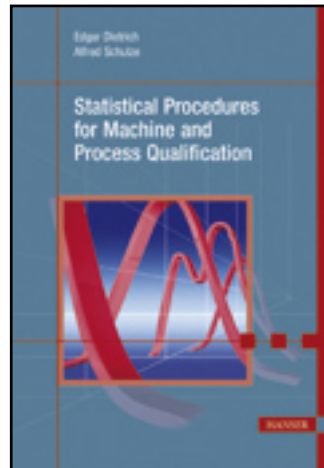Sample Pages

Edgar Dietrich, Alfred Schulze

Statistical Procedures for Machine and Process Qualification

For further information and order see

or contact your bookseller.

4.8   Pareto Analysis

The Italian economist Vilfredo Pareto (1848-1923) discovered more than 100 years ago that in many cases about 20% of the causes are responsible for about 80% of the effects. This is known as Pareto's law or the 80/20 rule. Pareto analyses (sometimes also called ABC analyses or fault frequency analyses) are performed with the aid of special diagrams to separate the "vital few" from the "trivial many" and thus enable the establishment of priorities for quality improvement actions.

Firstly, changes in process parameters and ambient conditions are recorded as events on a check sheet or stored in a computer with suitable input means. For subsequent evaluation, these events must be catalogued and provided with an unambiguous key. In general, it is possible to link a number of events to a single recorded value or sample. Figure 4.8-1 shows a chart of individuals with a number of assigned events.

In the present example the events

- tool wear (event) and
- machine adjustment (measure)

were assigned to value no. 11. The △ symbol in the value chart indicates that events are assigned to this value.

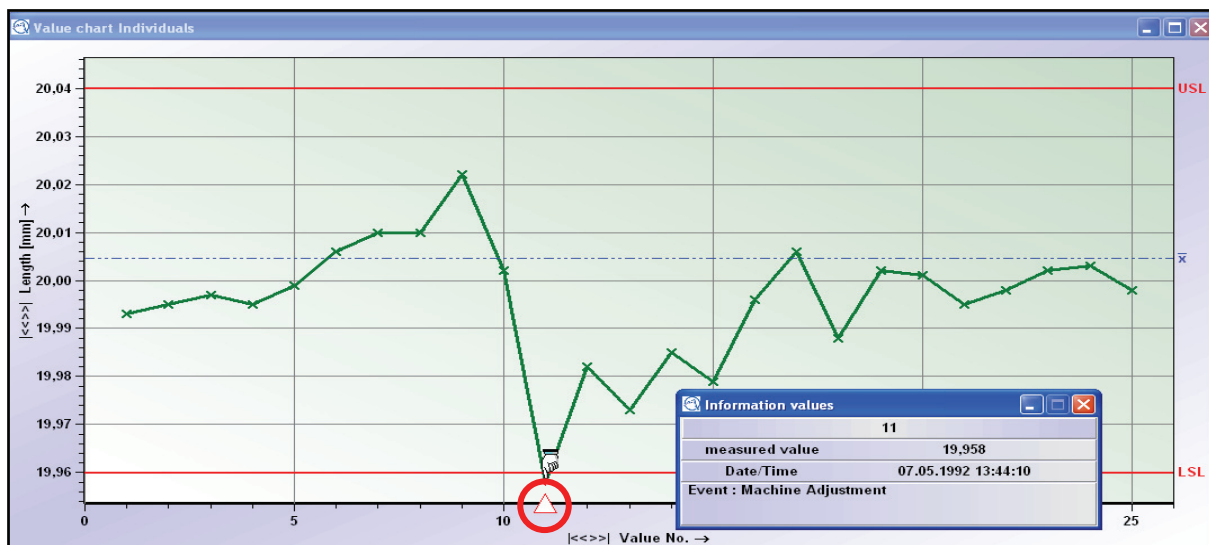*Figure 4.8-1:   Individual chart*

Note

When recording events, it is reasonable to distinguish between the event itself, taken measures and causes of the event right from the start. Thus later evaluations give more evidence for improvements.
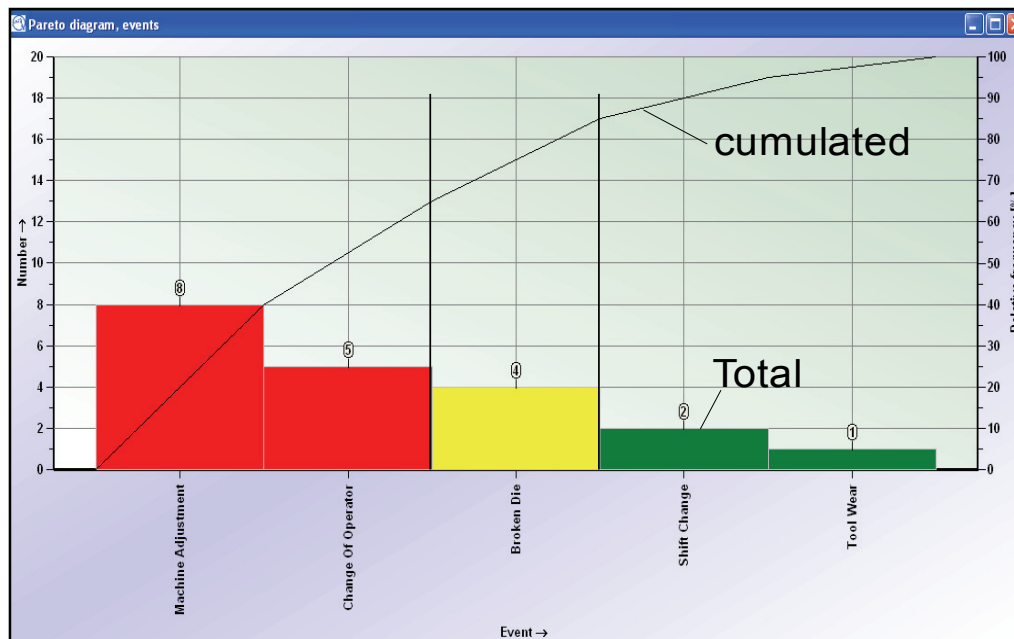
Figure 4.8-2:   Pareto diagram

If, over a given time interval, there are values with events linked to them, the individual frequencies and percentage frequencies of the events are determined and plotted in the form of a bar chart.

As the example shows, the Pareto's analysis rank, orders the events. Plotting cumulative percentages yields a typical Pareto curve (see Figure 4.8-2). The part of the curve with the steepest slope corresponds to region A. This is followed by region B, with a lesser slope. The part where the curve levels out corresponds to region C (see Figure 4.8-3).
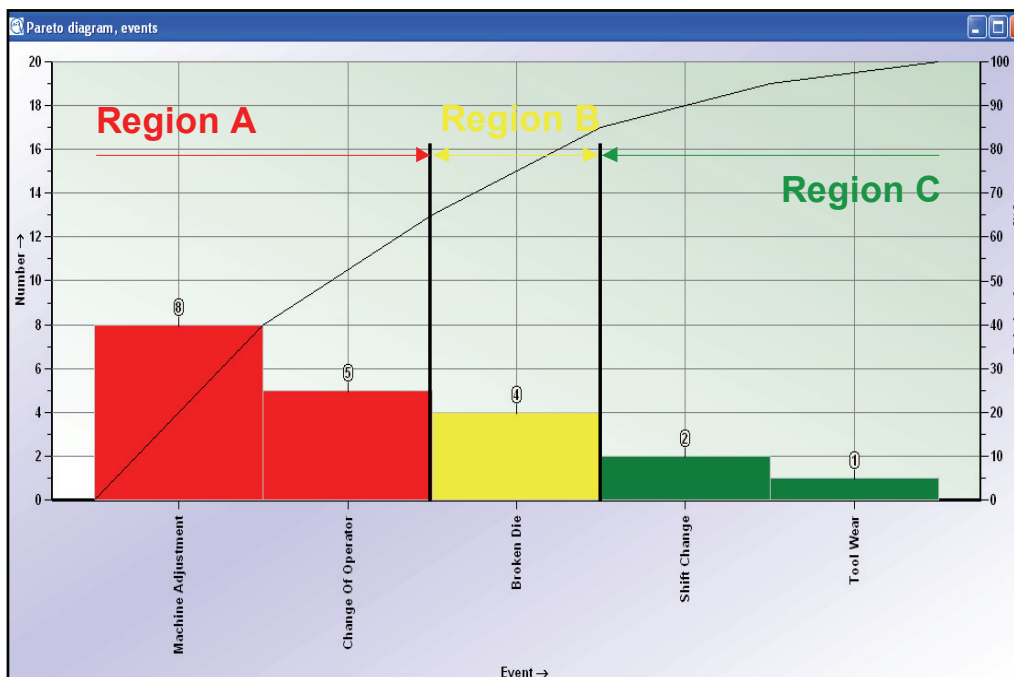

Figure 4.8-3:   Pareto diagam with ABC regions

Region A (red)

Here, two of the categories, namely "machine adjustment" and "change of operator", cause about 65% of the total rework time. Process actions to prevent defects and reduce rework should focus on these two points.

Region B (yellow):

The next category, "broken die," is responsible for a further 20% of the total rework time. The A and B categories together account for 85% of the rework.

Region C (green)

The remaining two categories cause only 15% of the rework time and thus, in comparison, constitute the "trivial many".

Thus the Pareto analysis provides a simple but very versatile method for determining and displaying priority items such as

- faulty components
- frequent process changes
- differing environmental conditions
- etc.

The effects of the important statistics are worthy of investigation and, where appropriate, this should lead to the establishment of rules and operational instructions. Figure 4.8-4 shows other ways of plotting Pareto diagrams.
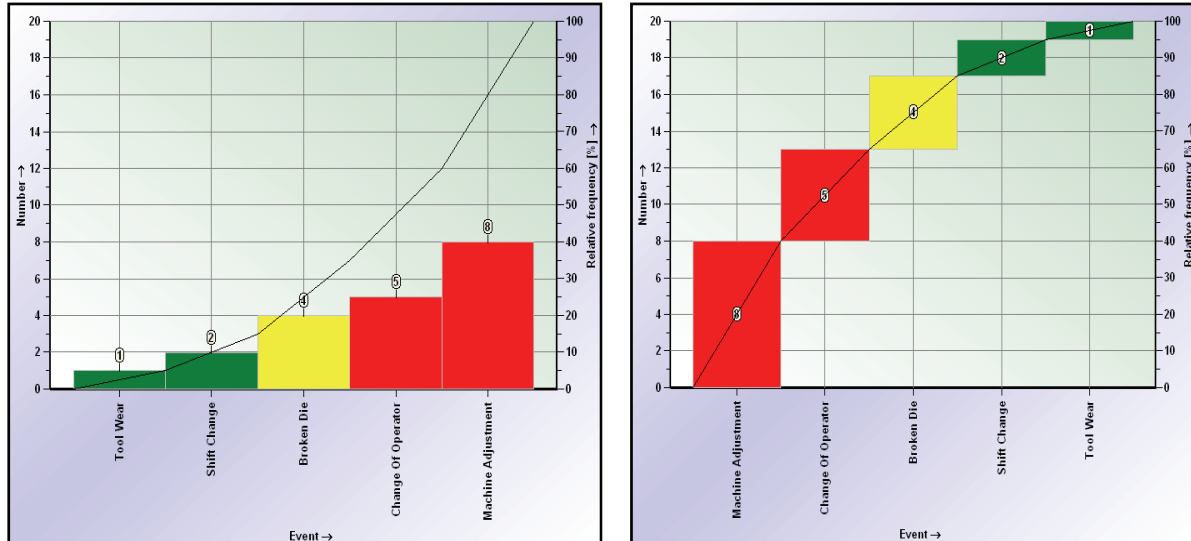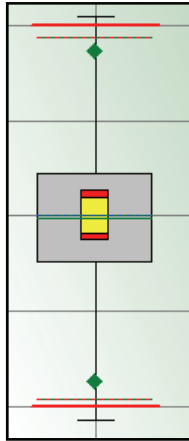

Figure 4.8-4:   Other possible ways to display the Pareto diagram

4.9   Box Plot

A clear representation of the most important statistics is of vital importance for the comparative analysis of product or process data. The box plot meets these requirements.

The height of the box shows the region in which 50% of the values lie on the basis of an assumed distribution model. In addition, the average and/or median may be shown with confidence intervals for 95%, 99% and 99.9%. The width of a box depends on the sample size. The graphic shows whether all characteristics include the same number of values or not.

Maximum or minimum value

The size of the interval indicates the range within which a certain percentage of the values lies. A typical range is the 99.73% range (corresponding to ±3s for a normal distribution). If the distribution is not normal, this range may not be centered on the box.

Table 4.9-1 contains the characteristic numbers, the specification limits and the statistical results as displayed in the box plot. The box plot shows clearly that the numerical values alone are almost meaningless in evaluations, particularly when many characteristics are concerned.

Char.No.	n_{eff}	LSL	x_{min}	$\bar{x}$	$\tilde{x}$	$X_{50\%}$	x_{max}	USL	R	X_{lo3}	X_{up3}
1	500	19,800	19.872	20.00176	20.0000	20.00176	20.124	20,150	0.252	19.87225	20.13126
2	100	19,900	19.855	19.99940	19.9990	19.99940	20.150	20,100	0.295	19.87245	20.12635
3	875	19,850	19.873	20.00058	19.9990	20.00058	20.150	20,200	0.277	19.87570	20.12546
4	500	19,870	19.892	20.00042	19.9990	20.00042	20.110	20,130	0.218	19.87923	20.12161
5	600	19,800	19.869	20.00111	20.0020	20.00111	20.113	20,200	0.244	19.87937	20.12284

Table 4.9-1:   Statistical values for the box plot

The box plot (Figure 4.9-1) presents the facts more clearly. Immediately, we can see that some values of characteristic 2 lie outside specification limits. Since the width of the boxes indicates the sample size, we can recognize that the sample size of characteristic 2 is considerably smaller than the sample sizes of the other characteristics.
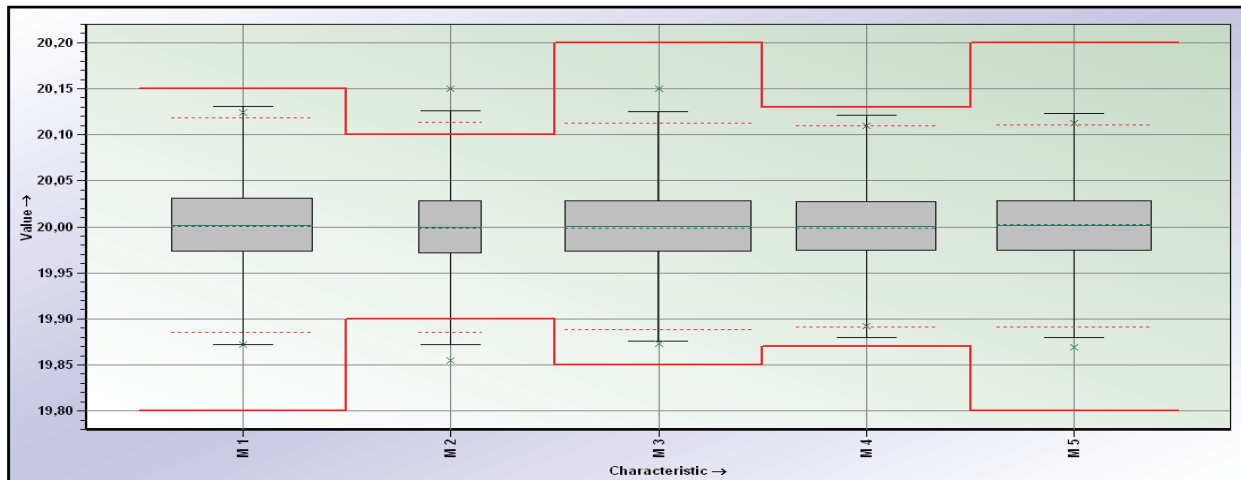
Figure 4.9-1:   Box plot showing the values from Table 4.9-1

In order to gain a representation that is independent of the value range of the characteristics, the display of statistical values should be standardized. The specification limits are determined: USL= 1 and LSL= -1. Consequently, statistical values such as process location and process variation can be compared. If one specification limit is not defined, the display of a characteristic cannot be standardized.
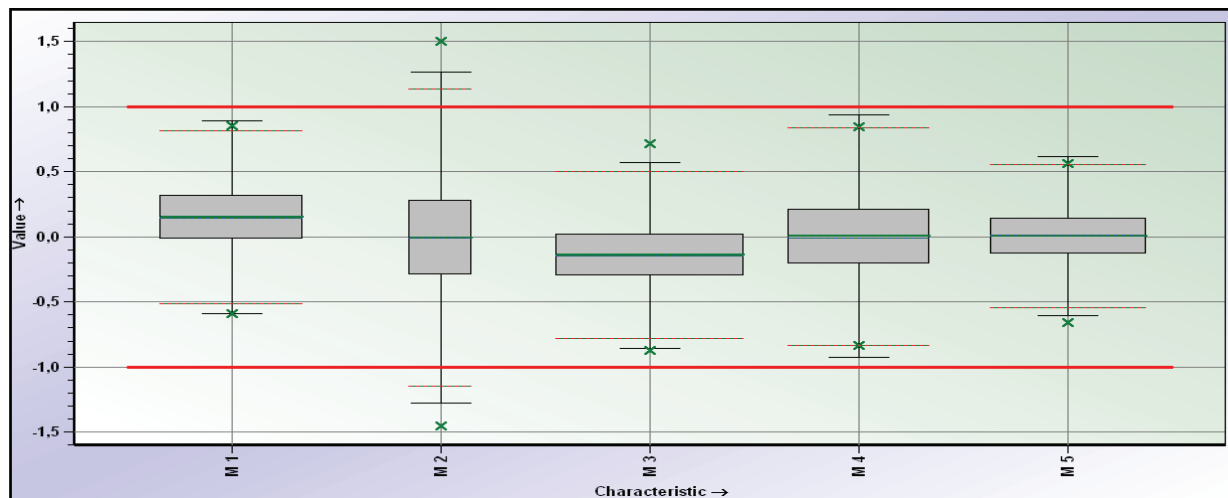

Figure 4.9-2:   Box plot with standardized specification limit

If the value ranges of the characteristics are not too different, it is possible to dispense with standardization (see Figure 4.9-3). In addition, the confidence regions of the averages are plotted. Figure 4.9-4 displays the median and the confidence regions for each characteristic.
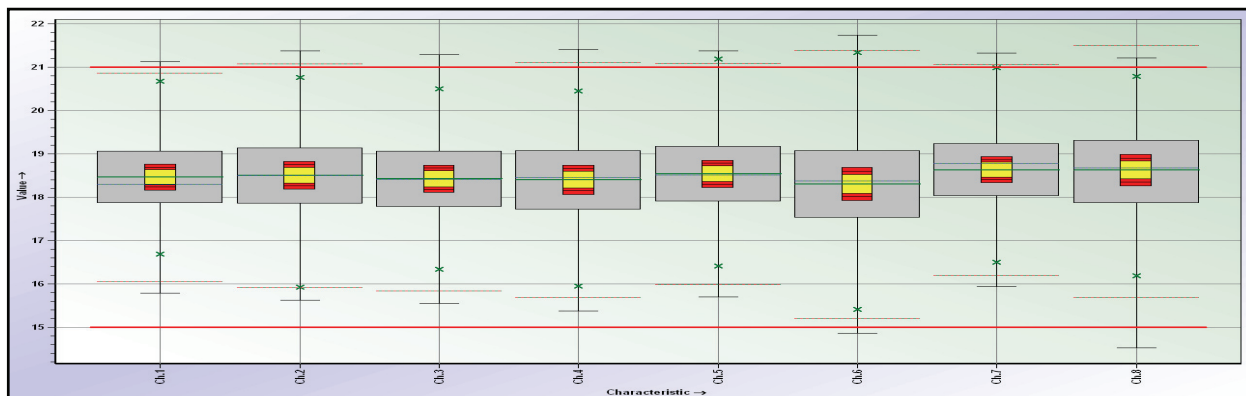
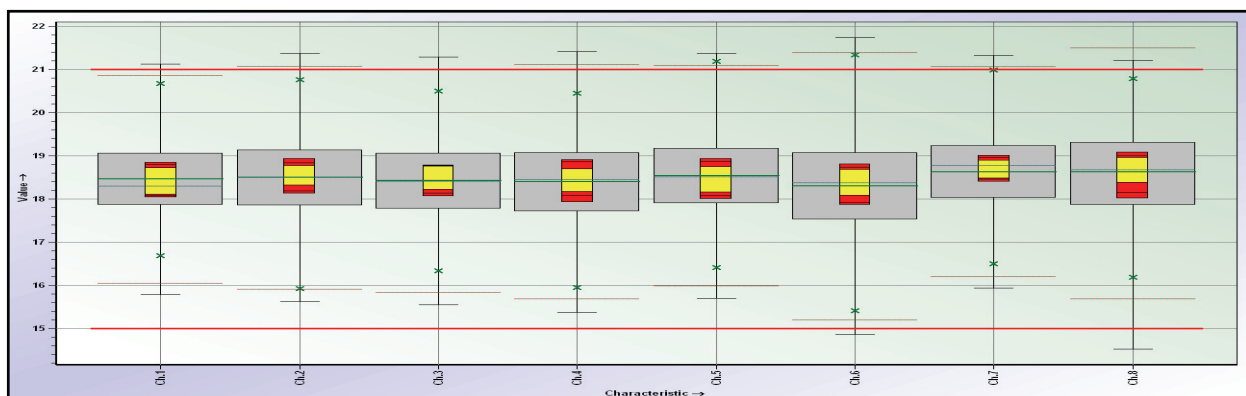*Figure 4.9-3:   Box plot with confidence range for averages*

Figure 4.9-4:   Box plot with confidence range for Median values

Figure 4.9-5 shows a practical example with 39 characteristics. We can see considerable differences between the characteristics. It is important now to observe the characteristics containing values outside specification limits and those with a process variation (99.73% region) exceeding the standard limits (± 1).
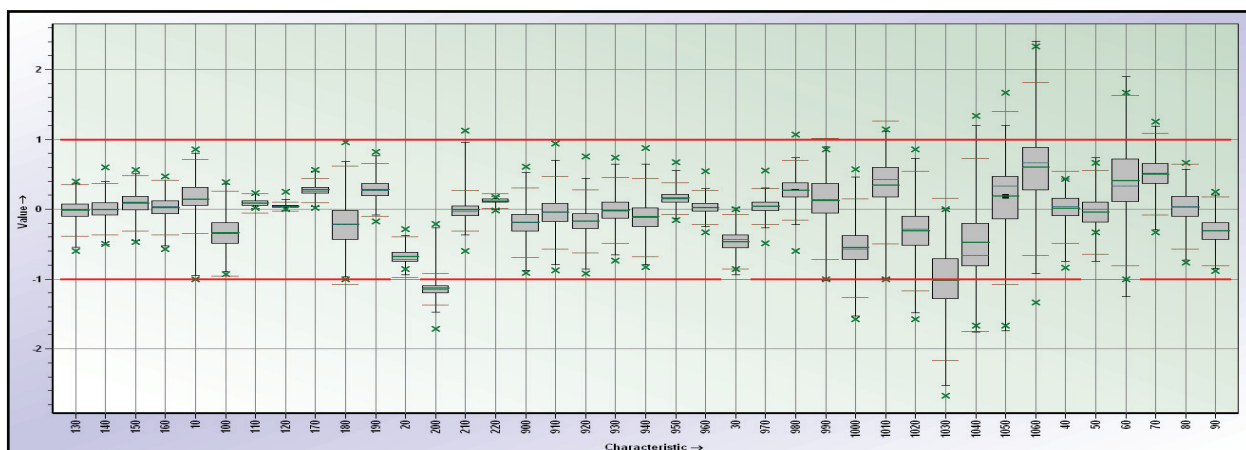

Figure 4.9-5:   Box plot for 39 characteristics

This procedure gives a quick overview and focuses on the essential characteristics, in this case on complicated ones.

4.10 Overview of Capability Indices

An important result of any process or machine assessment is a statement about quality capability as described by quality capability indices such as potential (e.g. C_p) and critical capability indices (e.g. C_{pk}). For comparison of a number of characteristics or of results taken at different time intervals, a bar chart of the capability indices, as displayed in Figure 4.10-1,is recommended. The left bar of each characteristic describes the potential and the right one displays the critical capability index. The name of the statistical values and of the characteristic can be found below the respective bars.
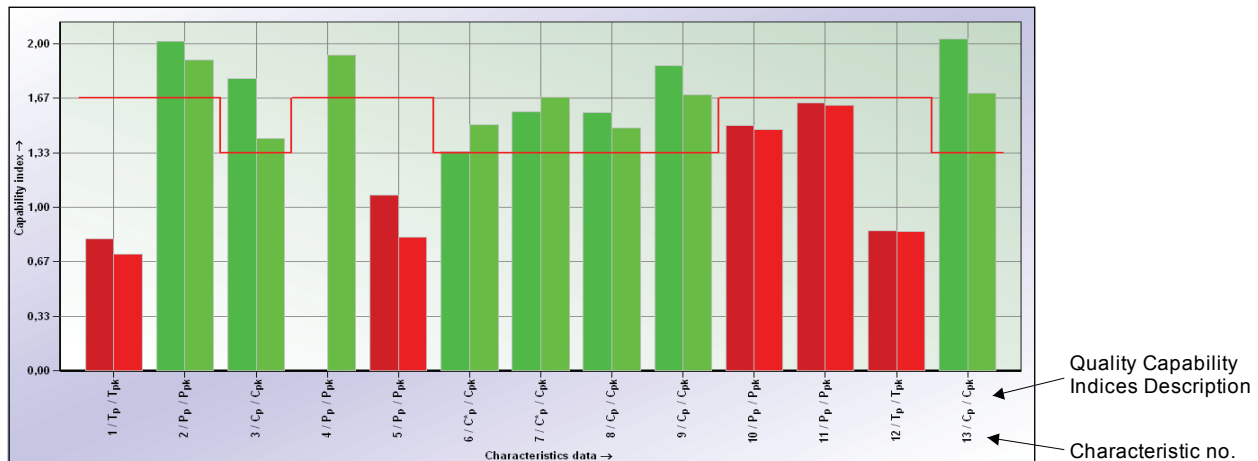

Figure 4.10-1:   Quality capability indices (potential, index separated)

Depending on the characteristics class, the number of evaluated values, and the evaluation result, the limit values may differ. In order to recognize the characteristics that do not meet the requirements, it is recommended to use different colors or to sort the characteristics by statistical value (in ascending or descending order). In particular, the distinction between red, green and yellow values has proved its worth. The yellow region can be defined variably. In order to simplify matters, a range of "±10% of the limit value" is recommended. As an alternative, the range of "limit value ± confidence region" is possible. In this case, the confidence regions for the individual process types must be calculated. However, this is not always feasible.

Notes

- In the case of characteristics with unilateral tolerances, only the critical capability index is shown (a potential index cannot be calculated, due to there being only one specification limit).

- If, however, there is a natural limit for a characteristic with unilateral tolerances (e.g. for zero-limited characteristics the lower specification limit is "zero"), the potential may be calculated (see chapter 9.11). In this case the potential can be smaller than the critical capability index.

In order to distinguish between "good" and "bad" (exceeding limit values) characteristics, the characteristics should be sorted by potential (Figure 4.10-2) or by critical indices (Figure 4.10-3).
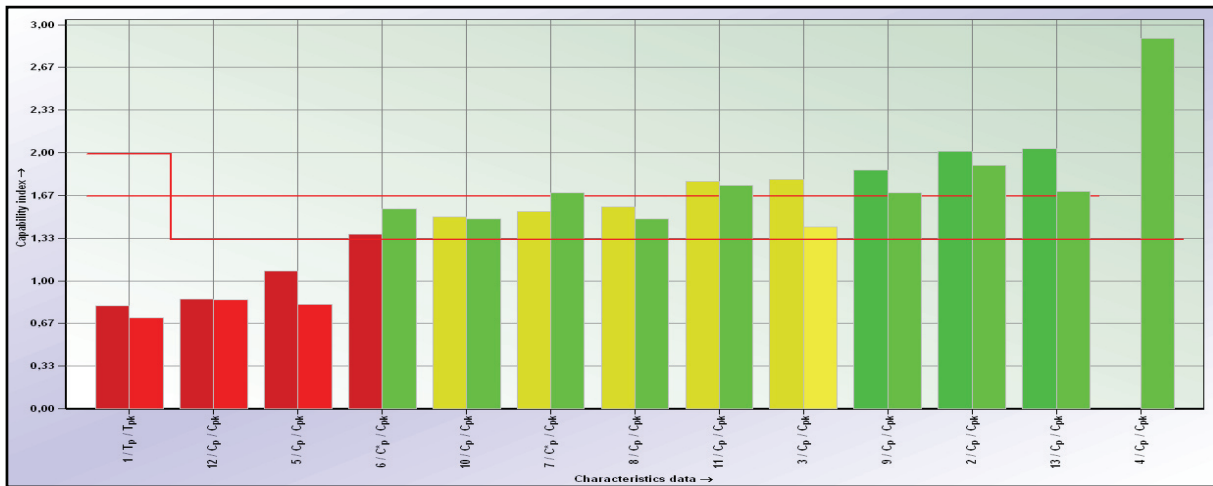
Figure 4.10-2:   Quality capability indices sorted by potential

If the calculated quality capability indices lie below the yellow region, they are displayed in red. However, if they lie above the yellow region, they are displayed in green. Thus it is easy to distinguish between

- complicated characteristics (red)
- uncomplicated characteristics (green) and
- characteristics under observation (yellow).

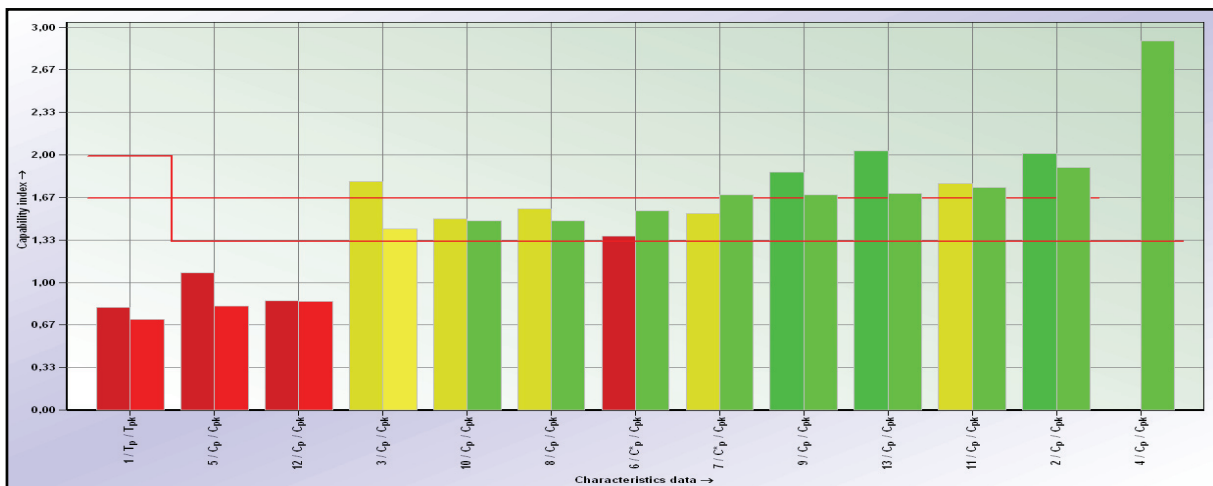The characteristics displayed in red require further analyses.

Figure 4.10-3:   Quality capability indices sorted by critical indices

The quality capability indices can also be displayed in three-dimensional graphics (Figure 4.10-4).
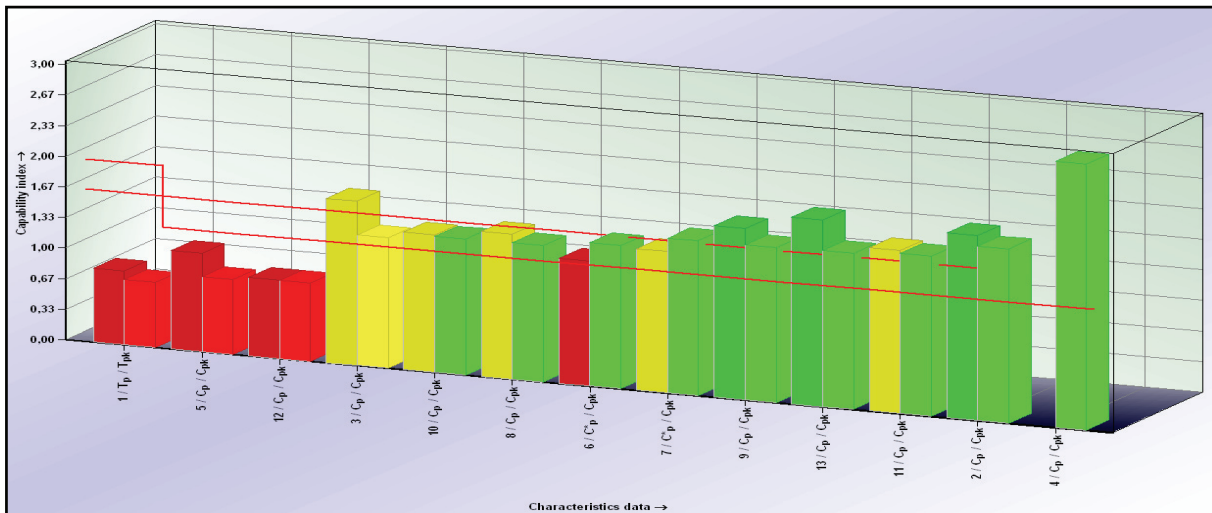
Figure 4.10-4:   Quality capability indices three-dimensional

In case of various characteristics, it is advisable to display the quality capability indices in a row and not next to each other (Figure 4.10-5). Thus the results of the characteristics are shown in space-saving diagrams. The maximum value of a characteristic indicates the potential and the value below the maximum value designates the critical capability index.
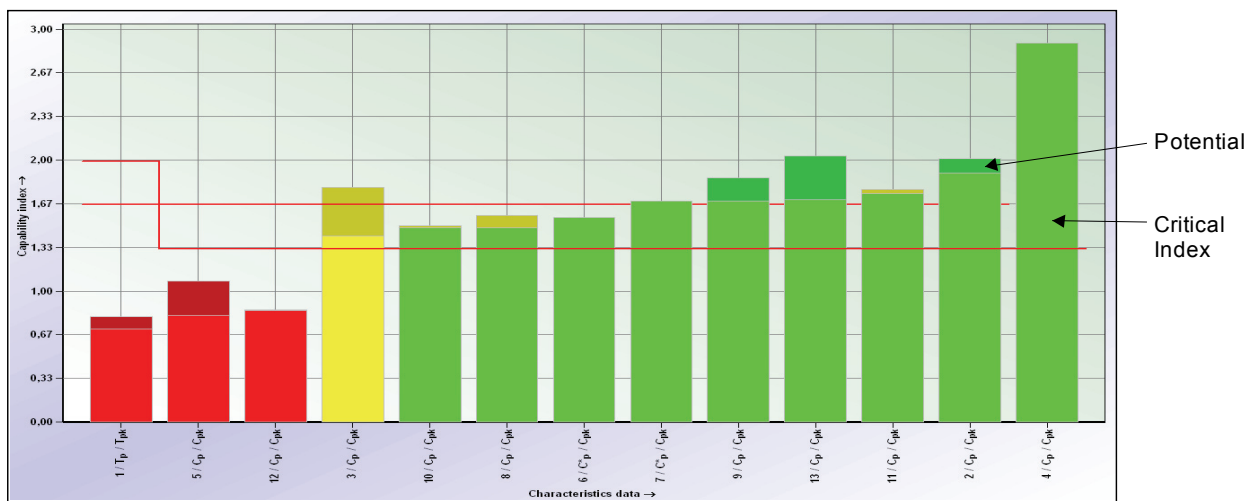

Figure 4.10-5:   Quality capability indices (potential, critical indices stacked)

Figure 4.10-6 illustrates the capability indices of 39 characteristics. This is a practical example which reflects a real process (see also Figure 4.10-7). It is easy to recognize that many characteristics do not meet the requirements. In such a case, it is recommended to sort them by result.
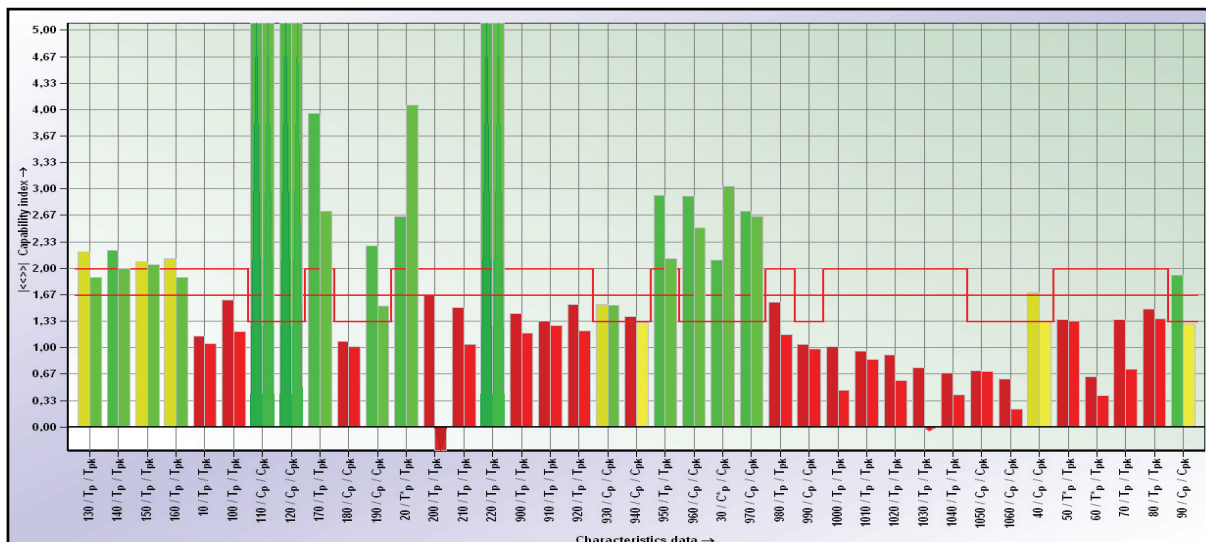
Figure 4.10-6:   *Quality capability indices for 39 characteristics*

Figure 4.10-7 shows the results by means of a histogram. Several classes were formed with the class width ω = 0.33. The quality capability indices are assigned to the respective class.
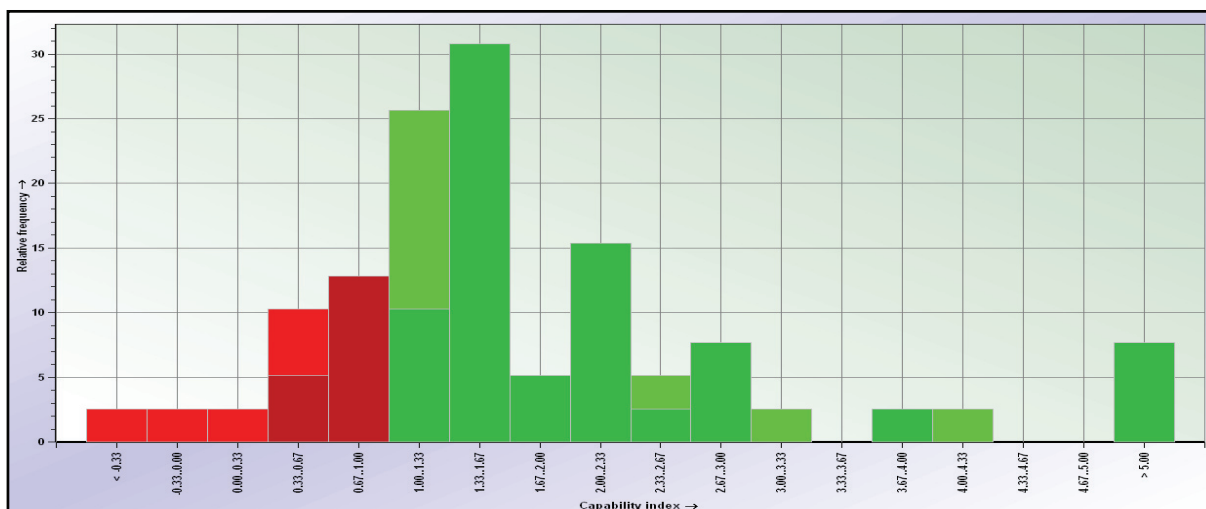

Figure 4.10-7:   *Classified quality capability indices (class width w = 0,33)*

In Figure 4.10-8, the classes are accumulated. This makes it easier to read off percentages such as "How many percent of the classes lie below 1.0?". The upper line relates to the potential and the lower line represents the critical capability indices.
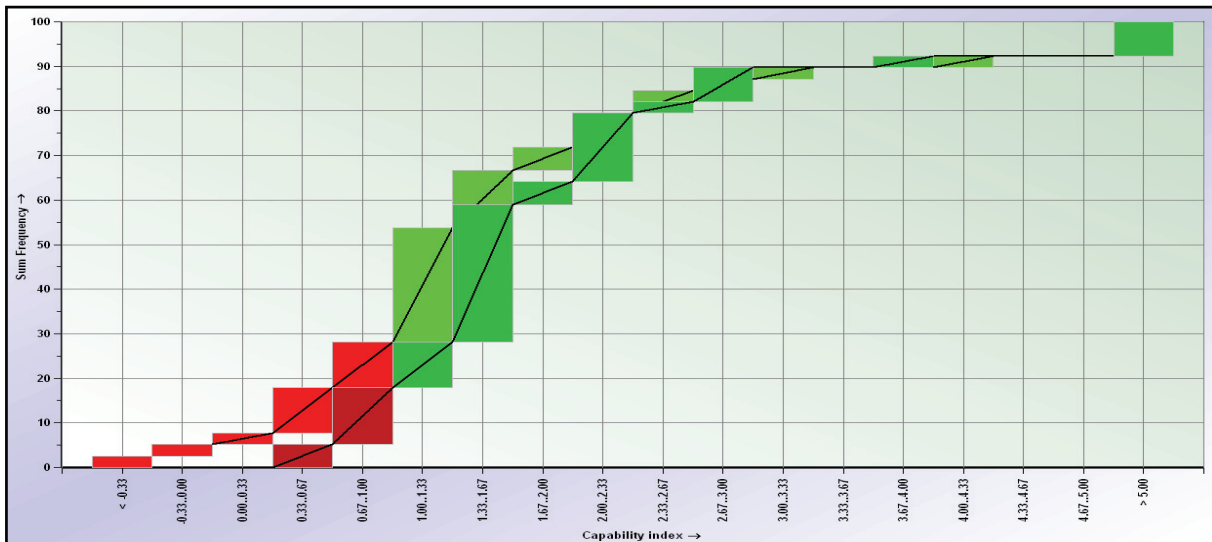
Figure 4.10-8:   Cumulated visualization of quality capability indices

In this case, about 28% of the characteristics do not meet the requirement "greater than 1" regarding the critical capability index. Moreover, about 18% of the potentials do not reach the defined limit value.

The preceding graphics only observe one part in detail and display the quality capability indices of the individual characteristics by means of a bar. Part-spanning graphics are also possible (Figure 4.10-9). In order to identify the parts, we need to enter the respective part number and the characteristic number.
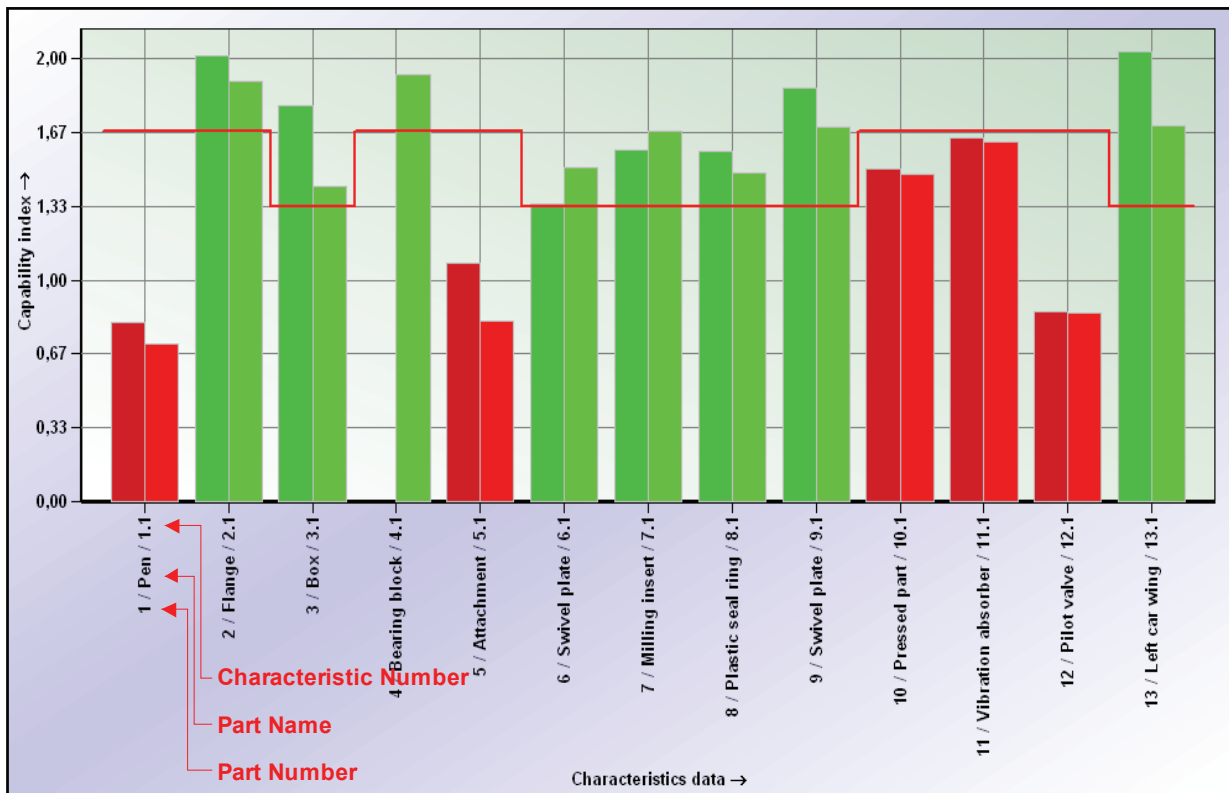

Figure 4.10-9:   Capability indices from several parts

If we want to compare more parts with many characteristics, the results have to be accumulated. Figure 4.10-10 shows the quality capability indices of several parts classified by means of a benchmark. These bars are displayed above the respective part.

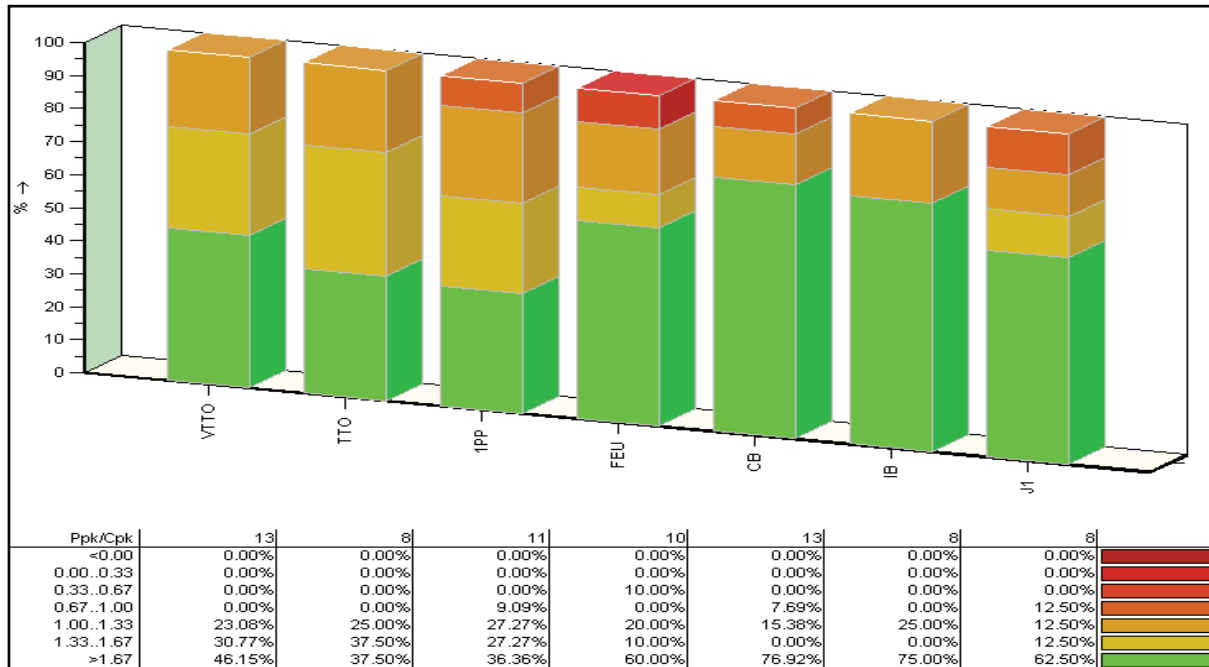| Ppk/Cpk	13	8	11	10	13	8	8	
<0.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
0.00..0.33	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
0.33..0.67	0.00%	0.00%	0.00%	10.00%	0.00%	0.00%	0.00%	
0.67..1.00	0.00%	0.00%	9.09%	0.00%	7.69%	0.00%	12.50%	
1.00..1.33	23.08%	25.00%	27.27%	20.00%	15.38%	25.00%	12.50%	
1.33..1.67	30.77%	37.50%	27.27%	10.00%	0.00%	0.00%	12.50%	
>1.67	46.15%	37.50%	36.36%	60.00%	76.92%	75.00%	62.50%	

Figure 4.10-10:   Benchmark

This kind of representation can also be used to compare the values of one part in different periods. The values of the different periods are displayed next to another. It might be helpful when comparing the quality capability of one part. We can also observe how a part changes over time.
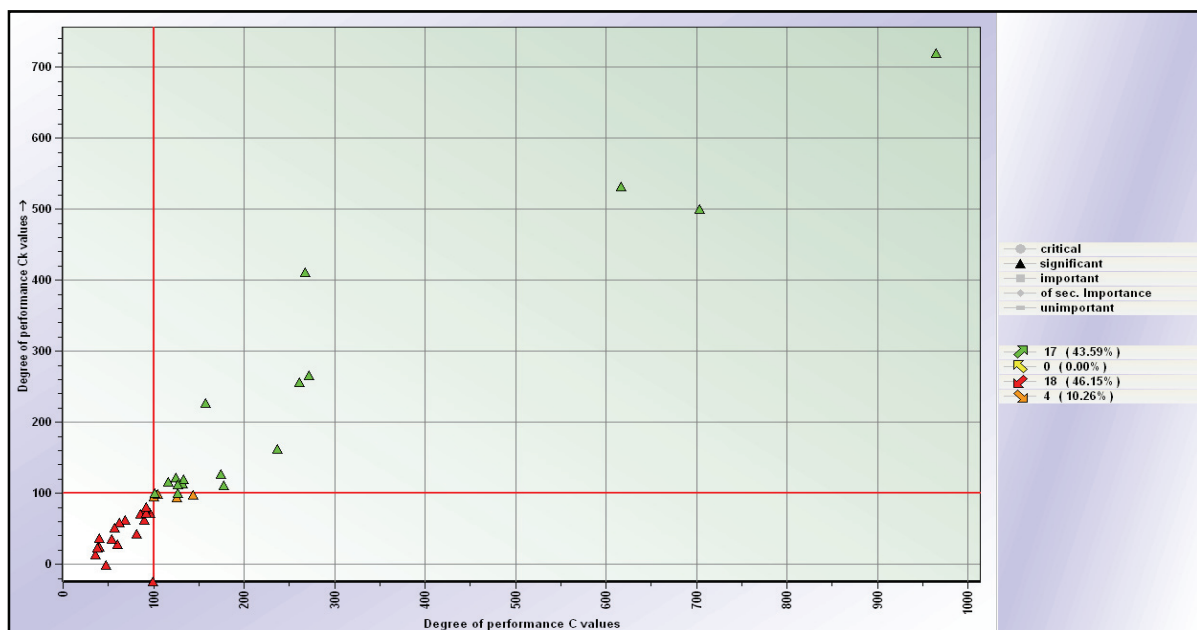

Figure 4.10-11:   Portfolio display

According to their degree of performance, the c values can also be displayed by means of a portfolio (see Figure 4.10-11).

Instead of c values, we may also plot the standard deviation of every characteristic against the corresponding average (see Figure 4.10-12). In order to recognize the capable characteristics immediately, borderlines are drawn according to the respective requirements (e.g. C_p=1.33).
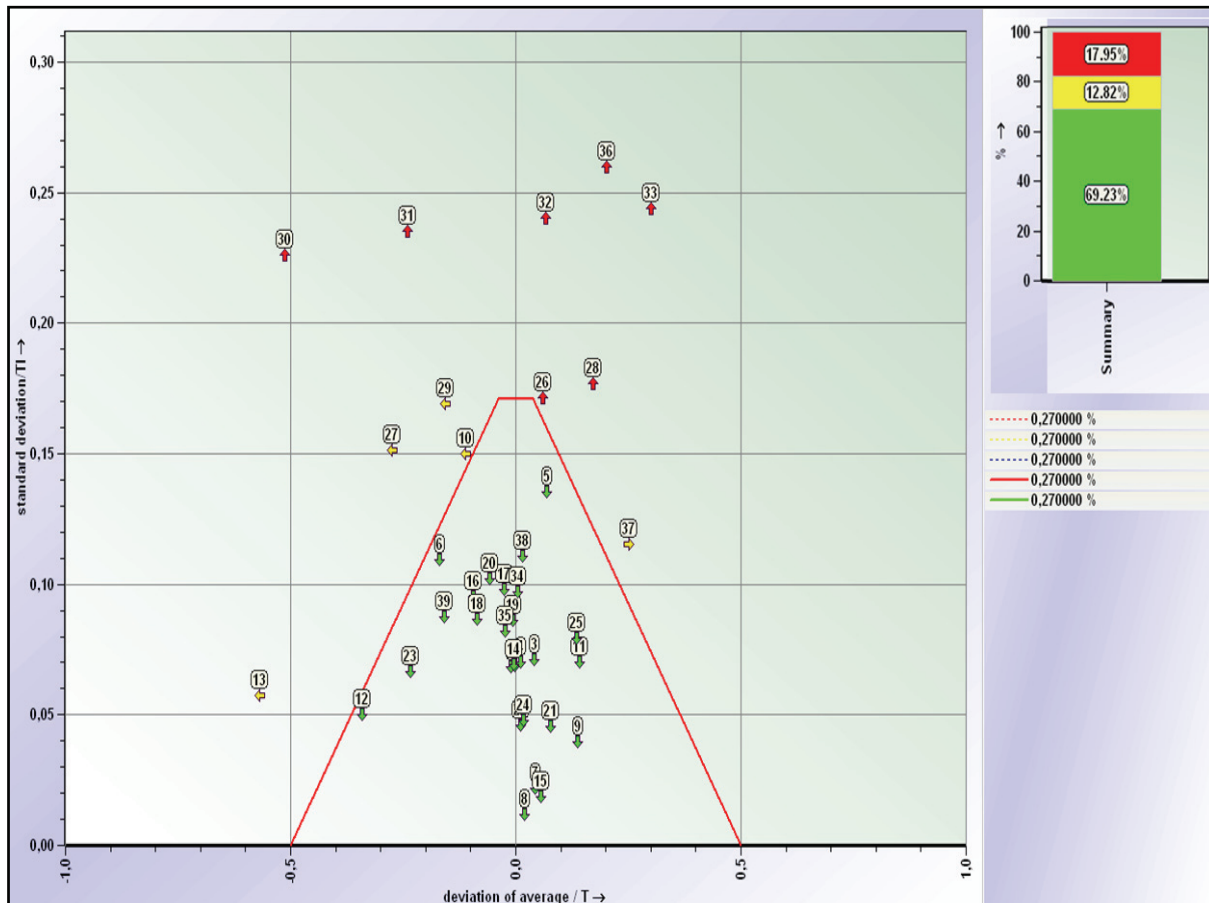

Figure 4.10-12:   s over $\overline{x}$ display

4.11 Graphical and Numerical Display

The advantages of the graphical visualization of statistical values have been demonstrated several times. However, the graphics were often supplemented by numerical results.

The results of one part for individual characteristics can be displayed by means of a parts protocol (Figure 3.3-3) We may decide how extensive this protocol shall be. In the first instance, the location of individuals is to be examined due to the parts protocol.