra, as do 2,3, 5 and 6. If the vertical lines between clusters are long then it means the spectra are very different in the two clusters.

Figure 3 can explain the classifications. Tags 2, 3, 5, and 6 all have a peak at about 0.018 on the frequency axis corresponding to the 56 minute oscillation. Tags 1 and 4 have a range of spectral features at lower frequency and the tree shows that these spectra are significantly different from the spectra of other measurements. These are the buffer tank level and the reactor level. The group of controller outputs Tags 7, 8, 10, 11 and 12 all have very low frequency features. They are more similar to each other than to any other tags but do not form a tight cluster.

Tag 9, the controller output of the reflux drum level, is shown to be greatly different from all other tags because it is joined into the classification tree by a long branch. No other controller error or controller output shares its dynamic behaviour. The reason is that the set point for Tag 3, the reflux drum level deviated up then down between sample 1 and about sample 1000 and the controller output responded to the set point trajectory in order to keep the level at its set point. The LC3.pv tag shows the same behaviour because the reflux drum level is under tight control, however it does not show in the spectral tree because the tree was constructed from the spectra of the controller errors and output, not from the *pv*'s.

The locations of the main disturbances are shown on the process schematic in Figure 5. The 56 minute oscillation is a plant-wide disturbance because it appears in all units, while the disturbance in Tags 1 and 4 with a broad range of spectral features at lower frequency is localized to the level controls of the reactor and feed buffer tank.

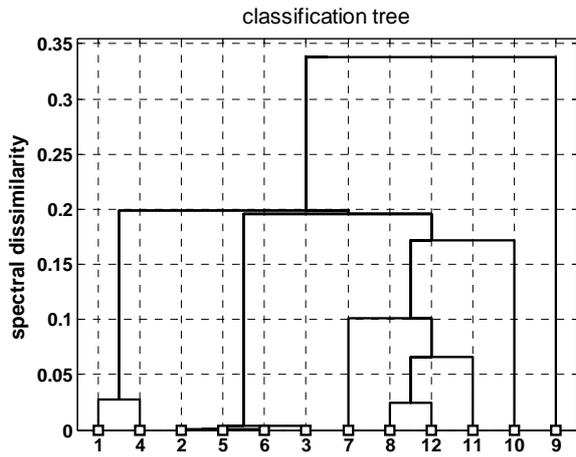


Figure 4. Spectral classification tree for controller errors (Tags 1 to 6) and controller outputs (Tags 7 to 12).

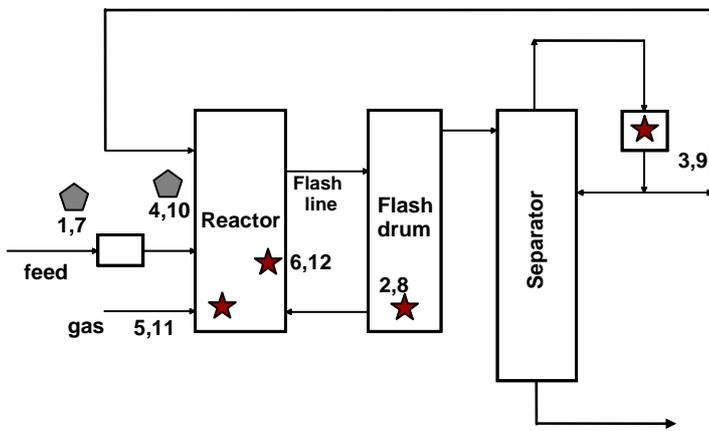


Figure 5. Locations of the main disturbances. ★ The plant-wide 56 minute oscillation, ⬠ Localized disturbance with a broad band of spectral features.

4.2. Oscillation analysis

Oscillation analysis: The purpose of oscillation analysis [6, 7] is to characterize the oscillations present in the data set. It determines the intervals between zero crossings of the autocovariance function and the regularity of the zero crossings. A benefit of the use of the autocovariance function to detect zero crossings rather than the time trend is that it is much less noisy while retaining the same period of oscillation. The oscillation index is high if the zero crossings are regularly spaced. An oscillation index below 1.00 means there is no regular oscillation.

Figure 6 shows oscillation analysis with the oscillation indexes (*O.I.*) presented on the right hand side. It confirms the oscillations present in Tags 2, 3, 5 and 6. Their period as determined from the intervals between the zero crossings was 56 samples per cycle.

The ACF is determined from the inverse Discrete Fourier Transform (DFT) of the two-sided power spectrum. An advantage of the method is that data may be filtered before analysis. For instance, high frequency noise may be removed by excluding high frequencies while slowly varying disturbances can be removed by excluding low frequencies before the inverse DFT step. Figure 7 show the filtered time trends after removal of

low frequency cycles longer than 100 samples. Table 2 is the oscillation analysis of the filtered trends. Filtering shows that Tags 8, 9, 11 and 12 (the controller outputs for Tags 2, 3, 5 and 6) are also participating in the oscillation and that the oscillation is also present in tags 1 and 4 (reactor and feed tank levels), although at a lower level of signal power.

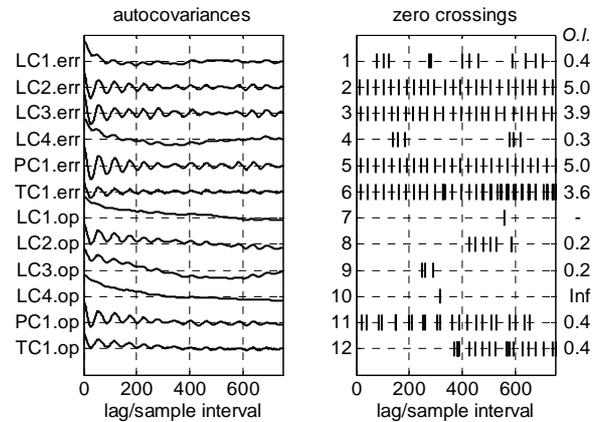


Figure 6: Oscillation analysis, no filtering.

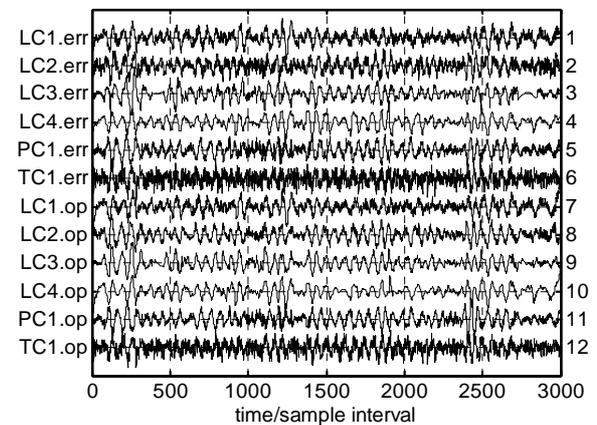


Figure 7: Time trends filtered using [2 100] filter to remove low frequency deviations.

Name	Tag No.	Period	OI	power%
1	LC1	55.3	3.1	34
2	LC2	56.7	5.0	99
3	LC3	56.2	6.4	89
4	LC4	56.0	2.3	24
5	PC1	56.2	4.8	96
6	TC1	56.2	4.6	97
7	LC1.op	55.3	2.1	12
8	LC2.op	56.4	5.1	57
9	LC3.op	55.6	4.7	28
10	LC4.op	56.0	1.5	9
11	PC1.op	56.0	4.3	68
12	TC1.op	56.2	5.8	67

Table 2. Oscillation analysis using [2 100] filter.

5. CAUSE-AND-EFFECT ANALYSIS

5.1 Testing for non-linear root causes

Non-linearity analysis: A non-linearity analysis is useful in determining non-linear root causes of plant-wide

oscillation when the cause is limit cycling in a sticking valve [3,5]

No non-linearity was detected in any tag. It is concluded that the source of the 56 minute oscillation is not a sticking valve in the plant.

5.2 Time delay analysis

Correlation coefficients: A simple test for causality is to find the lag τ at which a time shifted cross-correlation between two variables is maximised. The time shifted correlation coefficient is:

$$r_{x,y}(\tau) = \frac{1}{N-\tau-1} \sum_{i=1}^{N-\tau} \hat{x}_i \hat{y}_{i+\tau}$$

where \hat{x} and \hat{y} are mean-centred and scaled to unit standard deviation and τ can be positive or negative such that $-N+1 < \tau < N-1$. When the maximum correlation has been found then it is tested for significance using a standard significance test. The significance test calculates Pearson's correlation coefficient for the time series shifted by the lag with maximum correlation.

Results: Table 2 shows the significant correlations discovered by correlation analysis. An arrow symbol \rightarrow means that the first variable was ahead of the second while the \leftrightarrow symbol means they are synchronous. In general, the \rightarrow symbol suggests causal relationships which can be illustrated using the causal map representation of Figure 8.

	lag/sampling intervals
LC3 \rightarrow PC1	2
LC4 \rightarrow LC1	6
TC1 \rightarrow LC1	9
TC1 \leftrightarrow PC1	0

Table 3. Results from correlation analysis

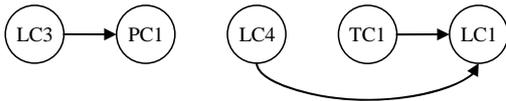


Figure 8. Causal map from correlation analysis

LC1 is the feed buffer tank level, LC3 is the level in the reflux drum of the separator, LC4 is the reactor level, while PC1 and TC1 are the reactor pressure and temperature. The interpretation of these findings will be discussed shortly.

5.2 Transfer entropy test

Method: Correlation analysis has shortcomings when the correlation coefficient is too low to be meaningful or if there is no detectable delay, for instance in the case of TC1 and PC1 in table 2. A new method based on conditional probabilities called transfer entropy has been recently proposed by Schreiber (2000) to measure dependency in time and thus which of two variables causes the other. Transfer entropy can find dependencies that are not detected by correlation analysis. Transfer entropy relates k previous samples of variable X and l previous samples of variable Y to predict the next value of X and thus incorporates the

time dependency. The concept is illustrated in Figure 1(a).

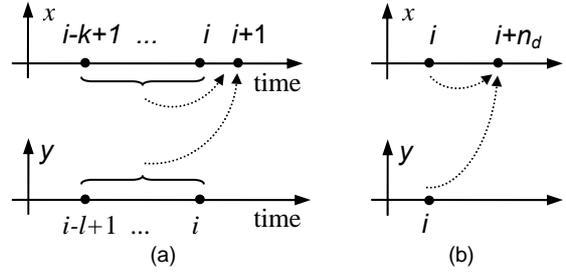


Figure 9. Illustration of the concept of transfer entropy

Transfer entropy is then calculated by summing the joint and conditional PDF of two time sequences logarithmically:

$$T_{Y \rightarrow X} = \sum_{x_{i+1}} \sum_{\mathbf{x}_i^k} \sum_{\mathbf{y}_i^l} p(x_{i+1}, \mathbf{x}_i^k, \mathbf{y}_i^l) \log \frac{p(x_{i+1} | \mathbf{x}_i^k, \mathbf{y}_i^l)}{p(x_{i+1} | \mathbf{x}_i^k)} \quad (1)$$

where the sums are over all amplitude bins. The joint PDF $p(x_{i+1}, \mathbf{x}_i^k, \mathbf{y}_i^l)$ is the probability that the combination of x_{i+1} , \mathbf{x}_i^k and \mathbf{y}_i^l have particular values. The conditional PDF $p(x_{i+1} | \mathbf{x}_i^k, \mathbf{y}_i^l)$ is the probability that x_{i+1} has a particular value when the value of previous samples \mathbf{x}_i^k and \mathbf{y}_i^l are known while $p(x_{i+1} | \mathbf{x}_i^k)$ is the probability that x_{i+1} has a particular value when only the values of previous samples \mathbf{x}_i^k are known.

The basis of the transfer entropy method is the determination of the ratio between $p(x_{i+1} | \mathbf{x}_i^k, \mathbf{y}_i^l)$ and $p(x_{i+1} | \mathbf{x}_i^k)$. If \mathbf{y}_i^l contributes no additional information about x_{i+1} then the two are equal which means the log of the ratio is zero and the transfer entropy is zero.

The index $Y \rightarrow X$ in (1) indicates that the influence of sequence Y on sequence X is measured. The reverse dependency $X \rightarrow Y$ is calculated by exchanging x and y of the joint and conditional PDFs (Fig 9(b)). To specify whether X influences Y more than Y influences X a measure for the directionality is introduced:

$$t_{X \rightarrow Y} = (T_{X \rightarrow Y} - T_{Y \rightarrow X}) \quad (2)$$

If $t_{X \rightarrow Y}$ is significantly greater than zero then X influences Y . Further details including the method for determining the significance threshold for $t_{X \rightarrow Y}$ are presented in [9].

Results: Transfer entropy calculations using the controller error data in Figure 2 showed the dependencies illustrated in Figure 11. Figure 11 shows the same information as Figure 8 using correlation coefficients together with other dependencies that were not captured by that analysis.

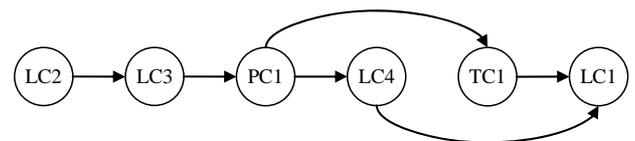


Figure 10. Causal map from transfer entropy analysis

A further calculation which included the controller outputs in the analysis showed, additionally, that the reactor temperature TC1 influenced all controller outputs except LC2.OP.

6. DISCUSSION OF DISTURBANCES

6.1 Slow deviations

The slow drifts: Spectral classification found coordinated low frequency deviations in Tags 1 and 4 (buffer tank and reactor level). The disturbance is local and is confined to those measurements. Figure 3 shows that their spectra are distinctive and different from all other tags and that they mainly occupy a band of frequencies from 0.001 to 0.01 (1000 minute cycles to 100 minute cycles).

The behaviour of these measurements is typical of level control loops tuned for surge control, i.e. to provide a buffer for variations in total hold-up in a process. The similar dynamic behaviour of Tags 1 and 4 suggests the reactor and feed tank are sharing the buffering duty.

Controller output tags 7, 8, 10 and 11 show very long term slow drifts which are principally the expected effect of integral action in these controllers. Tag 9 is the controller output of the reflux drum of the separator and represents the flow of cool recycled material into the reactor. The reason why it has unique behaviour is that the reflux drum level is tracking a varying set point and the distinctive spectral feature at 0.001 on the frequency axis is imposed by the set point trajectory.

6.2 The 56 minute oscillation

The 56 minute oscillation: The oscillation with a period of 56 minutes is a serious one because it is spread throughout the plant.

The absence of non-linearity (section 5.1) shows that the cause is not due to a limit cycle caused by valve friction. There is no evidence in the data set of any sticking control valves.

Process insight: The role of process understanding in a control loop performance study has been highlighted before [5]. Process insights greatly aid the interpretation of a data-driven analysis. Process insights from the BP co-author and information from the process schematic lead to the following additional comments:

- The recycle from separator to reactor contains an intermediate chemical species in the reaction;
- The reaction is sensitive to concentration of the recycled intermediate chemical species and to reactor temperature;
- Reactor pressure is key to the stability of the flashing stream, which is controlled via a DCS controlled valve, and hence the reactor inventory;
- Variations in separator reflux drum level (Tag 3, LC3) reflect changes in the inventory of the intermediate chemical species;
- Variations in separator reflux drum level (Tag 3, LC3) reflect variations in the composition of the stream leaving the reactor via the flash line;
- Liquid recycled from the reflux drum (Tag 9, LC3.op) will be cool and will disturb the reactor. Variations in flow also affect the composition in the reactor;
- Variations in flash tank level (Tag 2, LC2) reflect variations in the amount of an un-reacted component leaving the reactor.

The above reasoning shows that the reaction is important part of the explanation because it is known to be sensitive to composition and temperature both of which are affected by the recycle stream from the separator reflux drum.

6.3 Root cause of the 56 minute oscillation

Possible explanations: Explanations that might be considered for the 56 minute oscillation include:

- Structural oscillation caused by presence of recycle;
- Interaction between control loops such that they compete for control of the same physical quantity;
- A disturbance entering the plant from upstream because of variation in pressure of a gas feed or flow or composition variations in a liquid feed;
- A loop with oscillatory tuning causing the p_v to have a persistent oscillation, which then propagates to cause secondary disturbances in other loops.

To date there are no data-driven signatures that uniquely identify a plant-wide disturbance as originating from one or another of these categories. The present state of the art in controller diagnosis is that the process control engineers have to come up with an explanation that is consistent with both the data-driven results and their process understanding.

Inventory control explanation: The hypothesis that will be put forward to explain the 56 minute oscillation is an interaction between the pressure control and inventory of the recycled intermediate chemical species. The coupling is taking place through the reaction.

The recycling of an intermediate chemical back into the reactor means that a physical feedback loop exists. A plant with a recycle can oscillate just as a feedback control loop can oscillate if the gain is too high. Such an oscillation will be affected by process loop gain (i.e. the extent to which one physical variable influences others in the path). In this case the process loop gain would be affected by reaction rate and also by the presence or absence of buffering capacity within the recycle path.

Supporting evidence: Supporting evidence for the hypothesis is that transfer entropy testing highlighted a causal interaction chain $LC2 \rightarrow LC3 \rightarrow PC1 \rightarrow LC4 \rightarrow LC1$ involving flash tank level (Tag 2, LC2), separator reflux drum level (Tag 3, LC3), reactor pressure (Tag 5, PC1) reactor level (Tag 4, LC4) and feed buffer tank (Tag 1, LC1). All these measurements involve the inventory in the recycle and are explained by the process insights in the previous column, for instance it says that material in the reflux drum level (LC3) upsets the reactor pressure (PC1).

Other evidence is that the variables participating most strongly in the oscillation are reactor pressure and temperature and the inventories of the flash tank and separator reflux drum which contain the products of the chemical reaction and give an indication of their inventory. The reactor level was not found to be oscillating overall because its behaviour was dominated by long slow deviations from set point. However there was an oscillation comprising 24% of the total signal power superimposed upon its low frequency trend. In other words, the reactor level is participating in the oscillation though not dominated by it.

Transfer entropy also showed a causal path $PC1 \rightarrow TC1 \rightarrow LC1$ suggesting an influence from reactor pressure to temperature and feed tank level.

Transfer entropy found LC2 to be at the start of the causal chain and a close inspection of Figure 2 shows the peaks and valleys in LC2 are indeed slightly ahead of all other tags. Figure 2 also shows that the time-lagged correlation with separator reflux drum level (Tag

3, LC3) is positive. It is possible that they both reflect variations in the amount of unreacted and intermediate material chemical passing through the flash valve.

Plant testing: Step testing confirmed the interaction between the inventory in the separator reflux drum level (Tag 3, LC3) and the reactor level controller output (LC4.OP) and vice versa.

The BP co-author reported he has managed to settle the plant by making the level control in the separator reflux tank less tight thus including buffering capacity into the recycle. The effect would be to make the recycle flow more steady and thus to prevent the oscillation from feeding back and upsetting the reactor.

An advanced control solution: An advanced controller was commissioned in order to successfully stabilise the inventory controllers on the unit. This was based upon a FIR process model which showed strong interaction on the inventory controls based upon the step response data gathered on the unit. The decision to break the inventory controls was strongly influenced by the studies carried out in this report, which served to confirm operational experience on the unit.

8. SUMMARY AND CONCLUSIONS

The analysis of the data set to determine the source of the 56 minute oscillation has been conducted through data-driven methods such as spectral classification and oscillation analysis that highlight measurements with similar dynamic characteristics. A new method called transfer entropy was also able to make a causal map showing the direction of propagation of the oscillating disturbance.

All the analyses pointed to the reactor pressure, temperature and inventory of reaction products as the key variables involved in the oscillation. The most likely explanation is that the 56 minute oscillation is caused by an interaction between the pressure and inventory of the recycled intermediate chemical species. The coupling is taking place through the reaction and is sustained

through the recycle path. An understanding of the reaction and the variables which influence was needed in order to make sense of the data-driven results.

The study has shown that data-driven methods can successfully point out areas for investigation but that site expertise is still required to get to the bottom of them. A topic for the discussion session of today's seminar is the need for industrial tools which can capture such qualitative knowledge of the process, for instance in this case the major influences on a reaction. There seems to be an urgent requirement for tools that can link process understanding with a data driven analysis which can help company experts focus on getting early solutions to plant problems.

8. REFERENCES

- [1]. Horch, A., 1999, A simple method for detection of stiction in control valves. *Control Engineering Practice*, 7, 1221-1231.
- [2]. Horch, A., 2002, Patents WO0239201 and US2004/0078168.
- [3]. Choudhury, M.A.A.S., 2004, *Detection and Diagnosis of Control Loop Nonlinearities Using Higher Order Statistics*, PhD thesis, University of Alberta.
- [4]. Gerry, J., and Ruel, M., 2001, How to measure and combat valve stiction online, ISA, Houston, TX, USA. <http://www.expertune.com/articles/isa2001/StictionMR.htm>
- [5]. Thornhill, N.F., Cox, J.W., and Paulonis, M., 2003a, Diagnosis of plant-wide oscillation through data-driven analysis and process understanding, *Control Engineering Practice*, 11, 1481-1490.
- [6]. Hägglund, T., 1995, A control-loop performance monitor. *Control Engineering Practice*, 3, 1543-1551.
- [7]. Thornhill, N.F., Huang, B., and Zhang, H., 2003b, Detection of multiple oscillations in control loops. *Journal of Process Control*, 13, 91-100.
- [8]. Bauer, M., Thornhill, N.F., and Meaburn, A., 2004, Specifying the directionality of fault propagation paths using transfer entropy, *DYCOPS4 conference*, Boston, July 1-4 2004.
- [9]. Bauer, M., 2005, PhD thesis, University of London.