Introduction to Statistical Process Control

Primary Knowledge Unit
Participant Guide

Description and Estimated Time to Complete

This Primary Knowledge (PK) unit provides an overview of Statistical Process Control (SPC) and how it relates to MEMS fabrication. Statistical Process Control, often referred to as SPC, is a set of tools used for continuous improvement and quality control of an active manufacturing process. There are two (2) suggested activities that reinforce the material presented in this PK as well as a final assessment.

In this unit you learn the basics of SPC, its terminology, and some of the tools used to help ensure a quality production line.

<u>Estimated Time to Complete</u>
Allow approximately 30 minutes to read through this unit.

Learning Module Objective / Outcomes

<u>Objectives</u>

- To explain process variation and the need to identify special cause variation.

<u>Outcomes</u>
You should be able to describe why Statistical Process Control is needed when manufacturing a product and you should be able to apply the basic tools of statistics and Shewhart rules to interpret a control chart.

Terminology (Glossary at the end of this unit)

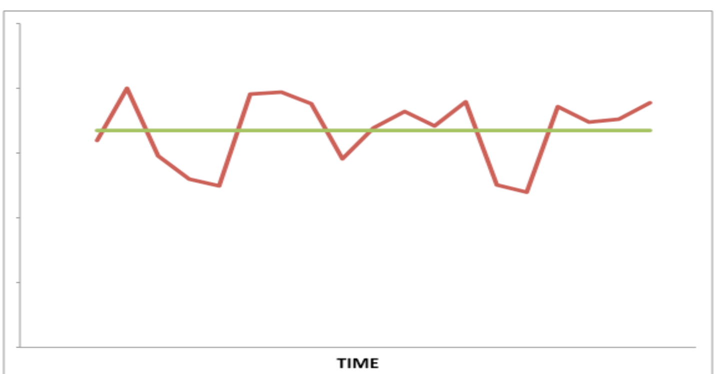
Statistical Process Control
Common or inherent cause variation
Special cause variation
Sample Median
Sample Mean
Sample Range
Sample Variance
Sample Standard Deviation

Introduction – Why is Statistical Process Control Important?

We can all agree that when manufacturing a product, it is desired to produce a "quality" product. This can be said whether we are talking about cars, food, medicines, or microsystems. There is no universally accepted definition for "Quality"; it is a subjective term full of meanings and connotations. Given the *needs* of a customer, we can say that quality is the *realization* and *control* of characteristics that determine whether the product will in fact satisfy those *needs*. *Realization* includes the design of the product. *Control* includes the control of deficiencies in the product minimizing the variation around desired nominal values or "targets".[Reference Mike Leeming]

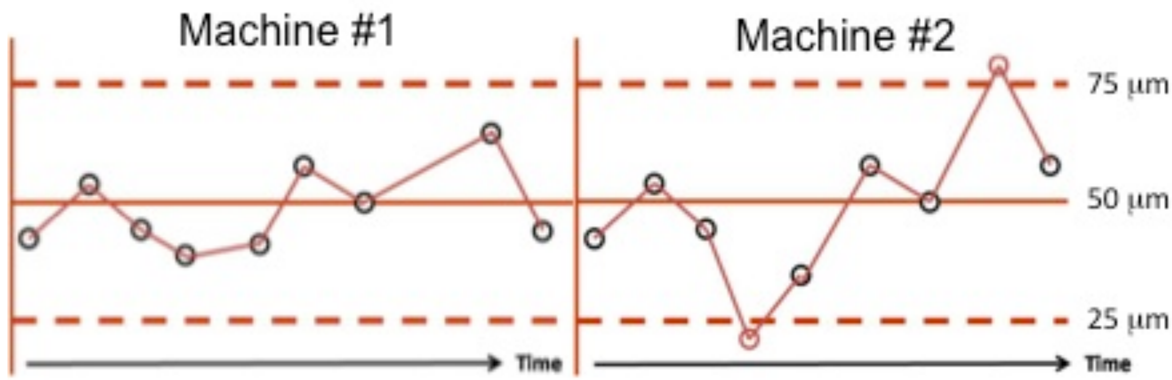
Statistical methods are used throughout the life cycle of a product, which are aimed at the *realization* and *control* of certain product characteristics. For example, methods of statistical experimental design or Design of Experiments (DOE) may be used in the design phase of the product life cycle or in efforts to improve the control of certain product characteristics. This is achieved by identifying key factors (e.g., thin film thickness, process temperature and pressure, line widths) that need to be controlled during the manufacturing process of the product.

Statistical Process Control (SPC) is used real-time during the manufacturing process where in-line data is attained from the processes that produce the products. Statistical methods are then used to assess whether or not the process is in a state of control. This statistically based process information can provide a greater understanding of the process by providing a graphical interpretation of the variation in the process. All processes have some variability over time, as illustrated below. The graph could represent the variation in oven temperature, photoresist thickness, or number of defective die on a wafer. Have you ever cooked a soft-boiled egg? Is the outcome "exactly" the same every single time you cook it or is there a little variation?



Variation is a natural and commonly occurring phenomenon but not all variation is created equal. A process may contain variation that is *common or inherent* to the process and, there may also exist variation that is NOT common or inherent to the process. Variation that is NOT common would be a result of a *special cause* outside of the normal process conditions.

Let's start off with an example of variation. Take a look at the following two graphs (which we will later call control charts). Each graph shows the resist thickness results for a photoresist application or "coat" process. The graph on the left is for Machine #1, and it shows the resist thickness results from one piece of equipment. The graph on the right (Machine #2) shows the resist thickness results from another piece of equipment. Both Machine #1 and Machine #2 are running the same process. Is the process variation over time the same for each piece of equipment? How would you respond to this data if you were working on this production line? Statistical Process Control and Control Charts are tools that help to graphically represent different aspects of a process. These tools are meant to assist engineers and technicians with producing a quality product.



Statistically, each piece of equipment shown in the previous graphs applies a target of a 50 micrometer (µm) thickness of photoresist to a wafer, but as the graphs show, the final resist thicknesses vary differently for each machine.

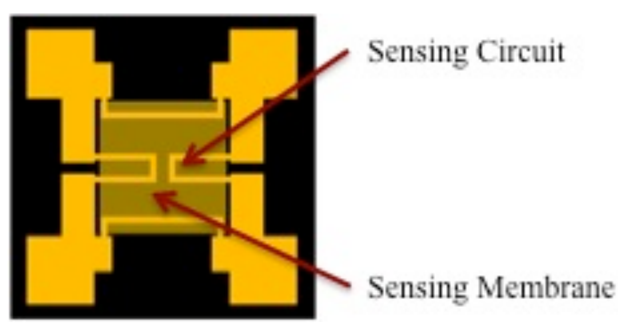
Studying process variation can provide insight into the sources of variation and ways to minimize the variation in the manufacturing process. This knowledge can help lead to greater consistency in the final product and less deficiencies or defects. The use of statistics makes good sense in quality, because even when all seems to be running well, there are many uncontrolled production factors that can affect product characteristics. When manufacturing a product, most of the factors are unknown, can vary, and may not affect the process all of the time. The unknown factors provide the ingredients of a probabilistic environment and natural "background noise" so that it is impossible to predict or calculate exactly how products and their deficiencies will vary. Under these circumstances methods of probability and statistics are applied so that predictions can be made and those involved in the manufacturing of the product know what to expect. Controlling quality is a science, and the mathematics of quality is probability and statistics.

A person does not have to be a statistician in order to correctly use and interpret the various SPC tools. However, one needs to understand and correctly apply statistics terminology and notation when using SPC for quality control. It is important to be mindful of accuracy when collecting and interpreting data. Understanding the statistical tools used in the quality control of a manufacturing process helps to formulate data-based predictions or decisions rather than just "guessing". Quality control provides mathematical clues and data which provide the framework to accurate problem solving. As in all problem solving ventures, communication is key; therefore, for SPC to be effective, all findings and results should be communicated with team members. Sometimes action is required and SPC can prove to be a valuable tool when troubleshooting process issues.

"*The long-range contribution of statistics depends not so much upon getting a lot of highly trained statisticians into industry as it does on creating a statistically minded generation of physicists, chemists, engineers, [technicians], and others who will in any way have a hand in developing and directing the production processes of tomorrow*." - Dr. Walter E. Shewhart, 1939

Variation

All products, whether being man made or nature made, are not exactly created equal. There is a *natural* or *inherent variation* in all processes. In a field of 3 leaf clovers, you won't have to look too hard to find either a 2, 5, 6, or even 4 leaf clover. When chickens lay eggs, the size, thickness of shell, color of the yoke, number of yokes, and the color of the shell all vary from egg to egg.

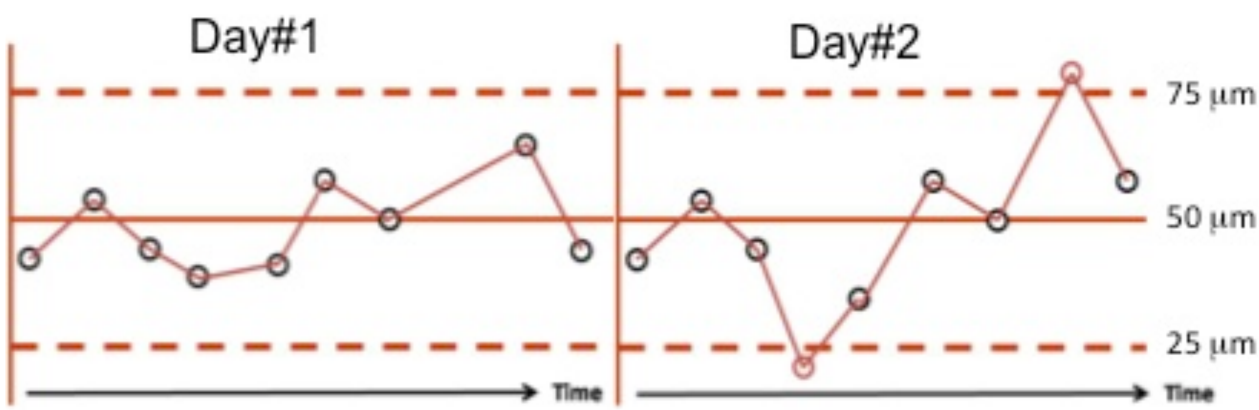


When considering a manufacturing process, variation becomes even more prevalent. Each manufacturing process contains one or more process steps, and each step has its own variation. For example, a micro-pressure sensor's process may include depositing a layer of silicon nitride (for the membrane) on a bare silicon wafer, followed by a photolithography step and an etch step that patterns the reference chamber hole on the backside of the wafer. Another photolithography step patterns the sensing circuit using photoresist on the front side of the wafer, followed by a metal deposition on top of the photoresist. Subsequent process steps remove the unnecessary metal and etch the reference pressure chamber on the wafer's backside. This is just a brief summary of a sample micro-pressure sensor process, but as you can see, there are many different process steps required to create this micro-device.

Each individual process step has many different parameters or variables. Each of these parameters can vary or drift during the process. For example, in the coat process of the photolithography step, it is desired to have a specific thickness of photoresist. The resulting thickness depends upon or is a function of the spin speed of the chuck on which the wafer sits, and the actual viscosity of the photoresist deposited on the wafer's surface. If one or both of these variables change during the coat process of a batch of wafers, then the final thickness of the photoresist will change from one wafer to another. Small changes in these variables may be acceptable as a *natural* or *inherent variation* of the process. Any change outside of this inherent variation is unacceptable and the cause must be determined and corrected. Consistency is key and any change in variables that is unacceptable can lead to defects either immediately or in a subsequent process step. This is where SPC comes into play. SPC helps us determine statistically what the inherent variation of the process is, and then when compared to the product specifications, determine if that variation is acceptable or not.

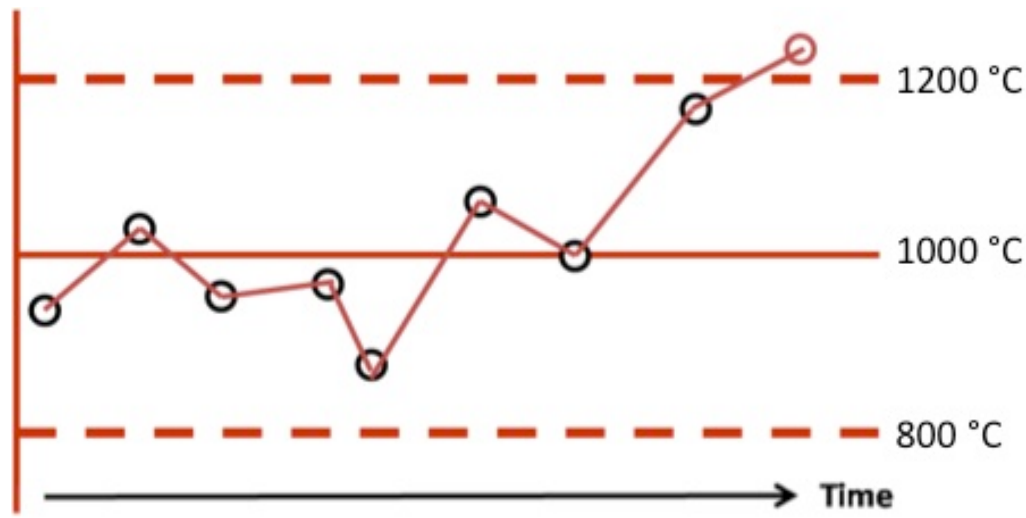
Types of Variation

With any process there are two types of variation: variation that is inherent to the process and variation that is not inherent to the process. For example, let's return to these two process graphs again. Assume that these two graphs or control charts represent the variation of photoresist thickness in a microsystems process over time. However, this time, let's assume that the graphs are from the same piece of equipment, but on two different days. Has something changed?



As you'll soon learn, the control chart on the left is "in control" and the variation that you see is expected because it has been statistically determined that this is the variability inherent to the process. This is the variation due to the "background noise" that we mentioned earlier. This variation is referred to as "*common cause variation*". However, the control chart on the right tells a different story. Notice the two points outside of the dashed lines. These points represent variation that is neither predictable nor inherent to the system. This type of variation is due to an "assignable cause" that can be known, or most often is unknown. Such variation is referred to as "*assignable cause variation*" or "*special cause variation*". A problem solving team should be able to determine the cause of this variation and eliminate it from the process. As you can see from this example, control charts used in SPC help determine which type of variation (common or special cause) is present. As a result, control charts can help determine the best course of action.

Let's look at another example. Take a look at the control chart shown below that is tracking process temperature. As you can see, there is a sudden increase in temperature and the last data point is outside the upper limit of 1200 °C.



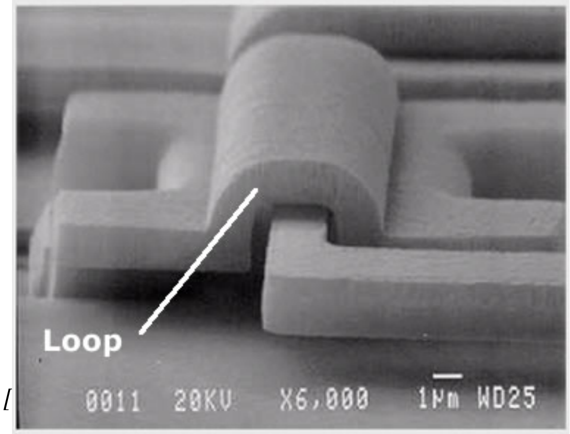
By monitoring the process this way, it is easy to pick up on a process change such as the temperature increase. After an investigation, an example of a *root cause* for this sudden jump in temperature may be due to not enough coolant in the cooling system. The cooling system could be checked, if a coolant is needed, than it would be added to alleviate the problem.

So would one ever want to decrease the variation inherent to the system? Of course you would! However, inherent variability cannot be reduced until the process is in control and all special cause variation has been eliminated.

How much variation is too much? That depends on the product. There are some cases where limited variation is intentionally built into the process. Can you think of a product where variation is desired? Many hand made products are valued because of the inherent variation, such as hand carved furniture, pottery, and jewelry. There are also products for which a certain amount of variation is acceptable or not recognizable by the consumer, such as the color of a car, the weight of a box of cereal, or the thread count in a set of sheets.

Microsystems, on the other hand, seeks to have as little variation as possible. Microsystems deal in micrometers and nanometers, not centimeters and inches. Because of the extremely small feature sizes (1 μm to 100 μm) and precise alignment (sometimes in the nano range), microsystems fabrication needs to be on "target" from step to step, over hundreds of steps. Remember the example of the micro pressure sensor process? There are so many steps involved in microsystems fabrication that when variation occurs at one step, it can perpetuate and affect other steps along the way.

Another microsystems example is this hinge and loop system shown on the right. The height of the loop and the hinge is approximately 1 μm. The hinge must be free to rotate within the loop. Can you see the importance of variability reduction or even elimination in the fabrication of this device? How much would the hinge's height and width be allowed to vary for the hinge to be able to rotate within the loop?



SPC and Variation

Statistical Process Control (SPC) assists in the proper management of variation. Certain types of variation require action by management. Other types of variation often do not require management intervention and can (depending on company policy and training) be remedied by people who are not in management (e.g., equipment operator or technician or both). Companies normally leave special cause variation in the reliable of hands of the people that are in daily contact with the process – operators, technicians, supervisors, and engineers. However, reducing inherent variation requires extensive knowledge of the process, the product, the equipment, and all of the factors that together, create this variation. It may also require process changes and even product redesign; therefore, reducing inherent variability is normally left in the hands of management; although, operators and technicians may be asked to be on the problem solving team.

It is important to emphasize once again, that communication is essential between operators, technicians, engineers, and management in order to understand and remedy both common and special cause variation. Changes should never be made to a process without properly following the protocol in place in a manufacturing facility. Changes could affect all processes run on each piece of equipment, and equipment is often shared between different processes and products.

In-line variation reduction is a major goal of Statistical Process Control. Interestingly, the way in which products and process characteristics vary about some desired target value can actually provide valuable information that can be used to understand, control, and reduce the variation. With predictable processes that also have less variation, the producer is better able to change or control things to provide improved products that are more consistent and on target. More consistent products that are on target require less inspection, scrap, and rework. An erratic process full of variation is similar to the automobile that has a lot of play in the steering mechanism. You have little control and less freedom in the driver's seat as you have to keep adjusting and hoping you will stay on the road.

In Statistical Process Control, although variation reduction is the goal for consistency, variation can also be a valuable source of information. For example, variation provides signals that can be used to assure us that whatever is being observed or measured is behaving (varying) as expected (naturally) or to warn us that it is not. Variation can provide an assessment of ongoing process predictability. The monitoring of variation provides early information that may prevent deficiencies from occurring at all. Variation can provide an assessment of process capability by showing how consistently a process will produce a product within any ideal target range. Understanding the process variations provides a deeper understanding of the product. With this improved understanding of what causes variation and how to control it, the developer is in an improved position to develop new, improved versions of the product and improved manufacturing processes.

The Data for Statistical Process Control

In order to understand the basics of Statistical Process Control (SPC), there are several statistical concepts you should be familiar with. These concepts are used to provide basic analytical descriptions of data. These statistical calculations include sample median, sample mean, sample variance, sample range, and sample standard deviations. Here, we briefly discuss these statistical concepts that you need to understand in order to employ SPC.

There are also many different types of charts that help to describe data in different ways in order to achieve quality control. Some examples include Run Charts, Frequency Distribution Tables, Histograms, Stem and Leaf Plots, Box Plots, Scatter Plots and Control Charts. This lesson discusses Control Charts and its uses in Statistical Process Control. For more information on the other charts mentioned, please review the *Basics of Statistics* provided as part of the **Statistical Process Control Learning Module**.

Data Samples

When employing Statistical Process Control, statistical analysis is usually done on a *sample* of data, hence the statistics which we cover here are for a sample set of data. All calculations use the data from the sample.

One set of calculations of SPC identifies the central tendency of the sample data. This set includes the *sample median and mean*. Remember, the word "sample" is used because the median and the mean are the median of the data in the *sample set* and the mean of the data in the *sample set*.

- The *Sample Median* represents the data value that is “physically” in the middle of the sample set when arranged in numerical order. (Examples provided below.)

- The *Sample Mean* is the central location of the sample data on a number line. This central location is the "target" that was referred to previously, or the center of the distribution. (Examples provided below.)

Another set of calculations addresses variability or dispersion in the sample data. The statistics that are usually calculated are the *sample range*, *sample variance*, and the square root of the sample variance which is called, *sample standard deviation*.

- *Sample Range* is the spread of the distribution of the data, or more simply put – the difference between the maximum value minus the minimum value. For example, given the following sample of voltage measurements, what would be the sample range?

  5,5.3, 4.8, 5.7, 6

  *Answer:*6 (maximum) –4.8 (minimum) = 1.2 (Sample Range)

- *Sample Variance* is an averaged squared difference between the individual measurements and the mean. (More on variance in the next section)

- *Sample Standard Deviation* is a measurement of how the data values are distributed around the sample mean and within the range of values. This is one of the most important calculations for SPC analysis and is used to develop control charts. Both common cause variation and special cause variation rely on the standard deviation for analysis. This will be discussed in detail shortly.

Let's get specific about our sample size now and define variables. Take a random sample of size *n* from a population of *x*'s ($x_1, x_2, \ldots, x_n$).

For example, say you have a random sample of five values from a population of *x*'s. These values are (x_1, x_2, x_3, x_4, x_5). For this set of data, your *n* would be equal to five (n=5). For the following calculations, let us use a numerical example to demonstrate the necessary calculations. Say you have obtained a random sample of voltage measurements, (2,4,9,1,4) volts.

Sample Median

To obtain the sample median, first put the numbers into ascending order as follows: (1, 2, 4, 4, 9). The sample median is the middle number, 4. The sample median is 4 volts. Had the sample been of even sample size, the median would be the average of the two center numbers. For example, suppose the sample was (2, 4, 1, 5, 1, 3). Ordering the data we have (1, 1, 2, 3, 4, 5). The two middle numbers are 2 and 3; thus, the median would be the average of 2 and 3, or 2.5.The sample median is a random variable $\tilde{X}$. The numerical value of the sample median is represented by $\tilde{x}$.

Sample Mean

The arithmetic mean, μ, of a set of data is the sum of the data values, x_n, divided by the number of data values (n):

$$\mu = \frac{\Sigma x_n}{n}$$

When a set of data is identified as a *sample*, the arithmetic mean is calculated for the set of data, but can sometimes be represented as $\bar{X}$.

Let us refer to the earlier sample of 5 voltages (2,4,9,1,4) volts.

$$\mu = \bar{X} = \frac{(2+4+9+1+4)}{5} = \frac{20}{5} = 4\ volts$$

A mean of 4 volts tells us that the central location of the distribution based on the sample data is 4 volts. In the case of the following equations, μ will be used as the sample mean, but keep in mind that it can also be referred to as $\bar{X}$ when referring to the arithmetic mean of a sample.

Sample Variance

When we are looking at a sample set of data, it is desired to quantify how spread apart the data values are with respect to each other. Understanding how far the numbers are from the mean is just one of the statistical characteristics which can help us to understand the nature of the process.

The equation for *sample variance* is

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1}$$

Where$\sigma^2$is sample variance, x_i are the individual sample data values, $\mu$is the sample mean, and n is the number of sample data values. Using our sample voltages, (2,4,9,1,4) volts and sample mean ($\mu = 4$), we will begin our calculation by looking at the differences between the individual sample data values, x_i, and the sample mean, μ.

(2-4, 4-4, 9-4, 1-4, 4-4)=(-2, 0, 5, -3, 0)

Note that each of the resulting numbers shows how far above or below each data value is from the sample mean. The sum of the differences (deviations from the sample mean) should *always*

be zero. If not, then there was a mistake in a calculation. In the case of our data, we have –2+0+5-3+0, which does sum to zero.
Continuing to calculate the sample variance, you now square each of the differences (-2, 0, 5, -3, 0) to get (4, 0, 25, 9, 0). Add the squares to get the *sum of squares*, 38. Finally, divide the sum of the squared differences by *n-1*. In the case of our sample data, we divide by 5-1 or 4 to get

$$\textit{sample variance} = \sigma^2 = \frac{38}{4} = 9.5\ square\ volts\ (V^2)$$

Remember that each of the data values represents a voltage reading; therefore, the squared value of the variance has the unit of V^2.

Sample Standard Deviation

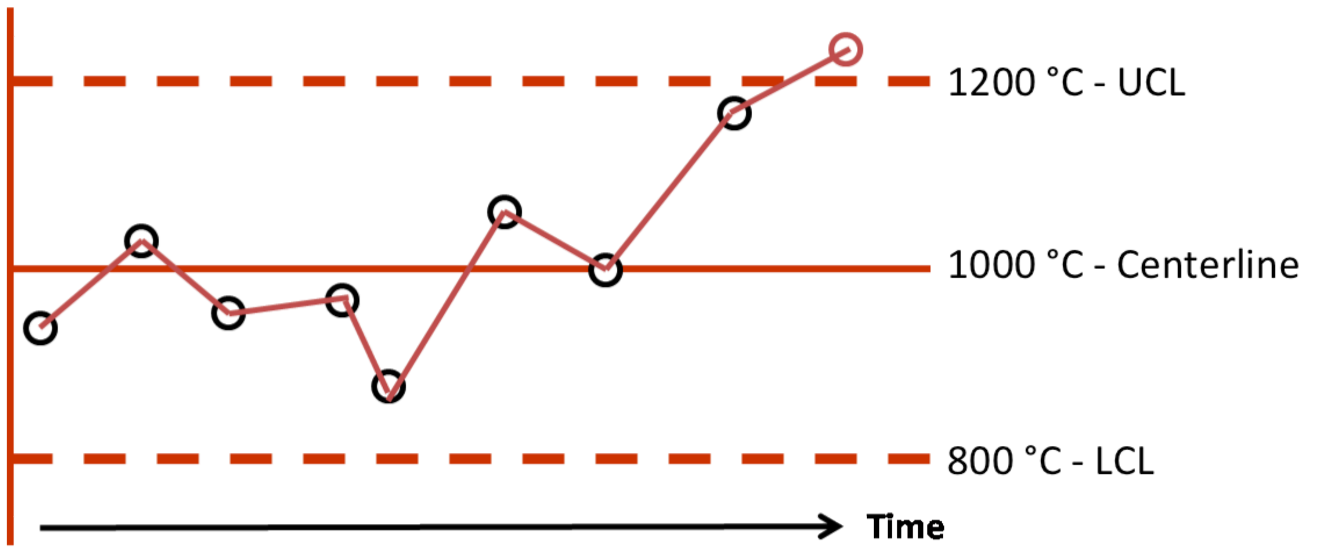
The standard deviation is simply represented as the square root of the sample variance. Sample standard deviation is an estimate of the root mean square deviation from the mean.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1}}$$

Using our sample data, the square root of the sample variance, 9.5 $volts^2$ is 3.08 volts. Therefore, the sample standard deviation σ, is 3.08 volts.

The Data and Control Charts

How does this statistical data from the sample help us create our control charts? Let's revisit the control charts we spoke of earlier. If you recall, the control chart monitoring the process temperature looked like this.



There are many types of charts used to monitor different types of data. It is very common in industry to base the limits for controlling the process on the sample mean, μ or as you recall, $\bar{X}$. This type of chart is commonly referred to as a Mean or $\bar{X}$-Chart. When this type of chart is used, the middle line, or the centerline is the sample mean, μ. The dotted lines are the upper control limit (UCL) and lower control limit (LCL). The upper control limit can be calculated or set many ways, but most commonly $\mu + 3\sigma$ is used for the upper control limit and $\mu - 3\sigma$ for the lower control limit.

As you can see, setting up a control chart can be very easy and practical, you just have to be familiar with basic statistical calculations and have access to the process data.

To learn more about the specifics of control charts and how to interpret the results, please read *Control Chart Basics*.

Summary

Statistical Process Control (SPC) is a scientific method that can provide much information about a process and how specific process parameters are varying with time. In order to produce a quality product, the amount of variation should be understood and controlled. SPC is a common set of tools used in industry to help manage this variation. Control charts help to identify variation and identify process issues before they begin to affect the product.

All processes have *common cause or inherent* variation. Most processes also experience some special cause variation. In order to maintain a quality product, it is important to be able to identify special cause variation in order to reduce or eliminate it. Statistical Process Control (SPC) can help to identify and monitor all types of variation.

Understanding the basic mathematical concepts involved with SPC is important in developing and utilizing these tools correctly. Some basic statistical concepts that you should be familiar with are *sample median, sample mean, sample range*, *sample variance*, and the square root of the sample variance, the *sample standard deviation*. Using these basic calculations will assist you in creating a SPC chart to monitor process variation.

References

1. Dr. Michael Leeming, University of Arizona
2. Dr. Richard Prairie, Statistical Professor, University of New Mexico
3. Montgomery, D.C. (1985). Introduction to Statistical Quality Control (2nd edition). New York: Wiley.
4. Devore, J.L. (1995). Probability and Statistics for Engineering and the Sciences (4th edition). New York: Duxbury Press.

Disclaimer

Support for this work was provided by the National Science Foundation's Advanced Technological Education (ATE) Program through Grants. For more learning modules related to microtechnology, visit the SCME website (http://scme-nm.org).