Statistical batch process monitoring

Ramaker, H.J.; van Sprang, E.N.M.

Citation for published version (APA):
CHAPTER 1  ♦ BATCH PROCESS MONITORING*
1.1 General introduction

Batch processes are present in many industries and therefore, play an important role. This chapter serves as an introduction to batch processes and statistical batch process monitoring. The advantages of batch monitoring are explained as well as the concepts of statistical process control (SPC). The current state of batch monitoring is outlined with respect to the application areas and research.

The first section gives an introduction of batch processes and for what products batch processes are deployed. In the second section, the principle ideas of statistical process control are briefly explained since statistical batch process monitoring originates from this. Also, some alternatives of statistical batch process monitoring are discussed. Furthermore, the advantages of multivariate control charts versus univariate charts are highlighted. The section ends with an overview of the current state of batch monitoring.

1.2 Batch reactors

Industrial batch reactors are in many ways the simplest reactors. Reactants are instantaneously administrated in a large stainless vessel, which is usually equipped with a stirrer to keep the reactants well mixed. Next, by adding heat or initiators the reaction is initiated and after a certain amount of time the reaction is finished and the reactants have been converted to the desired end-products. After that, the reactor is ready to produce a new batch. There are no reliable numbers concerning the market share of batch processes compared to continuous processes, but this market share is estimated to be around 50%.
1.2.1 Different Application Areas of Batch Processing

Batch processing is done in many industries because of several advantages (these will be discussed later). To show the diversity of products that are produced batch-wise, a list is given of different application areas. (This list is not intended to be complete).

- Chemical industry: polymer reactions (Latex, polystyrene, methylmetacrylate, PVC etc.), resins, coatings.
- Food industry: fat hardening, flavor products, beer.
- Pharmaceutical industry: antibiotics, powder mixing.
- Bio-chemical industry: waste water plant, fermentation, agricultural products.

Clearly, batch processes are common in a variety of industries. Every industry itself has its own characteristics. This is also reflected in the way batch reactors are operated. As an example, in the food industry hygiene is very important. In the pharmaceutical industry regulations for the production of e.g. medicines is very strictly controlled by the FDA (Food and Drug Administration). This makes the introduction of new measurements or recipes more laborious.

1.3 Features of a Batch Reactor

Various features characterize a batch reactor. The most important features are highlighted in the following.

- Batch processes are recipe driven. If during the development phase a completed batch run is known to have produced an on-spec product, the most important process variables from that batch are stored in a recipe. This recipe is then used for future batch runs aiming at consistent production of the on-spec product. This means that initial concentration; temperatures, pressures or dosing profiles of the reactants are fixed. The recipe itself is maintained using precise sequencing and automation of all required operations. During the lifetime of a batch reaction, the recipe itself is continuously adjusted and optimised.
- The reaction time of a batch is finite. The reaction time is often referred to as batch time.
- The batch duration can change a lot between different types of batch reactions. For example, a polymerisation reaction may only take half an hour while a fermentation reaction can take several days.
- The conversion of the reactants is usually high (~ 90%).
- The behaviour of a batch process is non-linear and non-steady-state.
- Batches processes can be operated in stages.
- Batches can also be operated in semi-batch mode. This means that reactants are being added into the reactor slowly instead of instantaneously. Semi-batch reactions have the advantage of good temperature control and the capability of minimizing unwanted side-reactions through the maintenance of a low concentration of one of the reactants.

1.3.1 ADVANTAGES OF BATCH PROCESSING
For large volume production, it is not advantageous to use batch processing. In such cases continuous production is preferred. However, batch processing is flexible in the sense that different products can be made using the same equipment. This enables the supplier to quickly respond to the customer needs. In principle, only the recipe needs to be changed in order to produce a different product. Some industries like the food and pharmaceutical industry are highly regulated.

The batch products can be tested and certified batch-by-batch, and from a regulatory point of view, it is easier to deal with. Furthermore, their low-volume, high-value products characterize these industries.

1.3.2 ON-LINE MEASUREMENTS OF BATCH DATA
Batch processes are often monitored using real-time measurements. The ideal on-line measurement is simple, disturbance insensitive, cheap and accurate. Unfortunately, the lack of instrument robustness or sample preparation time causes that most plant variables are difficult or even impossible to measure on-line. Process variables in a plant can be categorized according to M. Soroush [1]:

-8-
### Table 1

<table>
<thead>
<tr>
<th>Name of group</th>
<th>Brief description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic process variables or</td>
<td>Measured readily on-line, to ensure proper operation of</td>
<td>Temperature, pressure, liquid levels, flow rates, feed composition, stirring speed, compressor duty, power supply.</td>
</tr>
<tr>
<td>engineering variables</td>
<td>plant.</td>
<td></td>
</tr>
<tr>
<td>Plant-product quality indices</td>
<td>Monitored by process engineer to ensure proper operation of plant, rarely available on-line, obtained by laboratory sample analysis</td>
<td>Viscosity, melt viscosity, density, composition, pour point, flash point, octane number, molecular weight distribution.</td>
</tr>
<tr>
<td></td>
<td>specifications, complex and not well understood, until final product is ready, abstract variable (not quantifiable like colour or taste)</td>
<td></td>
</tr>
</tbody>
</table>

Definition of process variables.

Data from batch processes often concern the first category of Table 1. This set of data is often referred to as X-block or X-data. End-product quality indices (Y-data) are very interesting from a monitoring point of view. Unfortunately, X-data is often not completed with the Y-data. In the following, two types of engineering variables are briefly discussed. Some figures will be shown to give the reader a taste of what batch data looks like.

1.3.2.1 Engineering variables

It was already explained that batch data show non-linear dynamics. By looking at batch data this becomes obvious. The conditions over time are continuously changing, for example deactivation of catalyst in a reactor, fouling of a heat exchanger or fluctuation of concentration and/or flow rate of the feed, etc.
An average number of on-line measurements for a batch process lies between 7 to 15 process variables. In general most datasets are dominated by temperature and pressure measurements. However, a priori, it is not always known which variables are useful for statistical process monitoring.

The first example of typically engineering variables from a batch process concerns the production of paper pulp (TEMBEC-Canada).

For this process, twenty-one process variables are measured on-line. Because of company policy reasons, the time units for the x-axes are omitted as well as the measurement units on the y-axes. Some process variables show a dynamic behaviour with a good signal to noise ratio such as the digester pressure and bottom temperature. Other process variables show dynamic behaviour but suffer from a poor signal to noise ratio (e.g. steam header temperature, steam header pressure). It can be seen that some process variables represent the state of a valve (open/close).
Sometimes the state of a valve is given as a percentage between 0 % (close) and 100 % (open). Also, the batch-to-batch repeatability for one process variable can be better than another process variable. Another example concerns the production of polyvinyl chloride (PVC) (A.A. Tates et al. [2]). The on-line measurements of fifteen process variables are presented in the following figure:

![Figure 2](image)

*On-line measurements of a batch process for the production of PVC.*

Again, the x-labels and y-labels are omitted. In contrast to the paper pulp data, the heat management of this process is monitored (flow cooling water, reactor temperature, jacket temperature). Since the production of PVC is highly exothermic, it is important to monitor the heat management of the process for safety reasons. Sometimes it is possible to calculate new process variables from the measured ones. These (new) calculated process variables could be added to the data set.

By studying the trajectories of the process variables over time, valuable information can be obtained. This is the first step in analysing batch process data.
1.3.2.2 Spectroscopic data

Another type of batch data comprises the measurement of spectroscopic data (NIR, UV-VIS, Raman). Although commonly this type of data is also referred to as engineering variables, their nature is quite different from e.g. temperature or pressure measurements.

Spectroscopic measurements give mainly information about the chemistry of the process and to some extend physical information whereas temperatures and pressures describe mainly the physics. Also, temperature and pressure profiles are often examined separately, while spectroscopic data is examined as a whole. Furthermore, the number of measured process variables is much larger compared to the traditional engineering datasets. It is not uncommon to measure more than 300 wavelengths. The treatment of spectroscopic data for multivariate analysis is also different (e.g. no scaling because all channels are measured in similar units). In the following figure, spectroscopic X-data from a urethane resin polymerisation process is presented. The data represents the evolution of a single batch run and will be discussed in more detail in Section 5.3.

![Figure 3](image)

*On-line NIR spectroscopic measurements of a batch process for the production of a urethane resin.*

-12-
It is quite common to present spectroscopic batch data in these 3D pictures. It is obvious that visual inspection of spectral data for non-experts is cumbersome. The analysis of batch data from spectroscopic sensors requires some additional knowledge not commonly known by chemical engineers.

Besides engineering variables and spectroscopic measurements there are more process analytical techniques to monitor the evolution of batch processes. Examples of such techniques are:

- Microcalorimetric monitoring (A. Menert et al. [3]).
- Ultrasound spectroscopy (C. Turner et al. [4]).
- Electronic tongue (C. Turner et al. [4])

1.4 Why batch process monitoring?
Process monitoring is advantageous for many reasons. This is, of course, not only true for batch processes but for industrial processes in general. However, the focus here is on batch processes.

1.4.1 SAFETY
Some batch reactions, like the polymerisation of polyvinyl chloride, are highly exothermic. These exothermic reactions produce a considerable amount of heat that needs to be withdrawn from the reactor to avoid dangerous situations. In such case, it is important to monitor the temperature management. Therefore, process monitoring is used for safety reasons.

1.4.2 PROCESS DISTURBANCES
In the previous example, a process disturbance can lead to undesired side reactions due to temperature upsets. Sometimes the filling of the reactor completely fails because the wrong valve is opened and a different reactant is added to the reactor. As a result, undesired products are formed or unsafe situations can occur. Therefore, process monitoring is done to prevent that process disturbances lead to environmental stress and economic losses.
The batch time of biochemical processes such as the production of antibiotics can take up more than days. Early detection of process disturbances for these types of batch processes is highly wanted. First, detection of disturbances makes it possible to correct the process if possible. Secondly, if the process cannot be corrected anymore, the decision is taken to kill the current process in order to start a new one. If such a decision can be made supported by process monitoring in an early stage, valuable batch time is gained. Therefore, process monitoring leads to time and cost reduction.

1.4.3 PROCESS UNDERSTANDING
One way to optimise a process is to study its behaviour. Knowing exactly how the process reacts on changing process conditions can help process engineers to make adjustments to the recipe of a batch process. Process monitoring is a useful tool for these purposes since the interaction between the various process variables is monitored. The effect of process adjustments can immediately been monitored in multivariate control charts. Process monitoring therefore is helpful for improving the process understanding.

1.4.4 PROCESS AND PRODUCT CONSISTENCY
The consistency of how the batch recipe is realized is reflected in the measure process variables. There will always be a certain amount of variation present in the process parameters. This variation can influence the quality of the end-product. Take as an example the fact that the personnel work in shifts. Every shift will work with slightly different process tuning parameters. This will lead to variation in the batch time of the process and/or product quality. It is important for the producer to reduce this variation as much as possible so that his concurrent position improves. Products that do not fulfil the quality constrains are often mixed with the product from a good batch. In this way, the product still meets the product quality requirements. Process monitoring therefore helps to make the process under consideration more consistent.

1.4.5 CAPABILITY INDEX
In industrial statistics it is common to express the performance of the process consistency in relation with the product consistency in terms of capability indices. The process capability is the repeatability and consistency of the process with respect to
specifications limits of a product parameter. Next, the capability indices are used to analyse whether a process, operating under normal conditions, is capable of meeting the specifications limits of the end-product.

As can be seen from the previous examples it is advantageous to monitor batch processes. These advantages include topics such as environment, health and safety, cost reduction, process consistency and process understanding.

1.5 Statistical process control

The theory of statistical batch process monitoring originated from statistical process control. For that reason, it is interesting to take a look at the concepts of statistical process control (SPC) and the development of statistical batch process monitoring through the years.

1.5.1 Ideas of statistical process control

The basic idea of the statistical approach of process control is that every industrial process is liable to variations. This process variation will affect the quality of the end-product. When the fluctuations in the quality of the end-product are unacceptable, it is worthwhile to study the process itself to see where the process variation enters and how large it is. This thought was developed in the early 20’s of the past century (W.A. Shewhart [5]). Initially SPC was developed for the manufacturing industry where the end-product quality is directly measured (e.g. the diameter and length of a nail). Later on, the end-product quality measurements where substituted by measuring the process conditions. Ideally, the observed variance within the process conditions is related to end-product quality.

'Statistical thinking' has two starting points. First, it is assumed that improvement of the end-product is reached by reduction of process variation. To improve the quality of the end-product it is useful to study the variation in the process. Eventually, process monitoring leads to adjustments of the process.

Secondly, it is assumed that this process variation can be well monitored with the use of control charts. Shewhart distinguishes two different kinds of variation:
- Process variation that is inherent to the process itself (Common cause variation).
- Process variation that has an external cause (Special cause variation).

In short, it is aspired to operate a process that is stable, has a minimum of variation and produces an end-product that is on spec. In a process, the following situations can occur:

1) The process is statistically in-control and the end-product fulfils the predetermined specifications. This is the desired situation. Process faults are detected immediately. While detecting the process fault, the cause can be located.

2) The process is statistically in-control but the product does not fulfil the predetermined specifications. This is an unwanted situation. The process is stable, but adjustments are needed. There are two possibilities. First, the process variation is too large and needs to be decreased by e.g. revising the process design. Secondly, it is also possible that not all the required measurements are performed. In other words, the process measurements indicate that the process is stable and reliable while somewhere in the process some process variation is apparent that influences the end-product quality. Further analysis by e.g. experimental design and additional measurements is desired.

3) The process is statistically not in-control and the product fulfils the predetermined specifications. This is also an unwanted situation. During the running time of the process it cannot be predicted whether the product will meet the product specifications. It is also possible that the product will not fulfil these specifications at all. In such cases, localising the cause will be quite difficult.

4) The process is statistically out of control and the product does not fulfil the predetermined specifications. In these highly unwanted situations the whole process needs to be thoroughly examined.
1.5.2 DIFFERENCE BETWEEN PROCESS MONITORING AND PROCESS CONTROL

There is often confusion between process monitoring and process control. Sometimes process monitoring is actually done while it is named process control. This is for instance the case for (multivariate) statistical process control (MSPC). (M)SPC is actually a technique for monitoring processes but the designation suggest that process control is performed.

Process control from a chemical engineering point of view boils down to the automated monitoring with well-defined control actions of a process, in which a computer system is used to regulate the usually continuous operations or processes. Equipment that measures the variables of an industrial process, directs the process according to control signals from the process computer system, and provides appropriate signal transformation. SCADA (supervisory control and data acquisition) is a software program for process control where data is collected in real time from remote locations in order to control equipment and conditions. In the field of process control terms like P, PI and PID controller are commonly used. It is important to note that process control for batch processes is limited because of the non-linear behaviour.

However, the definition of statistical process control is different. That is, SPC is a well-established set of on-line, off-line and at-line mathematical tools, supported by a wealth of software and manual systems, to determine and monitor the level of statistical control demonstrated by a production process. In other words, the level of statistical control is monitored. This is where the term monitoring came in. Therefore, the terms SPC and process monitoring have the same meaning when used in the ongoing to avoid misunderstandings. As of now, multivariate statistical process control of batch processes is referred to as statistical batch process monitoring.

1.5.3 METHODS FOR BATCH PROCESS MONITORING

This chapter gives a summary of possible methods for batch process monitoring. These methods are for the majority based on SPC methods since SPC for batch processes originates from SPC. In order to give a more complete overview also some alternative methods for batch process monitoring are shortly mentioned.
1.5.3.1 Fundamental models
One choice for batch process monitoring is to use fundamental models (P. Nomikos & J.F. MacGregor [6]). These models are based on fundamental mass or energy balances that are commonly found in the field of chemical engineering. With the help of such models estimates of the underlying theoretical states of the process are provided. Often these fundamental models are based on state estimation methods. Such methods combine a fundamental model of the process with on-line process measurements to provide on-line, recursive estimates of the underlying theoretical states of the batch processes (e.g. using a Kalman filter) (P. Nomikos & J.F. MacGregor [6], K.A. Kosanovich et al. [7]).

However, using theoretical fundamental models have some serious drawbacks. Only for a very limited number of processes the models are accurate enough. The chemistry of e.g. a biochemical reaction is far too complicated to be captured in a good working fundamental model. Furthermore, constructing such models is often very time consuming. This makes the use of e.g. state estimators impractical for batch process monitoring purposes.

1.5.3.2 Knowledge based approach
The most pure form of a knowledge-based approach to monitor batch processes is the recipe driven approach that is commonly used. The recipe is composed and adjusted conform the experiences of process operators and researches. In case of process faults, e.g. a sensor failure, the knowledge of the process operator is utilized to take action.

It is obvious that such knowledge is transient since the loss of the process operator also includes the loss of process knowledge. However, each behavioral description can be associated with causal assumptions in a hierarchical structure. For each node in the hierarchy, diagnostic rules can be generated from the behavioral descriptions. Thus the process model is represented by a set of qualitative and quantitative descriptions based on the knowledge of process operators and engineers. These are called rule-based expert systems. This approach is advantageous since it requires no model. However, formulating the behavioral and causal descriptions is just as difficult and time consuming as constructing a fundamental model.

Another approach to the problem of fundamental modeling is the use of neural nets (P. Nomikos & J.F. MacGregor [6]). These neural nets can be used as
pattern identifiers in order to diagnose fault conditions in a batch process. However, such a model needs to be trained with data that contains numerous faulty batches. This is practically never the case for batch processes. Interrogation of neural nets to assign process faults is difficult. Therefore, neural nets are no good candidates for batch process monitoring in terms of fault diagnosis.

1.5.3.3 SPC
Statistical batch process monitoring originated from the field of SPC. The fundamentals of statistical process control were described by W.A. Shewhart [5]. H. Hotelling [8] was one of the first who introduced a multivariate approach for statistical process control to monitor a process of a multivariate nature. J.E. Jackson [9] then applied principal component analysis (PCA) to reduce the dimensionality of the multivariate data. J.V. Kresta et al. [10] proposed a multivariate control technique for continuous processes that uses PCA or projection to latent structures (PLS). Finally, P. Nomikos & J.F. MacGregor [6] extended the concept of multivariate statistical process control for batch processes.

It is important to understand the development of statistical batch process monitoring since terminology and concepts from different steps in this development are often mixed. This leads to confusion and unnecessary mistakes. Therefore, a short overview is given of the development of statistical batch process monitoring starting with classical SPC.

1.5.3.3.1 Univariate charts: $\bar{X}$, CUSUM, EWMA
Univariate charts are based on the measurement of only one property or individual observation such as the mean of e.g. the diameter of a nail. The univariate charts are widely used in e.g. manufacturing industries and make up an important part of quality improvements programs. In the following, the most common univariate charts will be discussed briefly (D.C. Montgomery [11], R.E. Walpole & R.H. Myers [12]).

1.5.3.3.1.1 $\bar{X}$ Chart
Walther S. Shewhart first proposed the general theory of statistical control charts in the early 30's of the last century. A typical control chart is a graphical display of a quality characteristic that has been measured from a sample versus the sample number or time (Figure 4).
In this figure, three lines can be distinguished. The centerline represents the average value ($\mu$) of the quality characteristic. This value corresponds to the in-control state. The other two horizontal lines are the upper control limit ($UCL = 3\sigma$) and lower control limit ($LCL = -3\sigma$). As long as the points fall between the UCL and LCL, the process is assumed to be statistically in-control. The opposite is true if a point plots outside these control limits.

The process is assumed to be out of control. Action is required to locate the cause of this abnormal deviation from the in-control situation. It should be noted that there are different rules for appointing a signal to be out of control (D.C. Montgomery [11]).

The setup for the $\bar{X}$ chart will be discussed according to a small example (D.C. Montgomery [11]). The quality characteristic for a manufacture process of piston rings is given by the diameter of the piston ring. It is known that the average value for the diameter ($\mu =$ process mean) is 50 mm and the process standard deviation is $\sigma = 0.015$ mm. Ideally, the outcome of the average diameter $\bar{x}$ measured for $n$ samples of pistons is 50 mm. In practice, this $\bar{x}$ average will always deviate from $\mu$. For sample sizes of e.g. $n = 10$, the standard deviation of the sample average $\bar{x}$ is calculated according to equation 1:

$$
\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.015}{\sqrt{10}} = 0.0047
$$

This is all the information needed to compute the UCL and LCL. In fact, the following hypothesis test is performed:
With a probability $\alpha$ of erroneously determining that the process is out of control, the null hypothesis is accepted if $\bar{x}$ falls between $50 - Z_{\alpha/2}(0.0047)$ and $50 + Z_{\alpha/2}(0.0047)$. Thus, the UCL = $50 + Z_{\alpha/2}(0.0047)$ and LCL = $50 - Z_{\alpha/2}(0.0047)$. Sometimes, the value for $Z$ is taken to be three. Typically, this choice is referred to as the three-sigma limits. The UCL and LCL then are plotted in the control chart. The centre line CL of the control chart reflects the ideal value of the average size of the piston diameter. An example of such a chart is given in Figure 5.

It is important to note that it is assumed that $\bar{x}$ is normally distributed due to the central limit theorem. Also, it is assumed that the process mean and standard deviation are known.

Closely related to the $\bar{X}$ chart are the $R$ chart (center line is the range for the data) and $S$-chart (monitoring the sample standard deviation). The disadvantage of the $\bar{X}$ chart is that small changes of the mean are not detected. Only shifts in the mean of magnitude $1.5\sigma$ to $2\sigma$ or larger are effectively detected.
1.5.3.3.1.2 CUSUM chart

The cumulative sum, or CUSUM chart, is a good alternative to detect small shifts of the mean and was first proposed by E.S. Page [13]. Consider a control chart for the mean with a target for the process mean $\mu_0$, with the following observations $x_1, x_2, x_3, \ldots, x_n$. The cusums are calculated according to:

$$
S_1 = x_1 - \mu_0 \\
S_2 = S_1 - (x_2 - \mu_0) \\
S_3 = S_2 - (x_3 - \mu_0) \\
\vdots \\
S_n = S_{n-1} - (x_n - \mu_0)
$$

Often, the starting value for the CUSUM chart is set to zero. Small changes in the mean will result in a relatively large increase in the slope (positive or negative) of the CUSUM chart. For this reason, tabular cusums and the V-mask form of the cusum are used. In the following figure, a CUSUM chart is presented using a V-mask see D.C. Montgomery [11] for a more detailed discussion.

---

Figure 6

*Example of a CUSUM chart with V-mask. Source: http://www.ub.rug.nl/eldoc/dis/eco/j.e.wieringa/c5.pdf*
1.5.3.3.1.3 EWMA chart

Another alternative to detect small shifts of the mean is the exponentially weighted moving average (EWMA) chart. The performance of the EWMA chart is equivalent to the CUSUM chart, although the EWMA is easier to set up and to operate (D.C. Montgomery [11]). The following chart statistic is plotted:

\[ z_t = \lambda x_t + (1-\lambda)z_{t-1} \quad (3) \]

Here, \( \lambda \) is a constant value between zero and one. The starting value of \( z_0 = \mu_0 \), where \( \mu_0 \) is the starting value of the average value \( \mu \) of the quality characteristic. If all the previous values \( z_{t-1}, z_{t-2}, \ldots, z_0 \) are substituted in equation 3, it follows that

\[ z_t = \lambda \sum_{i=0}^{t} (1-\lambda)^i x_{t-i} + (1-\lambda)^{t+1} z_0 \quad (4) \]

For the calculation of the chart statistic \( z_t \), the means of the past and current are considered. However, the ‘older’ the mean, the less weight it receives. The centre line, UCL and LCL for the EWMA chart can be calculated as follows:

\[
\begin{align*}
UCL &= \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda} \left[ 1 - (1-\lambda)^2 \right]} \\
CL &= \mu_0 \\
LCL &= \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} \left[ 1 - (1-\lambda)^2 \right]}
\end{align*}
\]

An example of an EWMA chart is given in Figure 7:
In the upper graph of Figure 7 a $\overline{X}$ (Shewhart) chart is given of an arbitrary process. In the lower graph of Figure 7 an EWMA chart is given of the same process. As can be seen, the EWMA chart gives a better performance in detecting a small shift from the mean.

1.5.3.3.1.4 Average run length

For univariate charts, decisions regarding sample size and sampling frequency are evaluated through the average run length (ARL) of the control charts (D.C. Montgomery [11]). The average run length is defined as the average number of observations up to and including the first-out-of-control observation. The ARL is a function of $\delta$ where $\delta$ represents a shift of the mean. The ARL ($0$) for any Shewhart chart can be calculated as follows:

$$ARL = \frac{1}{P(0)}$$

(5)

Where $P$ is the probability that any point exceeds the limits in the control chart.

For a $\overline{X}$ chart with three sigma limits ($P(0) = 0.0027$) the ARL ($0$) when the process is in-control is 370. In other words, every 370 observations an out-of-control
signal is generated while the process is in-control. This is also called a false alarm. The average time to signal (ATS) is calculated as follows:

\[
ATS = ARL \cdot h
\]

Where \( h \) is the time (in hours or seconds) that samples are apart.

1.5.3.3.2 Multivariate control charts

The univariate charts typically deal with single observations. However, when the variables are correlated (which is often the case for an industrial process), superimposing univariate charts is not a very accurate method of monitoring processes because relationships between the variables are not taken into account. In this section, the focus will be on multivariate data originating from continuous processes. The reason for this is that the concepts of multivariate process control for batch processes are similar to multivariate process control for continuous processes. However, in transferring the methodology from continuous to batch processes, there are some issues that are not straightforward. In the following, the concepts of multivariate statistical process control for continuous process are highlighted. This will be helpful when the theory on statistical batch process monitoring is extensively discussed.

Process measurements of \( J \) process variables are obtained at \( K \) time intervals on a continuous process and collected in the data matrix \( X (K \times J) \). Preprocessing the process data \( X \) is a first step as will be discussed later. For the following discussion, it is assumed that the data is properly pre-processed. Furthermore, \( X \) is believed to represent in-control observations.

1.5.3.3.2.1 Continuous processes without dimension reduction

In 1947 Hotelling introduced the \( T^2 \) statistic as a method to apply statistical quality control to correlated data with a multivariate nature (H. Hotelling [8]). This summary statistic enables easy control charting and allows for detecting faults of highly correlated measurements. In the case of continuous processes for chemical applications, \( K \) measurements of \( J \) process variables are measured every time interval. These measurements can be arranged in a matrix \( X (K \times J) \). For a new measurement \( x_{new} (J \times 1) \) the test statistic \( T^2 \) is calculated according to:
Statistical batch process monitoring

\[ T^2 = (\mathbf{x}_{\text{ev}} - \mathbf{x}_a) \mathbf{S}^{-1} (\mathbf{x}_{\text{ev}} - \mathbf{x}_a) \]  

(7)

where \( \mathbf{S} (J \times J) \) is the estimated covariance matrix of \( \mathbf{X} \) and \( \mathbf{x}_a \) is the target value or grand mean. Here, the sample size \( n \) of one single measurement equals 1. Therefore, if the goal is to obtain an in-control set of observations as described in the training phase, the UCL is given by (N.D. Tracy et al. [14]):

\[ \text{UCL} = \frac{(K-1)^2}{K} B(\alpha; J/2; (K - J-1)/2) \]  

(8)

where \( K \) is the number of samples, \( J \) is the number of process variables and \( B \) is the Beta distribution with a confidence level \( \alpha \) and \( (J/2, 1-J-1/2) \) degrees of freedom. The UCL is used to detect changes from the in-control process for new independent observations are given by:

\[ \text{UCL}_{\text{ev}} = \frac{J \left( K^2 - 1 \right)}{K(K-J)} F(\alpha; J; K-J) \]  

(9)

where \( K \) is the number of samples, \( J \) is the number of process variables and \( F \) is the F-distribution with a confidence level \( \alpha \) and \( (J, K-J) \) degrees of freedom.

An important note here is that all measured process variables are used and no dimension reduction takes place. With a large number of correlated process variables this approach breaks down, e.g., \( \mathbf{S} \) might become (nearly) singular.

### 1.5.3.3.3 Megavariate control charts

This section discusses the situation where dimension reduction is required for multivariate statistical process monitoring. This has some important consequences for e.g. the type of control charts that is used. Furthermore, the concepts of this approach are very similar to those of statistical batch process monitoring.

#### 1.5.3.3.3.1 Continuous processes with dimension reduction

In a chemical process, during process operation many process variables are collected and the dimensions of \( \mathbf{X} \) may be become large in the second mode \( J \). Since this large number may cause singularity problems, it is advantageous to reduce the dimensionality of the data. PCA is a multivariate statistical method well suited for this
purpose. J.E. Jackson [15] was one of the first who applied a principal component analysis in this context.

A PCA decomposes the matrix $X$ into the sum of $R$ outer products of scores $t$ and loadings $p$ (systematic part) plus a residual part $E$.

$$X = \sum_{r=1}^{R} t_r' p_r' + E = TP' + E$$  \hfill (10)

where $T (K \times R)$ is the score matrix, $P (J \times R)$ contains the loadings and $E (K \times J)$ is the residual matrix. A geometrical representation of equation 10 is given in Figure 8.

![Graphical representation of projecting a sample in a dimension reduced space.](image)

Once the model plane, spanned by the columns of $P$, is defined, the scores and residuals for a new-scaled measurement $x_{\text{new}} (J \times 1)$ are found by projection on the model plane. This is a regression problem:

$$x_{\text{new}} = Pt_{\text{new}} + e_{\text{new}}$$  \hfill (11)
The $D$-statistic describes the Mahalanobis distance from the projection on the model plane to the centre of the model plane, and describes the systematic variation in the data. The centre of the model plane corresponds to the average process behaviour. The $SPE$-statistic represents the Euclidean distance between the measurement vector and its projection on the model plane. This distance describes the variation in the data that is inconsistent with the model.

The test statistics are calculated according to

$$
\hat{t}_{\text{new}} = (P'P)^{-1}P'x_{\text{new}} \quad \rightarrow \quad D = t_{\text{new}}'S^{-1}t_{\text{new}} \quad (12)
$$

$$
e_{\text{new}} = x_{\text{new}} - P\hat{t}_{\text{new}} \quad \rightarrow \quad SPE = e_{\text{new}}'e_{\text{new}} \quad (13)
$$

Here, $S$ ($R \times R$) is the variance-covariance matrix of the scores $T$ ($I \times R$) and is diagonal since the scores are orthogonal. Subsequently, the test statistics are plotted in the control charts and $x_{\text{new}}$ is assigned to be statistically out of control if one of the control charts signals. The UCL for the $D$-statistic is the same as given in equation 9 except that $f$ is replaced by $R$. The UCL for the $SPE$ statistic is the critical value $SPE_\alpha$.

The derivation of this critical value is given by J.E. Jackson [15].

1.5.3.3.4 Statistical batch process monitoring
The principles of statistical batch process monitoring are very similar to the concepts for monitoring continuous processes in a dimension reduced space. First, a sufficient number of batches is collected that represent normal operating conditions (NOC batches). This data set is then modelled in a reduced space using PCA. The latent variables (scores and residuals) form the reference distributions of these batches. The scores form the basis of the $D$-chart whereas the residuals form the base for the $SPE$-chart. The control limits for these control charts are calculated from their reference distributions. Then, the multivariate measurement (at time interval $k$) from a completely new batch is projected on the same dimension reduced space as the NOC batches. The $D$ and $SPE$-statistic for this new batch are calculated from the scores and residuals. These statistics are plotted in the control chart. In this manner, a batch can be monitored in an on-line fashion. It can also be desirable to perform a so-called post-batch analysis. As soon as a new batch is finished, the collection of all the measurements over time are to be projected in the dimension reduced space at once.
The foregoing is a description of statistical batch process monitoring in a nutshell. In the next chapters, every aspect of this approach will be discussed in a step-wise manner. This is called the I.T.A.-trajectory.

**1.5.4 Multivariate charts versus univariate charts**

It was shown in the previous chapters that univariate charts could be used for process monitoring. For batch processes, a $\bar{X}$ chart can be constructed for each separate process variable. Then, the average value for each process variable is plotted at each time interval $k$. However, these univariate charts have some drawbacks that can be dealt with by using multivariate charts for statistical batch process monitoring. Some of these disadvantages are briefly discussed in the following.

**1.5.4.1 Abundancy of control charts**

The average batch reactor is equipped with ten to thirty sensors that measure typical engineering variables such as pressures and temperatures. If univariate charts are being used, one control chart is constructed for every single process variable. This leads to an overabundance of univariate control charts. This can be considered as a drawback since it is inconvenient to monitor 30 charts at the same time.

This problem becomes even worse when a spectroscopic instrument is mounted to the reactor. Such an instrument records more than 500 –1000 wavelengths every time interval. This can be overcome by using multivariate charts: the $D$ and $SPE$-chart. The performance of the process can be followed over time in only two control charts. When the chart statistic is out-of-control in one of the charts, the cause of the process fault can be localized using contribution plots (P. Miller et al. [16], J.A. Westerhuis et al. [17]). The ideas about contributions plots are discussed in more detail in the ongoing.

**1.5.4.1.1 False alarm rate**

If a chart statistic is plotted outside the control limits while the process is actually in-control, a false alarm occurred. The probability of a false alarm using $n$ multiple univariate charts simultaneously plotting in-control is not $1 - \alpha^n$. If a process is in-control, the probability of plotting $n$ means in-control is $(1 - \alpha)^n$ since each chart has a probability of $1 - \alpha$. Thus, the joint probability of a type I error is much larger: $1 - (1 - \alpha)^n$. This gives no trust in the control charts by the process operator. Often, the
process operator picks out only one chart that is of value to the operator. The other charts are being ignored because of the high false alarm rate.

1.5.4.1.1 Multivariate nature
Statistical batch process monitoring models capture the auto-correlation and cross-correlation among the process variables. This information is translated in the $D$ and $SPE$-chart. This gives the operator an extra tool and valuable information about the batch process. If the correlation changes due to a process upset, a signal is given in one of the control charts. The detection of a fault in multivariate control charts is illustrated in the following figure:

![Detection of a fault in multivariate control charts.](image)

Clearly, the fault that occurred around time interval 6 -10 is detected in the $SPE$-chart. Univariate control charts for the two faulty process variables are given in the following figure:
It can be seen from Figure 10 that the process fault goes unnoticed using univariate control charts. This happens because the multivariate nature of the process variables is not taken into consideration.

New events that are not captured by the PCA model are detected in the SPE-chart (e.g. a sensor failure). Sometimes the modelled correlation among process variables is still intact but the joint variation is larger than normal. This will result in an out of control signal in the $D$-chart. In this way, faults can be categorized or diagnosed according to the detection in one of the control charts.
1.6 Nowadays monitoring of batch processes

It is interesting to know how monitoring of a batch process is actually performed in practice. First, the theory of statistical batch process monitoring is only ten years old (T. Kourt [18]). It is common that academic research is implemented in industry with a time delay. Clearly, there has been a strong trend in the ascent of data acquisition systems (SCADA) and distributed control systems (DCS) in process plants. It seems that systems like SCADA is well suited for data driven approaches such as statistical batch process monitoring. However, SCADA software originally is designed in a process control environment. The demand for process control software is mainly reliability and a long lifetime. Therefore, developments in this type of software are slow, not to mention of incorporating monitoring tools. Also, the management of the large amounts of data that are being collected is cumbersome. This makes it e.g. almost impossible to withdraw historical data from a database system in a simply manner. These are serious problems that need to be overcome first. However, there are rules within the pharmaceutical industry, dedicated by the FDA, how to automate batch processes. This forces the producer more or less to build modular batch control software. This makes accessing the data a lot more efficient and therefore more attractive to statistical batch process monitoring.

The most common way that industrial batch processes are operated is according to a recipe. Customized software enables the producer to operate these batches fully automated. This is probably the most widely used method for batch process monitoring. Many applications exist where the implementation of these software modules lead to improved product quality and reduced loss of operating time. These software packages do not contain sophisticated monitoring tools like multivariate control charts and the success of these packages can be ascribed to the increased level of knowledge about the batch process itself then due to the software.

However, also commercially software packages are available that offer the opportunity for on-line statistical batch process monitoring. The major drawback of these software packages is that quite often they are not compatible with the platforms used in the process plant. It is foreseen that the next generation software packages will
overcome this obstacle. It also requires a mentality change of the personnel that often sticks to the classical way of monitoring processes.

Considering this positive development together with the increasingly becoming important data-flood problem, it seems the share of statistical batch process monitoring tools will grow compared to conventional monitoring systems.

1.6.1 LITERATURE & SOFTWARE
The pioneering work in this field is presented in the work of Nomikos and MacGregor in the early nineties. Based on their work, several extensions and modifications have been proposed in the last 10 years. In the following, an overview is given of the most relevant publications in this field. The publications are listed under the headings: pioneering articles, statistics, extensions/- methodologies and comparison and evaluations.

Table 2

<table>
<thead>
<tr>
<th>Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pioneering</td>
</tr>
<tr>
<td>Statistics</td>
</tr>
<tr>
<td>Extensions &amp; Evaluation</td>
</tr>
</tbody>
</table>

Literature overview statistical batch process monitoring.
There are many commercial statistical software packages available. However, commercial software focussing on batch processes is limited. The two leading software packages are SIMCA-P+ developed by Umetrics in Sweden and MSPC+ developed by MDC technology in the U.K. Both packages can be used for off-line and on-line process monitoring.