Process Dynamics

Part 1

Operation stability in time series data

Introduction

In this series, “Process Dynamics” is described and explained. Recently, we created a "check-list" for process engineers to investigate their maturity level as process & operation engineer in Oil & Chemical industries, and asked engineers to answer all thirty questions in it. As typical result, it was found that there is one item with low maturity both domestically and internationally. It is knowledge, experience, and understanding of “Process Dynamics”. All units do not function in a static state. The unit functions, in running. And if you move the unit, there will always be process dynamics there. Specifically, all equipment related to process operation, such as pumps, compressors, heating furnaces, and distillation towers, does not function at all when stopped. These units perform a predetermined function only when moved. And, the dynamics of the unit are generated at that time. Whether the operation of the process is stable or unstable, it is represented by dynamics. However, it is not so simple to say whether the stability / instability of the process operation can be if the dynamics of individual equipment are examined. This is because the dynamics of a single equipment affect the dynamics of a continuous unit. Also, the dynamics are not constant. It may change depending on operating conditions and equipment abnormalities. If the operation state of a unit is unstable, the instability may affect the entire process and cause the operation of the entire process to become unstable. In addition, instability may increase by continuous equipment. If there is recycling in the flow of the process, instability may increase in the recycling loop. Therefore, if process operation is to continue stably, it is necessary to understand the dynamics of the entire process and consider identifying unstable factors. Unfortunately, with little understanding of process dynamics, it is quite difficult for process engineers to get guidance on how to catch the destabilizing factors of process operation indicated by their departments and companies. In this training, we would like to explain how to grasp the stability /instability of process operation in situations where the dynamics of individual equipment are difficult to grasp. It would be appreciated if this content could be useful even a little in examining the instability of process operation.

1.DCS and PIMS

DCS stands for Distributed Control System and is now the most common designation for process unit operation systems. The functions of DCS are roughly as follows.

1) Operation variables (temperature, pressure, flowrate, level, etc.) of the process unit are

collected and displayed on the screen as a number or trend graph.
2) When the measured value of operation variables exceeds the upper and lower limits set in advance, an alarm is displayed due to changes in the display color, etc., and at the same time, an abnormality is noticed to the operator with a beep code.
3) It has a control function such as feedback and feedforward based on PID control, and the output value to field controller is calculated from the measurement value of the variable, and the value is transmitted to move the controller.
4) A sequence for safely responding to abnormal situations is incorporated, and the output value is changed to the control unit according to the sequence.
So, common DCS features does not include the "understand dynamics from variable measurements" feature. The driver sees the trend change displayed on the screen and finds an out-of-the-changing movement to grasp the instability of the operation. However, since the number and duration of variables that can be displayed on the trend screen are limited, it is impossible to constantly monitor the operation stability of the entire process.
PIMS stands for Plant Information Management System and is one of the subsystems defined in Level 3 (MOM: Manufacturing Operations Management) of ISA-95 (ISA: the International Society of Automation). It is often referred to by product names such as PI (OSISoft), IP21 (Aspen Tech), PHD (Honeywell), and dataPARC (Capstone). Pre-registered Tag measurements are obtained from DCS via the OPC server at a specified period and stored in database. In addition, it is incidental and generally has the following functions.
1) Data display function: Display the value of multiple variables in trend. The difference from DCS is that it is possible to freely display amount of data to be stored in past several years. In some PIMS, it is also possible to display and compare past data with current data on the same trend screen. Some PIMS have no limit on the number of Tags displayed.
2) Data processing function: A calculation formula using multiple measurement values can be defined using four arithmetic operations, etc., and the result can be stored as a calculation Tag and displayed on the trend screen. The defined formula also can be applied to historical data. Some PIMS can be programmed using C languages and the like.
3) Applications: Some PIMS have various applications such as calculator to correlate between data in the data groups and PLS (Partial Least Square) based soft sensors to expand functions. Also, some PIMS applications are based on frequency analysis such as FFT (Fast Fourier Transformation).
4) Interface: In addition to the interface with the OPC server, PIMS has an interface with the SQL server and an interface with various programming languages. This function allows the use of data collected on user-created programs.
These above is a function of PIMS, and it has a primitive function to understand the process

unit dynamics such as FFT (Fast Fourier Transformation), etc., but the judgment of the stability / instability of the process operation as a result is left to the operator or engineer. However, at most of factories, PIMS is rarely applied as a system to assist in real time operation, even though it collects data online. The most commonly used function is as a data storage system for analyzing data in the vicinity of occurrence retroactively when the trouble is noticed.

2. What can be seen on the trend screen?

In DCS, it is impossible to grasp the dynamic characteristics, and the period of data that can be displayed is limited. So, from here on, we only focus on PIMS and its data.

In typical PIMS, the number of Tags that can be displayed on a single screen is limited. However, if the stability / instability of the process operation is to be observed and evaluated, it is necessary to display the data of all Tags related. In some PIMS, there is no limit to the number of Tags displayed on the trend screen, so we will use that function to display trends and try to evaluate stability/instability.

(Example)

The tags related to the B unit of the A plant is 63 tags in total. The all data for the week was displayed on the trend screen. (Fig. 1, Fig. 2)

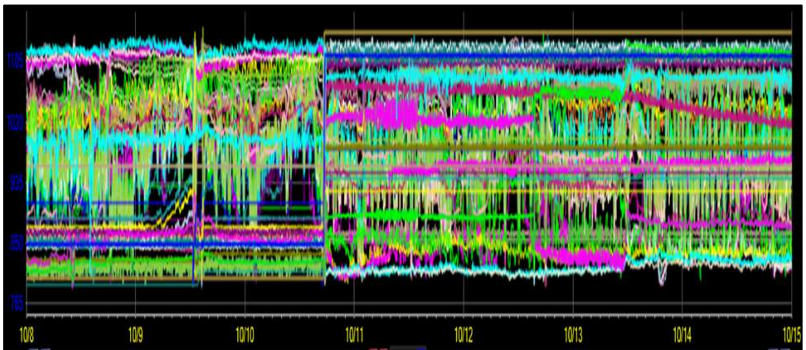


Figure-1 Operation Data (Oct.8th -15th)

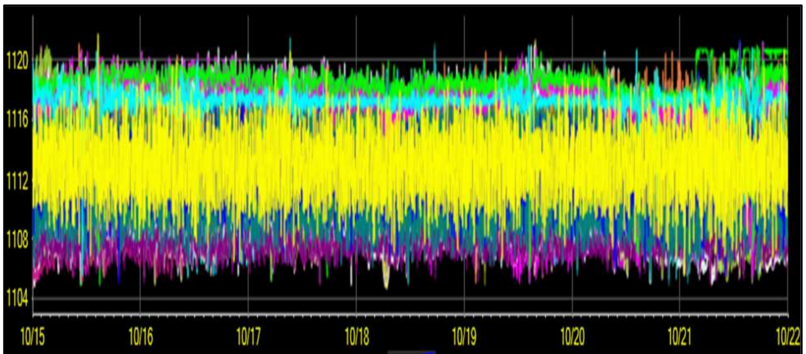


Figure-2 Operation Data (Oct.15th -22nd)

Both graphs are composed of the same group of 63 Tags and in Figs. 1 and 2, the data collection period is one week off. In fact, it was operated with unstable condition during one period and relatively stable during another period. If you are a troubleshooter of process operations who are given these trend graphs with their datasets and have to report which period is under unstable operation, how do you determine and report to your boss?
It is the period of Figure 1 that the trouble occurred in which. Have you seen the answer clearly? We think most of you were wrong in your judgment or that you would have chosen with your senses without clear reason. Then, why was clear judgment not made? This is because there is a limit (lack of information) for engineering to examine and decide stability/instability only with the information on the current trend display. The information displayed on the trend screen alone does not provide enough information for the stability/instability study. So, what information should be obtained other than the trend display information, and can the stability/instability judgment be accurately performed? There is no doubt that one of these technologies which can achieve the accurate judgement is "clustering" that will be explained in next section.

3. Clustering

Since the current pandemic situation, we have often heard the word "cluster". It was a word rarely used before the pandemic. The meaning of this cluster is "when the incidence of disease is a number that exceeds the average or expected value in a certain place", and it is a word that was previously used for the occurrence allocation of cancer etc.
For example, a cluster as a technical term represents a group of computers that operate as if it were a single machine by combining with a "computer cluster", and "cluster sampling" is a sampling technique that performs statistically exact grouping in the population, and "cluster analysis" is used to represent statistical data analysis methods.
Clustering taken up here is a method of doing "cluster analysis". Clustering is to classify the group of data in a similar chunk. The calculation is carried out by the following procedure.
1) Randomly allocate each data to the cluster.
2) Calculate the center for each cluster. In general, arithmetic means are used.
3) Determine the distance from each data to the center of each cluster, and if the distance to the other cluster center is shorter in the distance calculation result, change the affiliation of the data to the shortest cluster.
4) Repeat the procedure of 2) and 3) using the allocation of new clusters.
5) If there is no change in the data for each cluster, the processing is terminated there.
There is a dataset called “iris”, which is often used as an example dataset for cluster analysis. This is involving four characteristics for each of the three types of irises: "sepal length", "sepal

width", "petal length" and "petal width", to perform clustering. Fig. 3 shows the execution result by the petal length (X axis) and the petal width (Y axis).

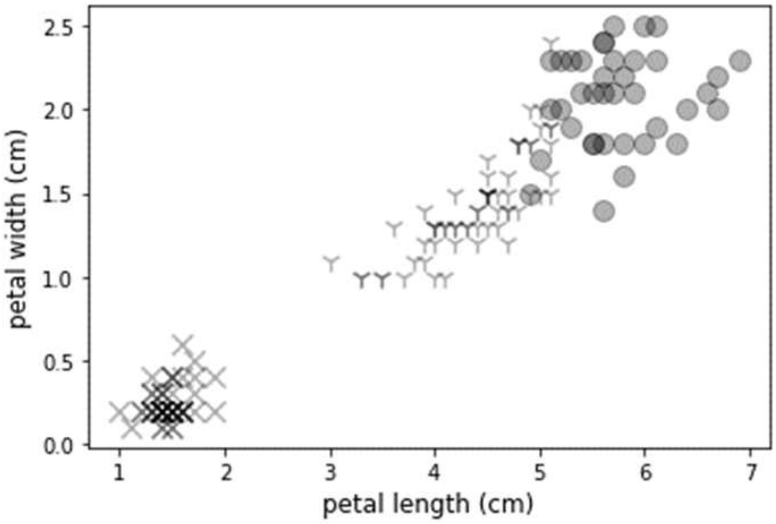


Figure-3 Petal length (X) vs Petal width (Y)

In this figure, it can be seen that it is classified into almost three types of iris species. (Reference: Machine learning by Python (ed., System Planning Laboratory, Inc.) Ohmsha) In this way, even if you do not know which of the three iris species belongs, you can estimate approximately from the length and width of the petals. If it is likened to process operation data, even if data at the time of stability and data at the time of instability are mixed, by classifying the data into several clusters, it is possible to estimate the stable period / unstable period and the ratio of stable / unstable operation period to the entire period from the operation data group. In other words, clustering enables quantitative fault detection using process operation data.

4. Application of clustering to fault detection

Applications for fault detection currently commercialized often use clustering. This is because the change in operation state can be confirmed in chronological order by classifying time series data from the results of clustering.

(Example-1)

Fig. 4 shows the degree to which APC (Advanced Process Control) was stabilized when it was implemented in a place with unstable operation as a result of clustering.

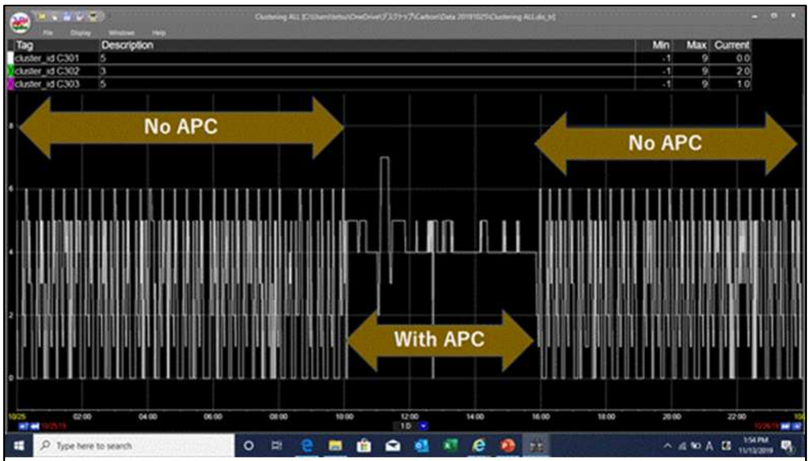


Figure-4 Stabilization with APC

In this unit, normal operation was performed without implementing an APC in the first half. After that, APC was implemented for several hours, then the APC was removed to come back to normal operation. In the time series data expressed in this clustering results, when the cluster change is intense, "operation is unstable", and when the cluster change is small, "operation is stable". Therefore, from this graph, it can be seen that the operation stabilization has been greatly improved when implementing APC.

(Example-2)

The above trend in Fig. 5 shows raw data of time series and the below trend shows the clustering calculation result.

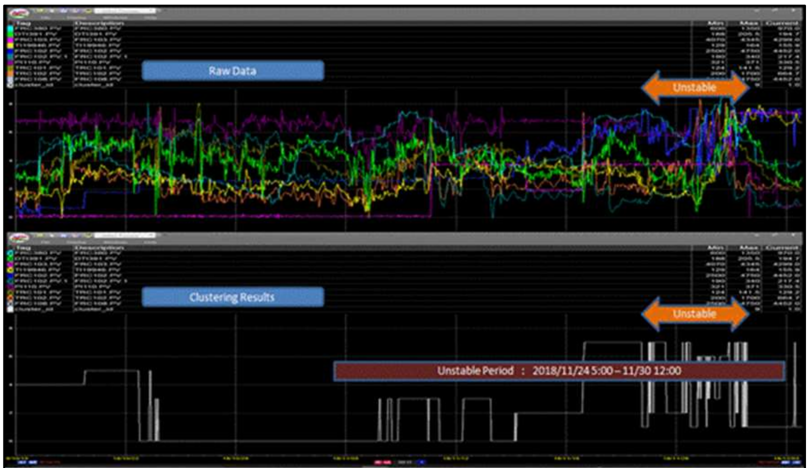


Figure-5 Identification of exact period with instability

From the operation data trend shown above, it will be difficult to identify the period during operation is unstable. However, the clustering results in the figure below can be used to accurately identify how long the operation data is fluctuated (i.e., the operating instability period). If this operation instability period can be identified, it will be possible to check what was happening at this period from the operation logbook etc.

There is always a cause for unstable operation as found by clustering results. Therefore, after identifying the period of unstable operation, it is possible to eliminate unstable operation by identifying the cause and responding.

The main factors that cause instability in operation are as follows.

1) Poor adjustment of the control system (including PID tuning defects)

2) Overcapacity of each equipment and piping (operation that does not reach the design range)

3) Capacity shortness of each equipment and piping (operation beyond the design range)

4) Unbalance flow due to depositing of each equipment and piping (including coking of reactor)

5) Flooding and weeping in distillation towers

6) Wrong control system design

7) Others

If these destabilizing factors are examined individually for the unstable operation period identified by clustering, it is possible to find and confirm the cause of instability.

5. Examples of operation instability that can be identified by clustering

Let's look further at the case of operation instability by applying clustering.

(Example-1)

Fig. 6 shows the operation data of a process that performs with batch operation.

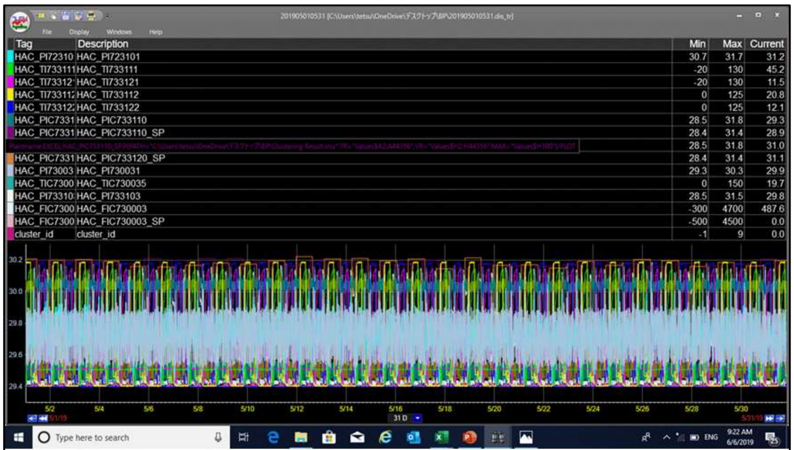


Figure-6 Batch operation data

Operators and process engineers thought this batch operation was stable because the data was displayed in the same cyclic way on the screen. Therefore, clustering was carried out using these data. The result is Fig.7.

Figure-7 Clustering results of batch operation

From this result, it was found that there is a stable operation / unstable operation portion in the one cycle even in batch processing that seems to be stable operation. In this factory, it was made to consider why unstable operation occurs in this period.
Thus, it is easy to be missed in batch operation that there is a partially unstable fluctuation in one cycle.
(Example-2)
From the results of clustering, it is often possible to see the operation dynamics that was not noticed in the trend screen display. Fig. -8 is the clustering calculation result of a unit in some factory.

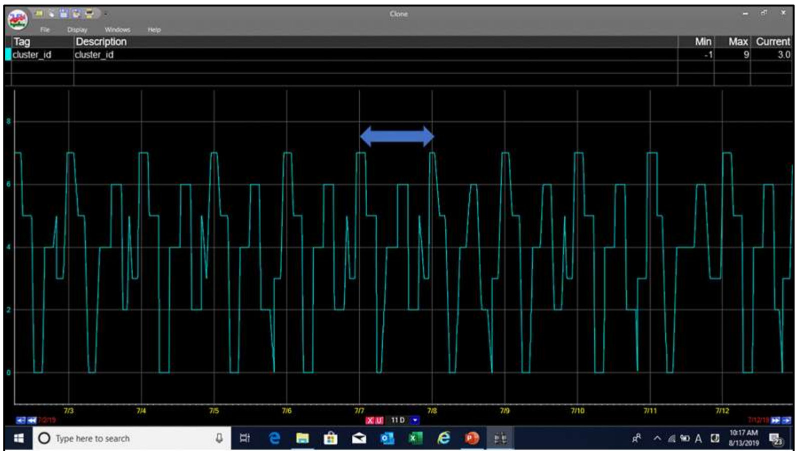


Figure-8 Cyclic fluctuation of operation data

From this result, it was found that a 24-hour cycle fluctuation occurred in the operation of this unit. When the equipment of this unit and the related equipment of other units were examined, it was found that the fuel gas flow introduced from the outside had 24 hours period. Afterwards, it was confirmed that the fuel gas supply unit in the factory was affected by the outside temperature, and it was thought that the influence was the cause of generating the

24-hour cycle. However, there are some movements that cannot be explained only by the outside temperature, so consideration is continuing. As can be seen from the various examples described so far, clustering can accurately identify if the operation of the unit is stable / unstable.

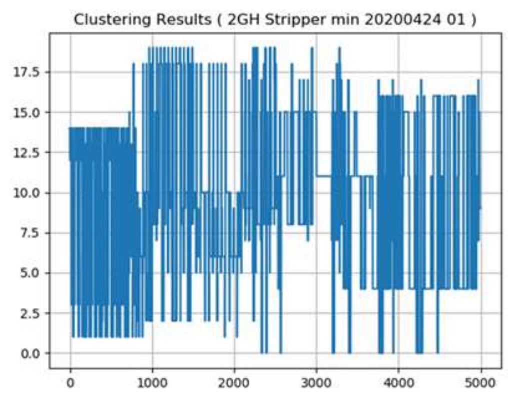
However, it is not possible to confirm how stable/unstable it is. Therefore, Process Stability Indicator (PSI) was developed to quantify operation stability/instability.

6. PSI (Process Stability Indicator)

PSI is an indicator of process operation stability/instability based on clustering calculations. Based on our former experiences, we calculate percentage of the stability points obtained from the clustering results, and further evaluate it on a 5-step basis. These five-step evaluations are as follows.

100% ≧PSI>90% Stable
90% ≧PSI>80% A little Unstable
80% ≧PSI>70% Unstable
70% ≧PSI >60% Very Unstable
60% ≧PSI Extremely Unstable

After calculating Clustering and evaluating PSI (%), our program automatically generates files that report all information, calculation results and evaluation. Fig. 9 is an example of clustering results and PSI calculation and evaluation results in an automatically generated file. In this case, the PSI is 88.4%, and the rating is "A little Unstable".



######### 2GH Stripper min 20200424 01 #########
Number of dataset = 4999
Number of Stable Point = 4419
Process Stability Indicator = 88.4 (%)

Stability Evaluation = A little Unstable

Figure-9 PSI evaluation results (Example)

7. Examples of examination using PSI

Here, we will explain examples of how the evaluation result of PSI is utilized. The following examples are actual examination results at factories that have already implemented this program.

(Example-1)

Fig. 10 is an example in B plant of A Company which is divided into several sections, and clustering, PSI calculation and evaluation are performed using data in April and October.

Section	PSI (Apr20)	Status	PSI (Oct20)	Status
HVN Feed	90.6 %	S.	85.3 %	A.L.U.
Fuel gas and Fuel oil system	93.1 %	S.	**37.6 %**	**E.U.**
Heater	81.2 %	A.L.U.	90.7 %	S.
Tube Skin	93.2 %	S.	85.5 %	A.L.U.
Reactor	92.6 %	S.	94.6 %	S.
Regenerator	67.4 %	V.U.	95.6 %	S.
Spent cat.	**24.2 %**	**E.U.**	**25.6 %**	E.U.
Reduction Gas Heater	94.2 %	S.	93.7 %	S.
Debutanizer Feed-Bottom exchanger	91.8 %	S.	65.0 %	U.
Fines and lift blowers reactors No.3 and No.4	**34.1 %**	**E.U.**	**46.0 %**	**E.U.**
First stage discharge drum	87.2 %	A.L.U.	74.3 %	U.
Second stage discharge drum	**68.5 %**	**V.U.**	**42.4 %**	**E.U.**
Third stage discharge drum	**62.5 %**	**V.U.**	**38.7 %**	**E.U.**

Figure-10 PSI results at B factory of A company

In April data results, there were two sections where the PSI was less than 60%, or Extremely Unstable. However, according to calculations using October data, sections below PSI 60% have increased to six sections. In other words, in this plant, there are significantly more sections of unstable operation in October than in April, and it is considered that product loss and energy loss are increasing, or factors leading to trouble are increasing in October.

(Example-2)

Fig. 11 is a result of examination of D plant of C Company C which is divided into fourteen sections. From this result, it can be seen that the operation of this plant is relatively stable except for Section 14. In addition, even if Section 14 is unstable, subsequent sections are operated stably, and it is inferred that the fluctuation of Section 14 does not affect the stability / instability of the entire plant.

Stability	Section	PSI
Stable	HDO Pump Around	99.47 %
	LDO Pump Around	99.20 %
	Stabilizer	98.91 %
	Kerosene Pump Around	98.41 %
	Hot Reflux	97.79 %
	OVHD & Bottom	97.77 %
	Naphtha Pump Around	96.57 %
	Heater CIT & COT	94.74 %
	Draw Off Level	94.14 %
	Naphtha Splitter	97.67 %
	Heater 011F-101 B	93.87 %
A little unstable	Flow Product	83.04 %
	Heater 011F-101 A	82.82 %
Extremely unstable	Feed	41.50 %

Figure-11 PSI results at D factory of B company

As in these examples, PSI makes it easy to consider what the stability of operation has become in time series data, and how unstable operating sections affect other sections.

In addition, we have recently developed a program that links clustering and PSI calculation and evaluation programs with various PIMS. This program automatically collects data from PIMS at any time or at a pre-reserved time, and automatically performs PSI evaluation using collected data. It also automatically generates ppt file (PowerPoint), xlsx file (Excel) and doc file (Word) containing the results. An overview of the application and stability/instability analysis methods in the frequency band will be described in the next part.