

Notes:



**Evaluations with
qs-STAT**

Notes:

Contents

- Why do we do statistics?
- Systematic sequence of evaluations
- What's a population and a sample?
- Usage of the complete tolerance vs. target value orientation
- Means and deviations
- A picture tells more than a thousand words
- Usage of graphics for evaluations
- Capabilities revisited

Why do we do statistics?

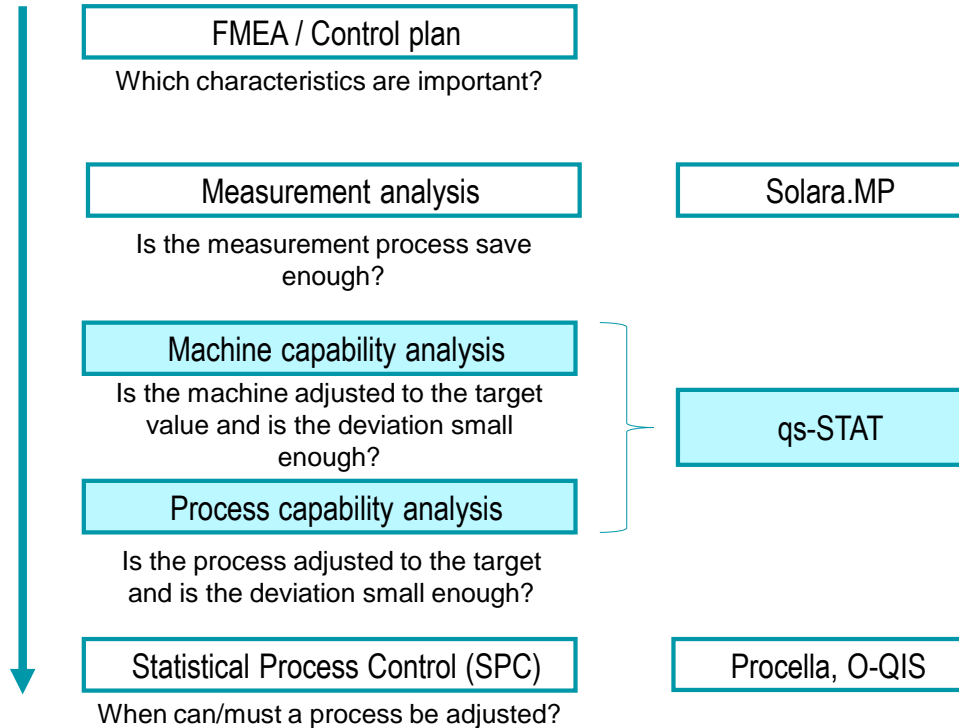
- To estimate the quality of a process, machine or measurement process
 - To see a process where we do not measure everything more clearly
 - To ensure quality
 - Avoid rejects and eliminate waste
- Reduce costs
 - By measuring less frequently
 - By measuring less parts
- Create documentations for process improvement and for the case of problems
 - Create a knowledge base / provide a prove of quality
 - To get more transparent processes
 - Compare similar processes
 - Prove an assumption
- ...



Notes:

Notes:

Systematic sequence (example)



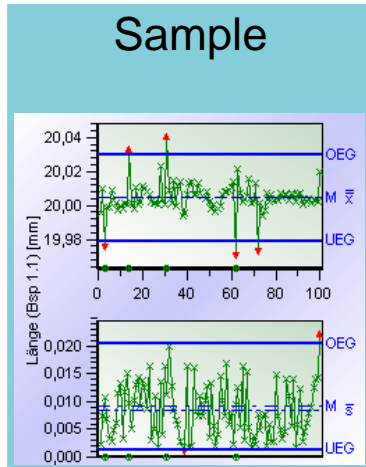
Notes:

Standard approach (not complete)

Methode	FMEA	MSA	MCA	PC	SPC
Procedure:	Team of experts decides	Method 1-3 Linearity Stability...	50 (25) parts are measured in production sequence	Lots of parts (i.e. 1000) are produced and i.e. 25 x 5 are measured as samples	After the control limits of a QCC are exceeded the process is directly adjusted
Conditions:	Team must consist out of members of all relevant divisions	<ul style="list-style-type: none"> Measurement systems produce data, but no parts 1/20 of the tolerance is the base resolution Them measurement must take place where it normally is done(!) The operators responsible for the measurements must perform the analysis 	<ul style="list-style-type: none"> Observation of production machines which produce parts, not measurements One operator One machine One Cavity No tool changes One batch No adjustments ... 	<ul style="list-style-type: none"> Each part should have the same chance to get sampled. The production runs on serial production conditions: <ul style="list-style-type: none"> More operators More batches Different machine settings With tool changes 	<ul style="list-style-type: none"> Samples need to be taken regularly The measurements must be truthfully reported The measurement must be directly visible The process needs to immediately be changed
Why are these conditions necessary?	To recognize problems before the process is started	<ul style="list-style-type: none"> Production can only be as good as their measurement processes are. If the measurement process is not working correctly the customer could get scrap or could reject the parts To provide a prove for a capable measuring method 	<ul style="list-style-type: none"> To find out if the location of the machine (mean) and the variance (deviation) of the machine are good enough for the requirements To reduce process influences to only get the natural deviation of the machine To provide a prove for a capable machine 	<ul style="list-style-type: none"> To find out if the process influences even under serial conditions are small enough to still ensure a good quality To provide a prove for a capable process 	<ul style="list-style-type: none"> Because the process changed The change could produce waste / rejectable parts Because the risk for following processes increases ... Otherwise it shouldn't be called SPC
Conclusions / Results	<ul style="list-style-type: none"> 100 Merkmale wichtig 10 Merkmale kritisch 	<ul style="list-style-type: none"> Cg/Cgk %GRR Stable QCC Good linearity capable Limited capable Not capable ... 	<ul style="list-style-type: none"> Cm/Cmk Newer standards use Pm/Pmk capable Limited capable Not capable 	<ul style="list-style-type: none"> Cp/Cpk or Pp/Ppk In seldom cases Tp/Tpk capable Limited capable Not capable 	<ul style="list-style-type: none"> Working processes Timely determination of process changes Reject costs Lower hidden quality costs

Confidence intervall

Notes:



indirect conclusion
Confidence intervall

Based on the samples the real Cpk is between 1.63 and 2.17 with 99% probability

Population	
$\bar{x} =$	$14,067677 \leq 14,067923 \leq 14,068169$
C_p	$1,73 \leq 2,01 \leq 2,29$
C_{pk}	$1,63 \leq 1,90 \leq 2,17$

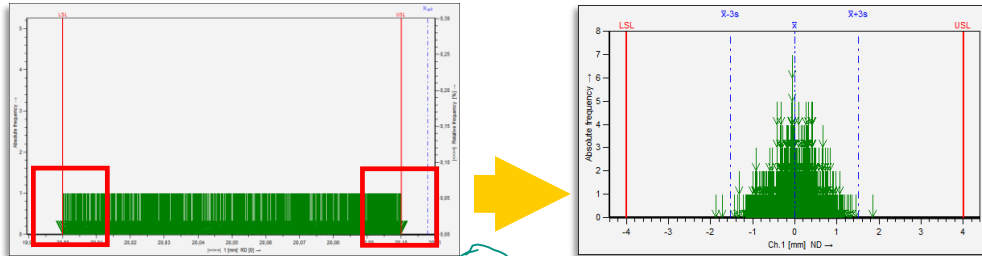
Direct conclusion
Confidence intervall

The samples will be with 99% probability within the upper and lower control limits if the process doesn't change

Notes:

Usage of the complete tolerance vs. target value orientation

„The tolerance belongs to production!“ „We want security and therefore to reach the ideal value as good as possible.“
 „...it was always like that!“



Troubleshooting costs
 Prevention costs
 Inspection costs

**Hidden costs
 because of bad
 quality**

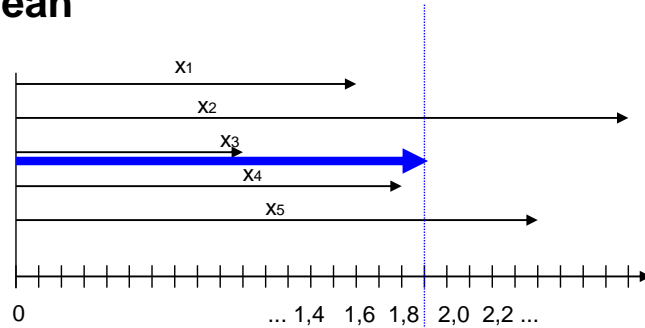
Error prevention has priority
 to troubleshooting / fault
 elimination!

Notes:

Some statistical numbers...

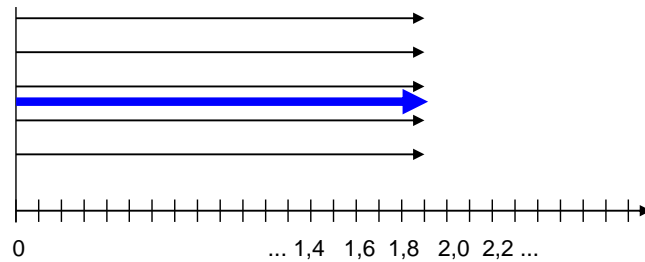
- Parameters for Mean
 - Arithmetical mean
 - Median (50% value)
 - Modalwert (Wert der höchsten Häufigkeit)
- Parameters of variances
 - Range
 - Standard deviation
 - Quantile limits / Percentiles
- Capability indices
 - C_p / C_{pk}
 - C_m / C_{mk}
 - ...

Arithmetic Mean



Calculation:

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

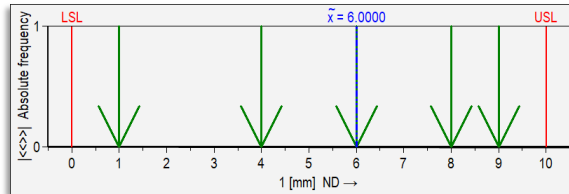


Notes:

Notes:

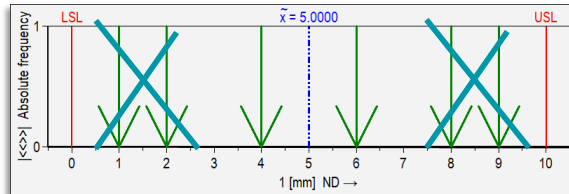
Median (50%-Value)

- Median for an odd # of values:
 - The values (4, 1, 9, 8 and 6) are sorted by value (1, 4, 6, 8 and 9).
 - The value $(n+1)/2$ is the median. Here: 5 values + 1 = 6; $6/2 = 3$; 3rd value is the median.
 - The amount of numbers left and right of the median are the same



$$\tilde{x} = x_{(n+1)/2}$$

- Median for an even # of values
 - The highest and lowest numbers are cancelled until 2 values remain. These 2 will be added and divided by 2. ($4+6=10$; $10/2=5$)

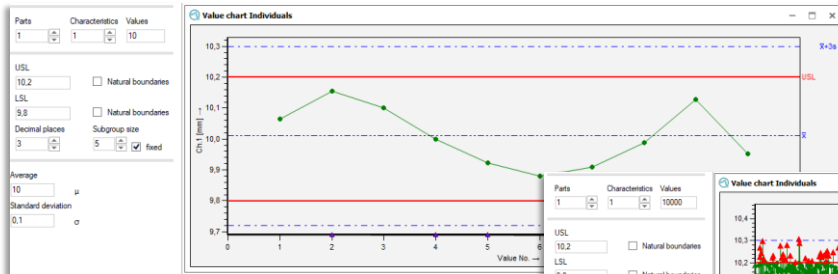


$$\tilde{x} = \frac{x_{n/2} + x_{(n/2+1)}}{2}$$

Range

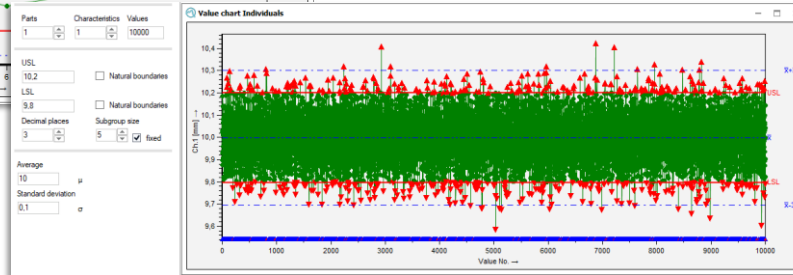
- Highest value – lowest value: $R = X_{max} - X_{min}$
- The Range is easily calculated, but sensitive to outliers
- The Range is often used in quality control charts
- Range can easily be misinterpreted
 - Fewer values: $R <$ than the real deviation of the population
 - Many values: $R >$ than the real deviation of the population
 - 2 Examples with: mean = 10, sigma = 0,1:

Notes:

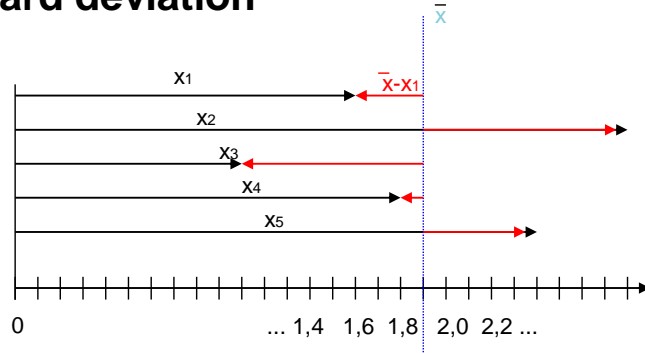


10 values

10000 values



The standard deviation



Standard deviation:

$$S = \sqrt{\frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n - 1}}$$

Variance

$$S^2 = \frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n - 1}$$

Notes:

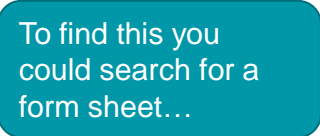
Exercise:

- Open the module sample analysis and load Test_02.dfq
- Find the arithmetic mean
- Find the standard deviation
- Find the Cm/Cmk bzw. Pm/Pmk values
- Find the Range



Simple, isn't it?

- Find the number of values of the analysis
- Open the graphics value chart, value plot, histogram, cumulative line and probability plot.



To find this you could search for a form sheet...

- Time: 5 minutes

Notes:

Notes:

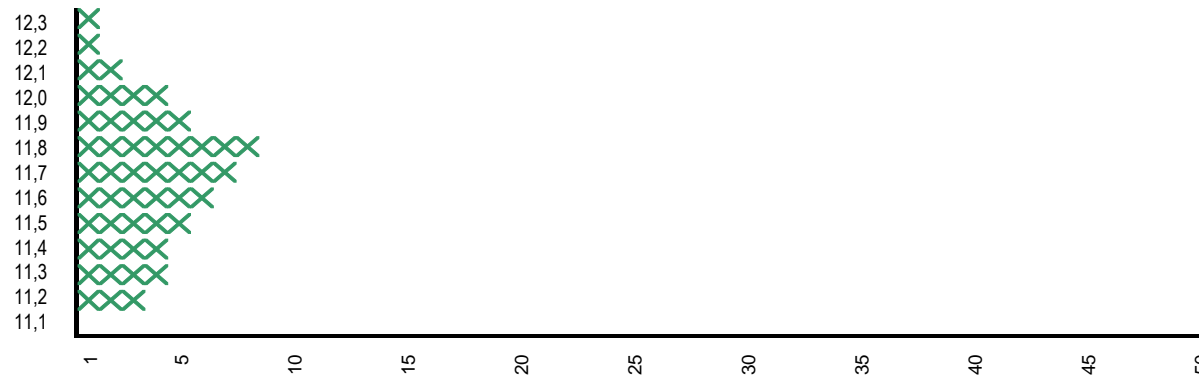
Origins of graphics

- In this chapter we try to understand how the graphics evolve and clarify some functions within the software
 - Value chart individuals / Value plot individuals
 - Value plot individuals / Histogram
 - Histogram / Cumulative line
 - Cumulative line / Probability plot

Notes:

Value chart individuals -> Value plot individuals

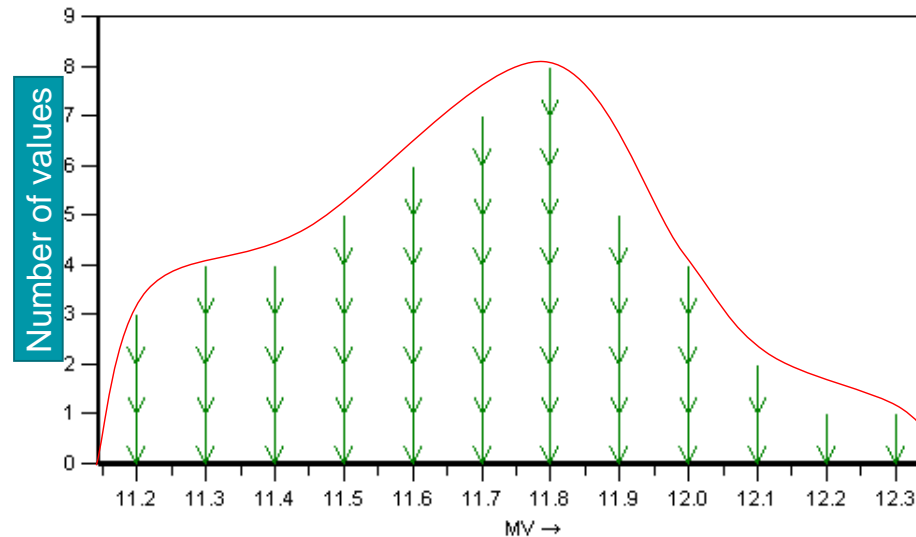
- Which value is measured when is not visible anymore
- The frequency of each measurement will be visible
- If a value is measured more often or less than normal will be visible
- After moving all values to the y-axis the plot is turned by 90 degrees



Notes:

Value chart individuals -> Value plot individuals

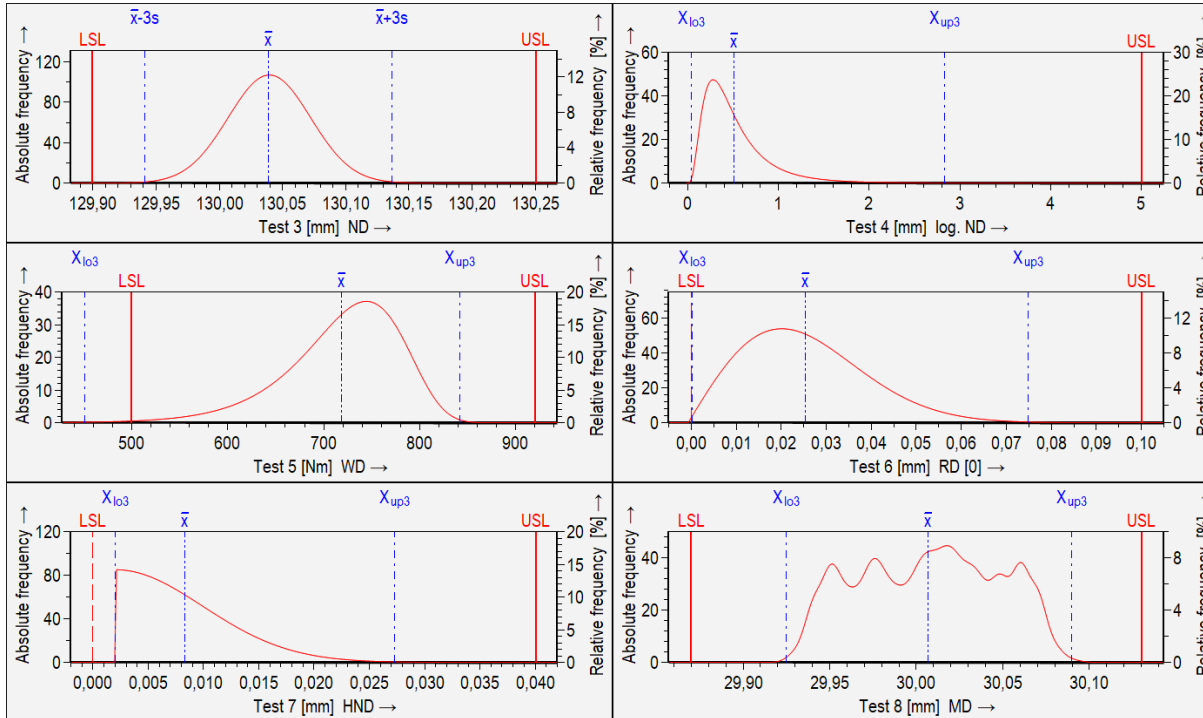
- The more often a value is measured, the more likely it is in the population
- The frequency of occurrence of values should not be interpreted as the likelihood of the distribution



Notes:

Value plot is not about distribution models

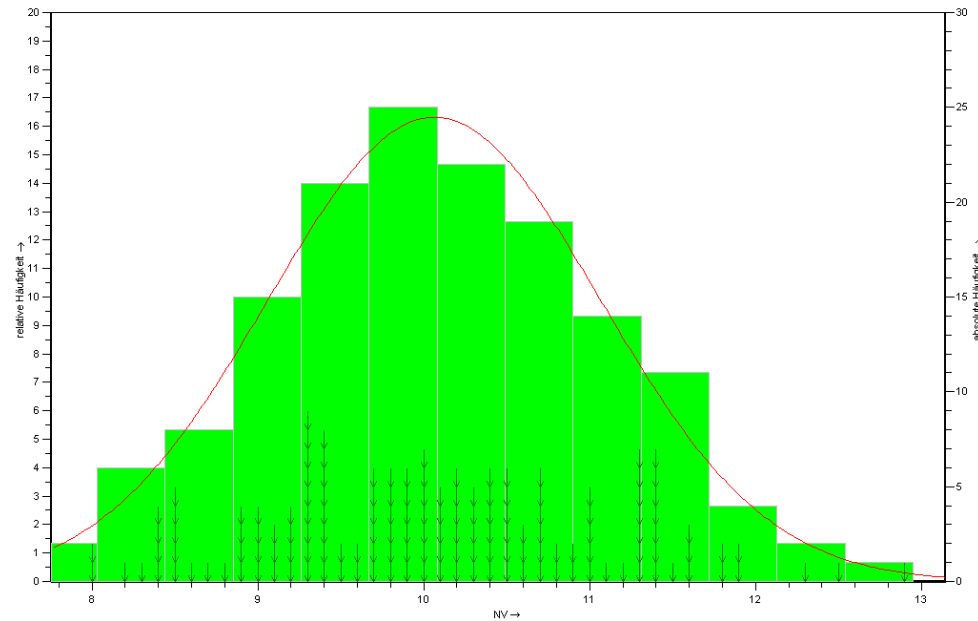
- Here are some distribution models which could be found:



Notes:

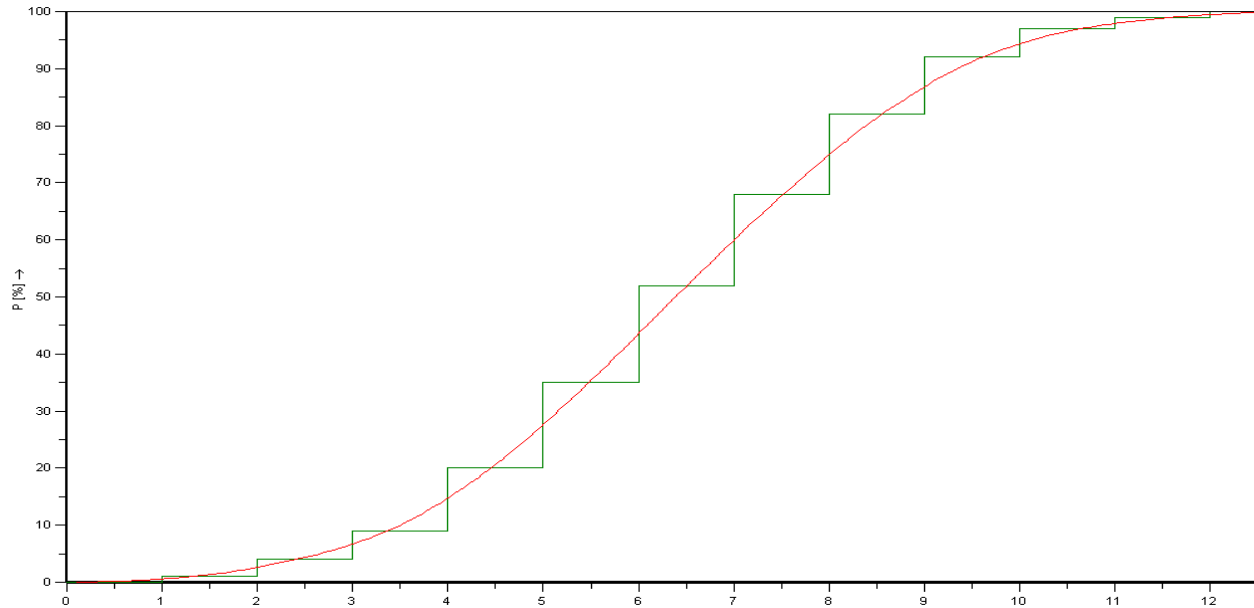
Value plot individuals -> Histogram

- First you will only see values: which distribution model could be assumed?
- The absolute and relative frequencies in the histogram lead to different model assumptions
- Because of that an evaluation strategy chooses for you the best model for the data



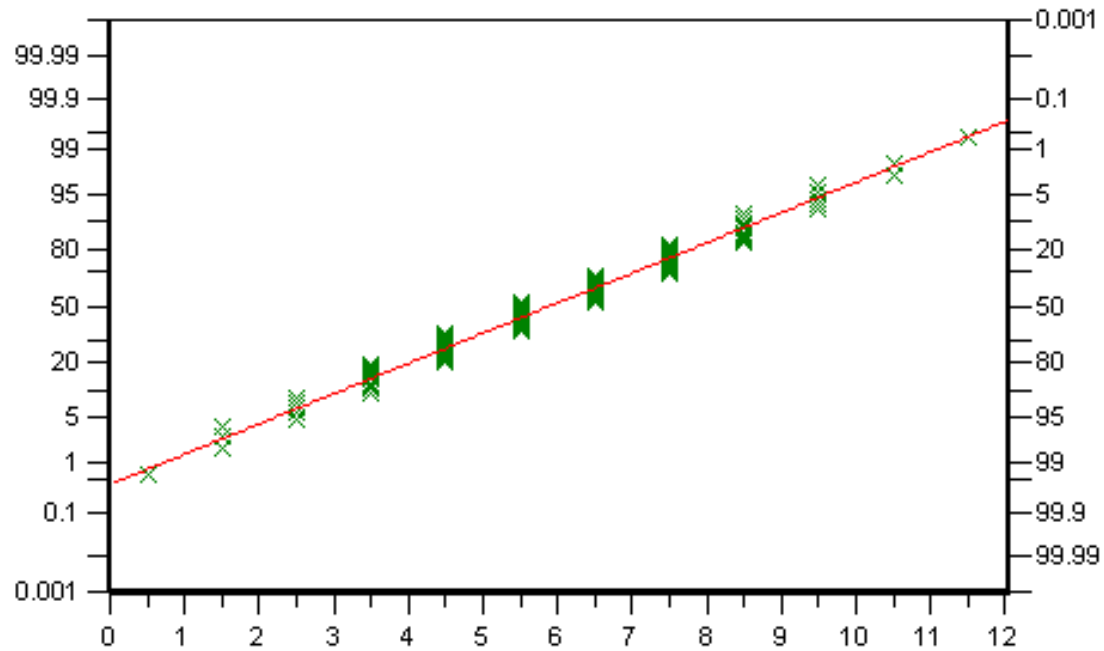
Histogram => Cumulative Line

Notes:



Cumulative Line -> Probability plot individuals

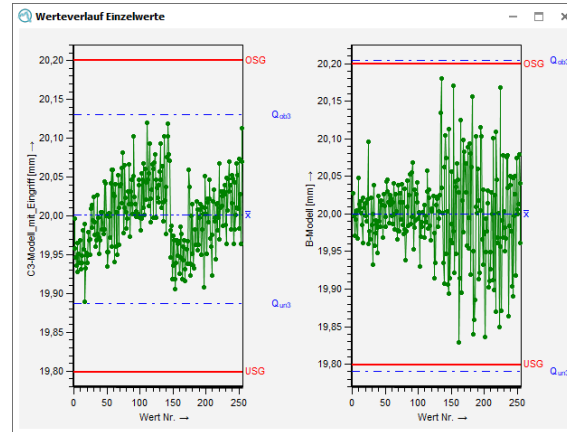
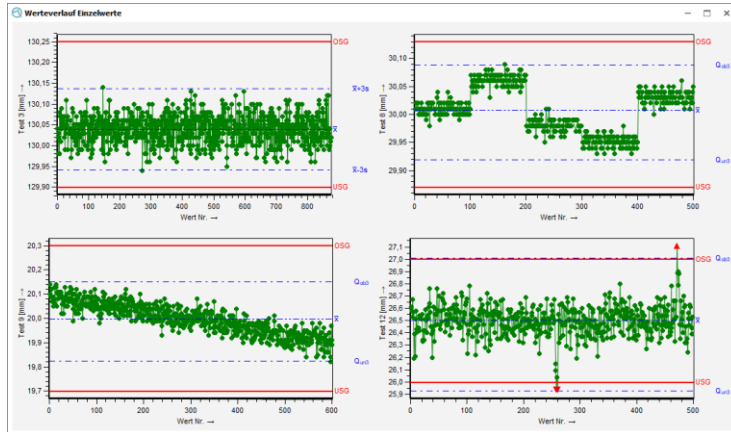
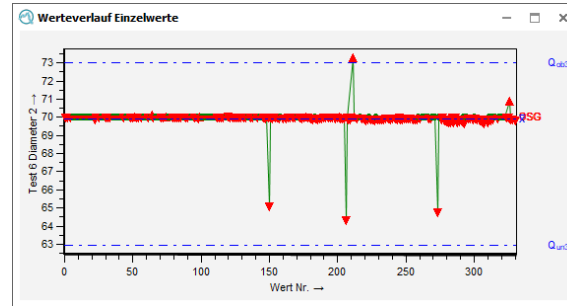
Notes:



Value Chart individuals: Time Sequence of single values

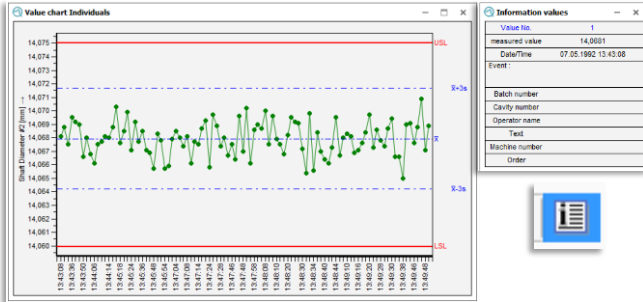
- Values outside specification and possible outliers
- Position of values
- Time sequence
- Trends
- Change of means (constantly, suddenly)
- Changes of deviation

Notes:

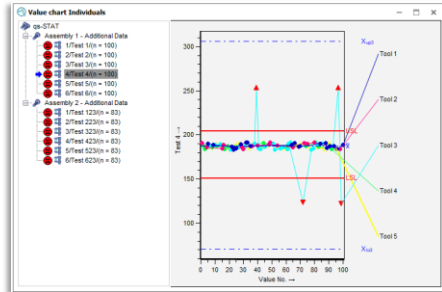


Different displays of additional data

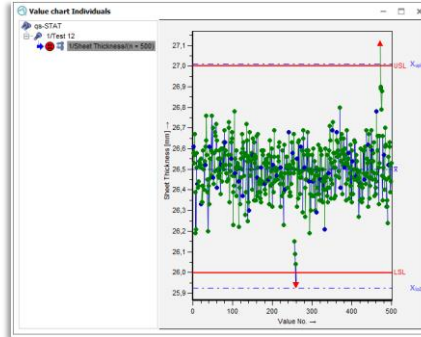
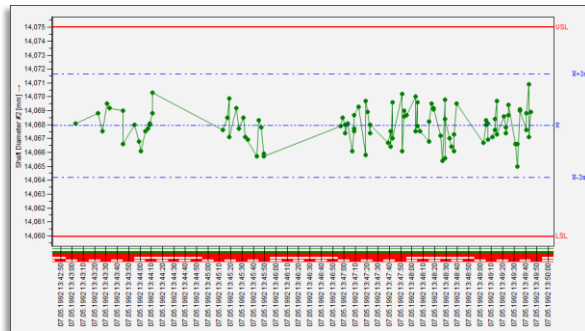
- Graphical Settings for the value chart



Time order normal
Time order relative



Divided by cavity
Divided by sample




Notes:

Additional data axis



Visualization of additional data

- Open the file Test_01.DFQ
- Activate | Graphical settings | value chart the “info” option 
- Divide the values by activating | Graphical settings | Value Chart „allocation for additional data“ by “Batch number”

Notes:



Allocation for additional data

Activate

Activate

Allocation for

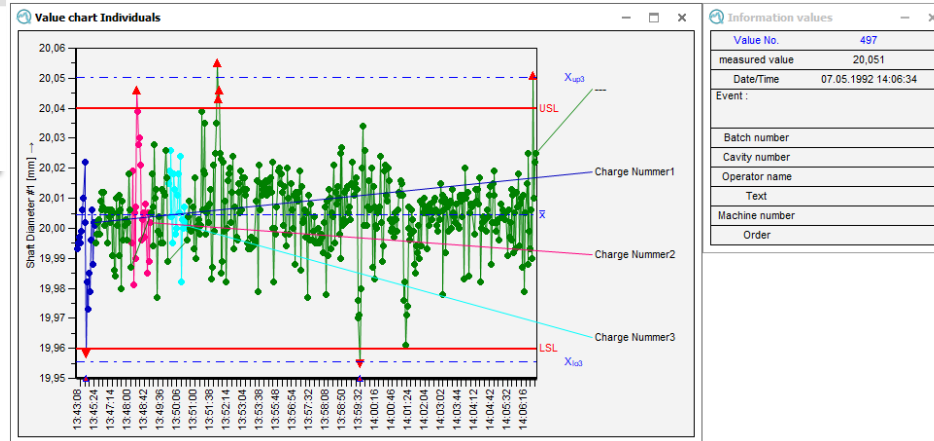
Machines

Cavities/spindles/clamping locations

Gages

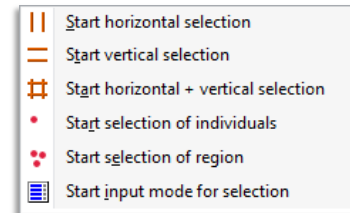
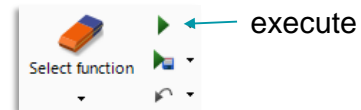
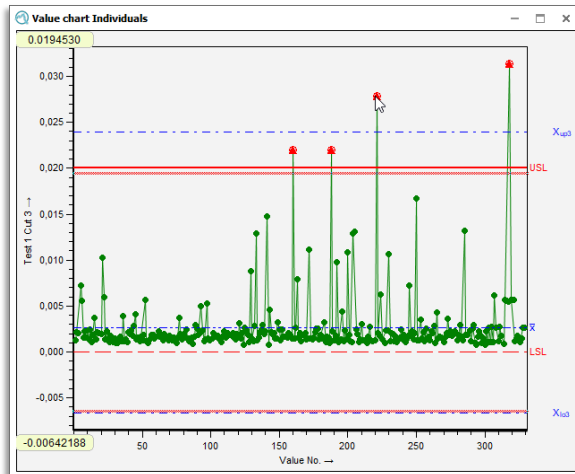
Inspectors

Batch number



Exercise: Data selection

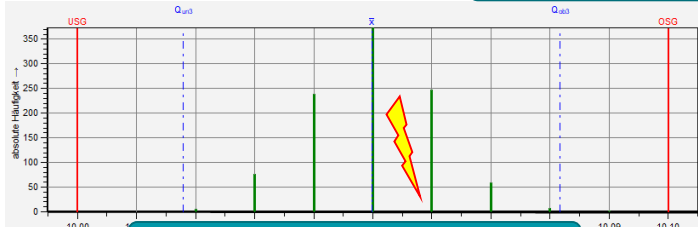
- If the data quality is bad some measurements could lead to a wrong result
- If only a few data are wrongly recorded these could be deselected (Out_of_Spec_1.dfq)
- If that is performed wrongly one may think the process is OK even if it isn't
- To select them open the value chart and go to | Part / characteristic | Select function
- Use for example the vertical selection and right click the falsely measured values



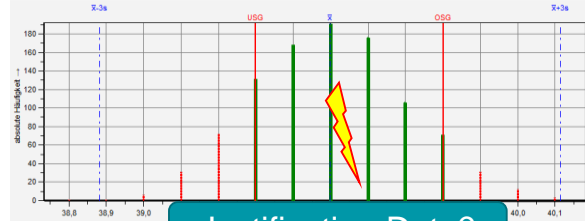
Notes:

Value plot individuals

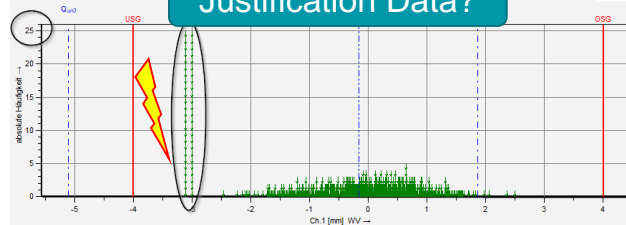
Insufficient resolution



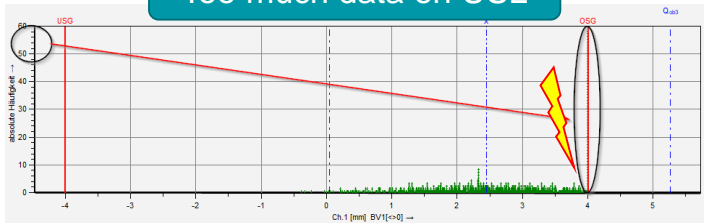
Mixed resolutions



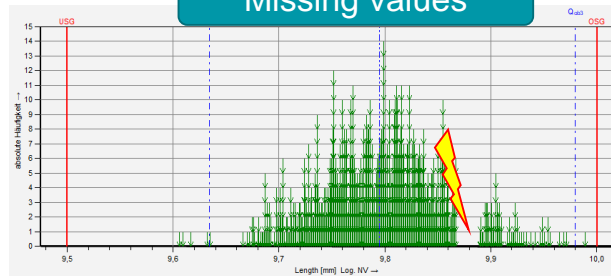
Justification Data?



Too much data on USL



Missing values



Notes:

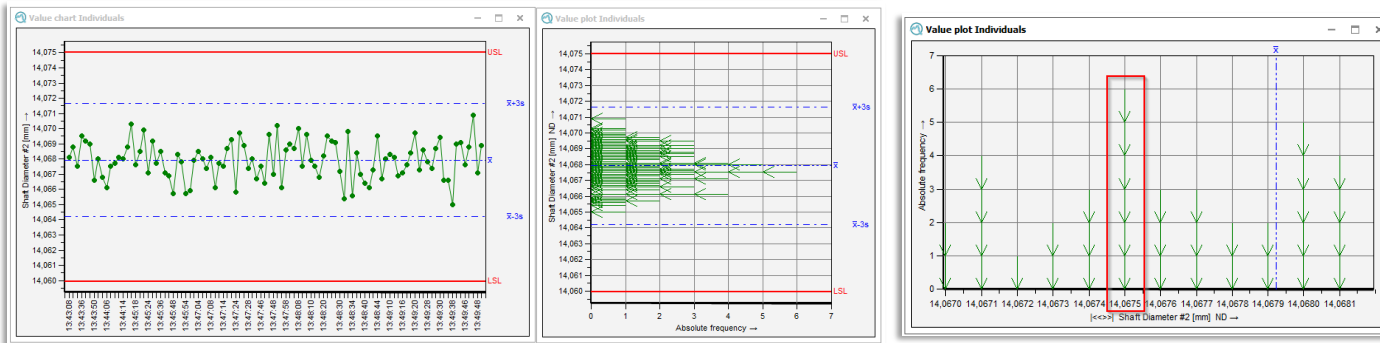
Exercise:

- Display value chart individuals and value plot individuals near each other
- Turn the value plot by 90° (| Graphical settings | Display | rotated)
- Which Value is the most frequent? (| Part / Characteristic | Zoom functions)

Rotated
The graphic is rotated, i.e. it is displayed with swapped axes.

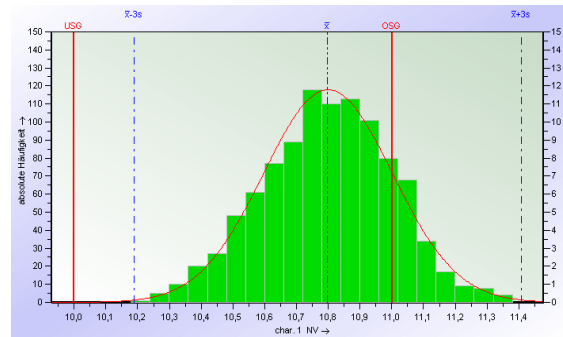
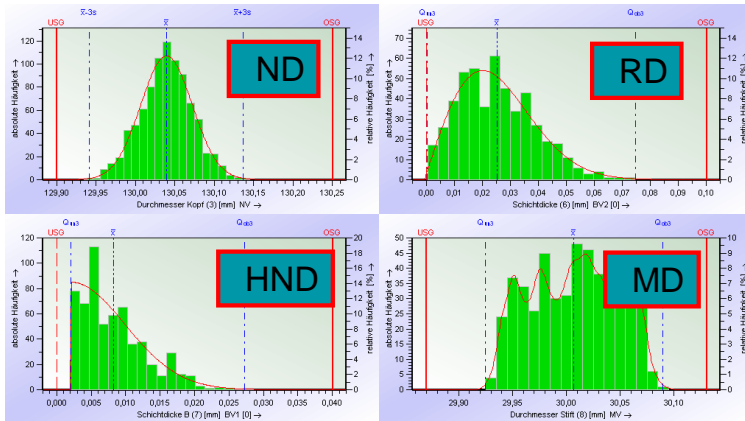
Zoom functions

- Start Zoom function (horizontally)
- Start Zoom function (vertically)
- Start Zoom function (horizontal + vertical)



Histogram

- Displays the deviation of the distribution (quantiles)
- Displays different means of the distribution (adjustable)
- The distribution model is automatically calculated
- The classification is set in the evaluation strategy
- It visualizes the ratio within/outside the specification

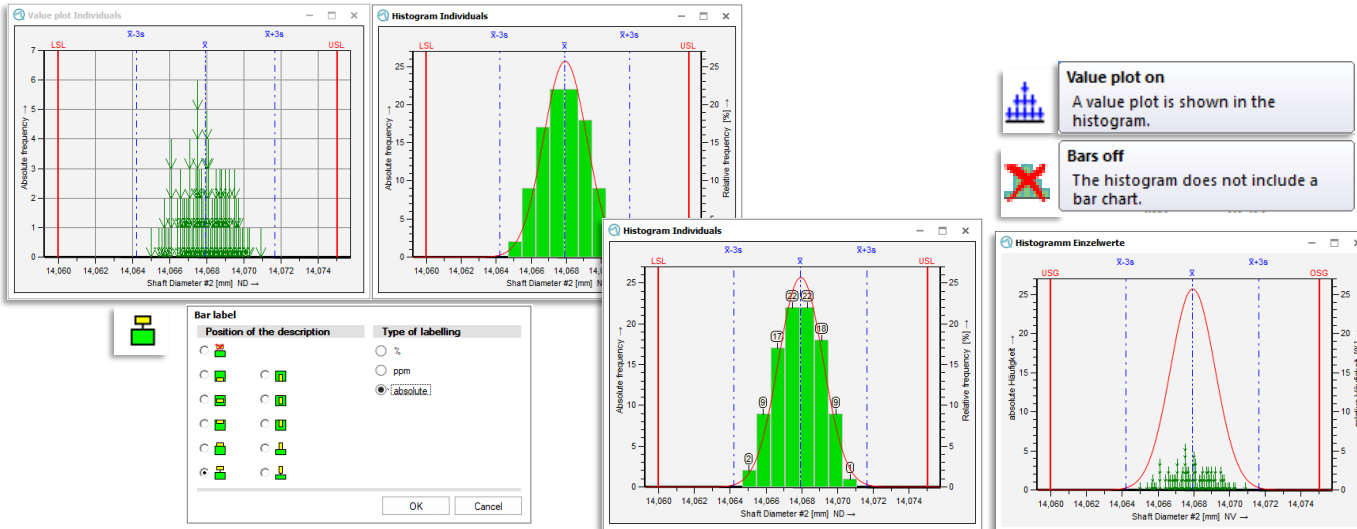


Notes:

Exercise:

- Display value plot and histogram next to each other
- Activate the value plot into the histogram
(| Graphical settings | Histogram |value plot on |)
- Display the number of values per class in the histogram
- Remove the display of bars in the histogram

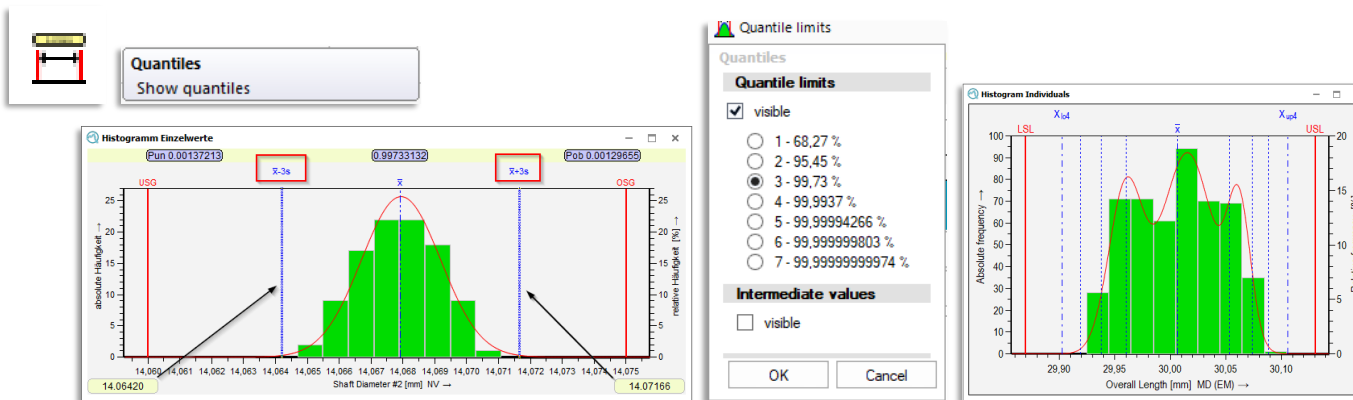
Notes:



Exercise:

- In a normal distribution are 6 σ the quantile limits (99.73%)
- Other distributions use quantile limits which are not calculated by 6 σ
- With “quantiles” it is possible to display different quantile limit areas
- The option |graphical settings | axes and limits | quantile limits changes the display of these limits (i.e. +/- 4 σ or X_{lo4} / X_{up4})

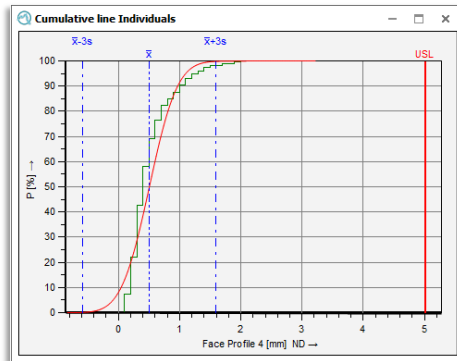
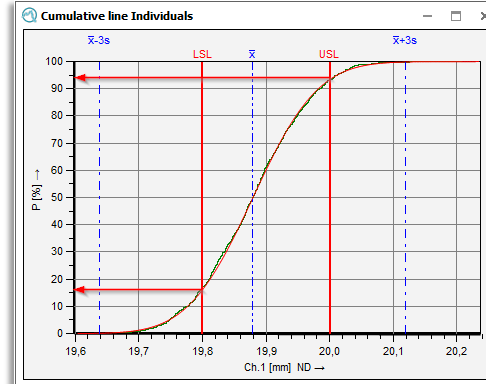
Notes:



Cumulative line

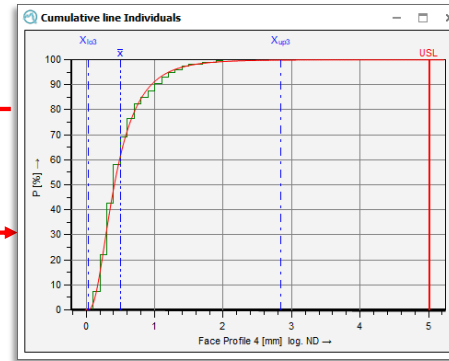
- % Scrap may be visible
- Is the distribution model plausible?
- Quantiles or standard deviations
- Value area

Notes:



Bad fit of the distribution

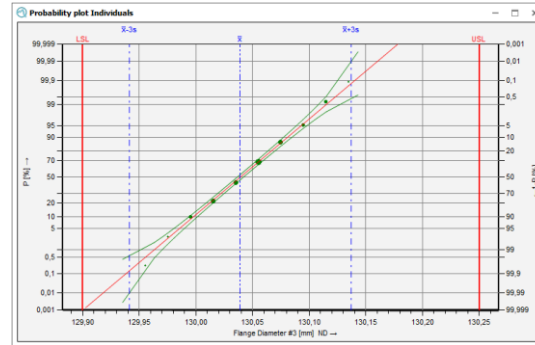
Good fit of the distribution



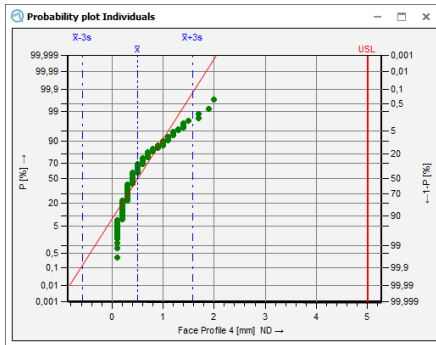
Probability plot individuals

- % Scrap may be visible
- The confidence interval can be displayed
- Distribution
- Comparisons between different distributions
- Wrong distributions

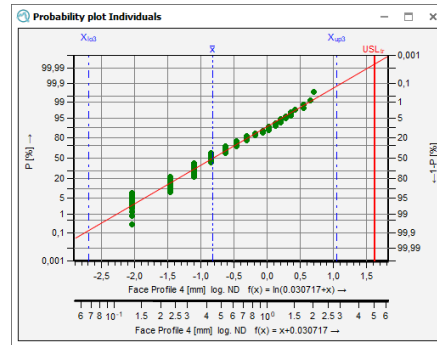
Notes:



Bad fit



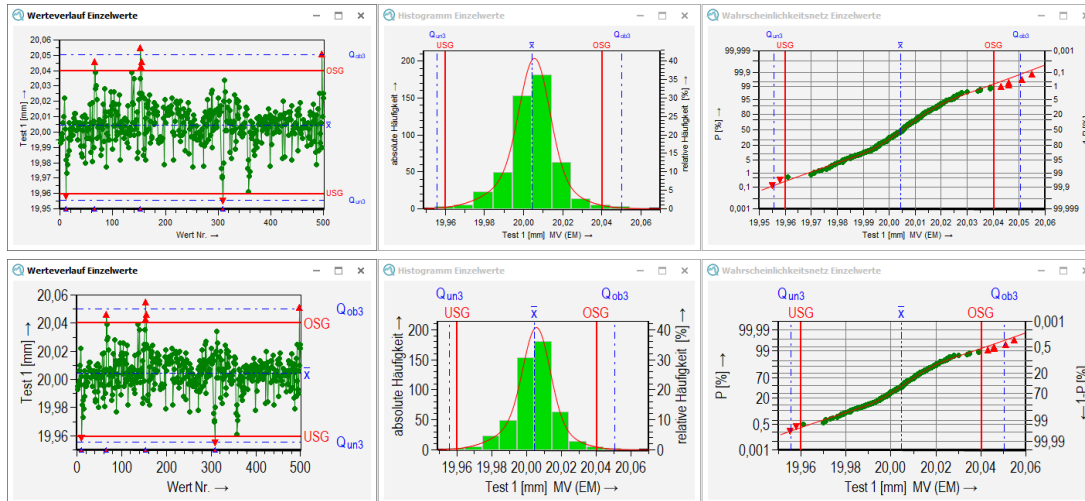
Better fit



Graphical settings

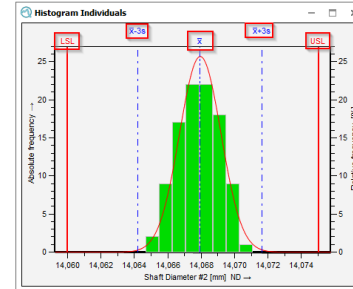
- Each graphic / form sheet / mask has its own graphical settings
- In in the ribbon “window” the option “all windows” is active a setting is active for every open graphic / form sheet / mask which could inherit this setting
- To circumvent this activate current window or close other graphics

Notes:



Capabilities in qs-STAT

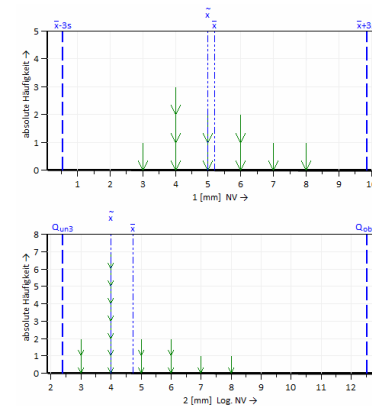
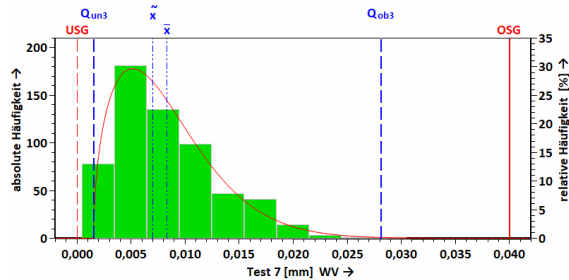
- What is needed to calculate the indices?
 - Mean
 - Deviation (Quantile/Standardabweichung)
 - One or two tolerance values
- A process / machine is capable if:
 - The expected no. of values outside spec is very low
 - It has a small deviation compared to the tolerance
 - It is located at the tolerance middle
- The description and requirements are depending on module, evaluation strategy and standard
- Pm/Pmk or Cm/Cmk (i.e. 1,67)
- Cp/Cpk, Pp/Ppk, CAM/CAP oder Tp/Tpk (i.e. 1,33)



Notes:

Process location

- Estimators for mean are:
 - Arithmetical mean
 - Median
- Depending on the formula the results may differ
 - The Median is less sensitive to outliers
 - The arithmetic Mean is more sensitive to outliers
- In a normal distribution the results may be more equal



Notes:

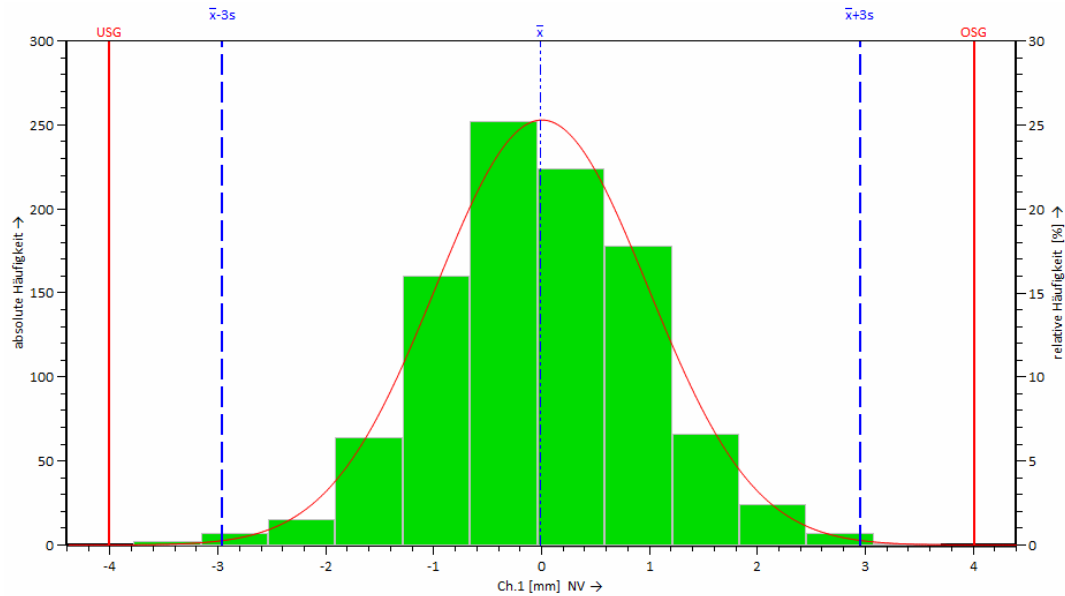
Deviation

- Which quantile limits describe the process best?

$$\pm 1s \triangleq X_{up/lo1} \triangleq \text{ca. } 68\%$$

$$\pm 2s \triangleq X_{up/lo2} \triangleq \text{ca. } 95\%$$

$$\pm 3s \triangleq X_{up/lo3} \triangleq \text{ca. } 99,73\%$$

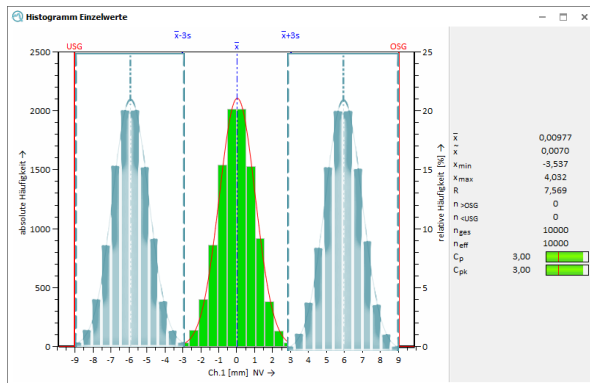


Notes:

Notes:

The potential capability index

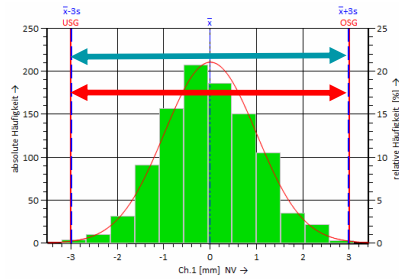
- The C_p value describes how often the process theoretically fits into the specification
- The C_p value describes the proportion between specification and deviation of the process (intention / ability = C_p)
- Necessary parameters are Specification (intention) and deviation (ability)
- In this case the process fits 3 times into the tolerance: $C_p = 3$



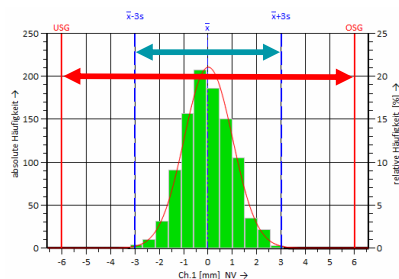
Notes:

Exercise

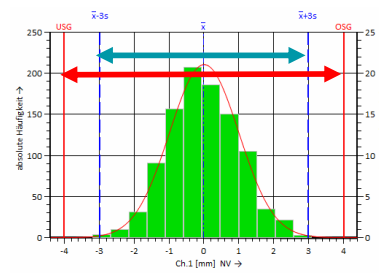
- Guess the Cp value (do not calculate!)



Intention = specification
 ability = deviation
 Intention / deviation = 1



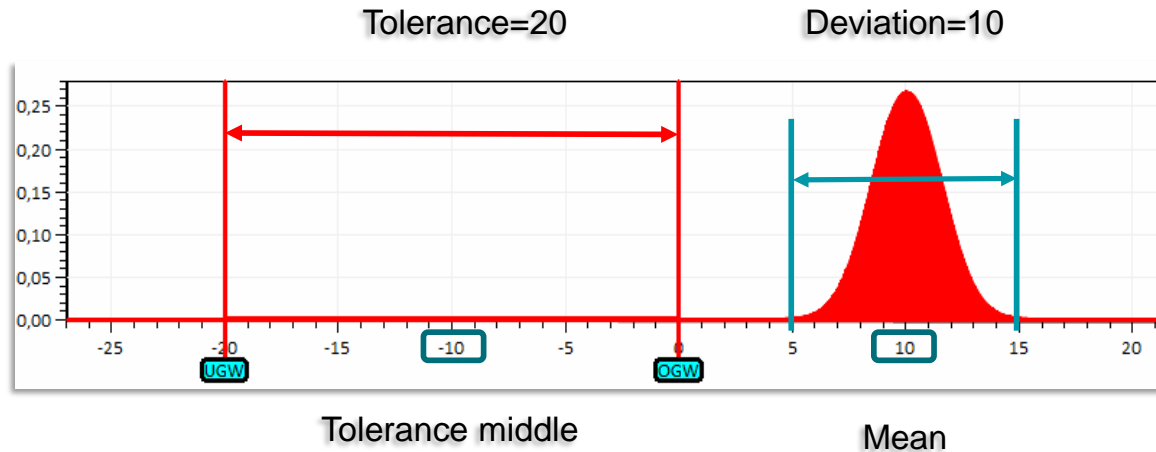
Intention / deviation
 Deviation fits two times
 into the specification
 $C_p = 2$



(intention) 8 / 6 (ability)
 =
 $1,33 = C_p$

Why do we need a Cpk?

- These are the reasons:
 - Specification / deviation = $20 / 10 = 2 = C_p$
 - The potential is OK, but no value lies within the specification
 - Cause: the mean is not taken into account
 - If the mean would be changed to -10 everything would be fine



Notes:

Notes:

Calculation of C_{pk}

- Divide the available tolerance from mean to one limit by half of the deviation
- Repeat that to the other side
- The C_{pk} is the smallest value of the two.

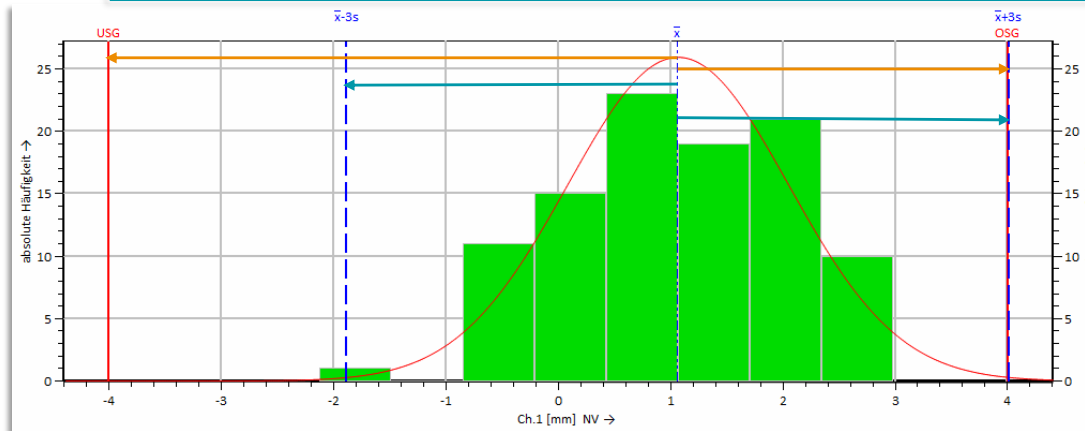
$$C_{pk} = \text{MIN}(C_{pkU}; C_{pkL})$$

$$C_{pkU} = \frac{\text{mean} - \text{USG}}{3s}$$

$$5 \text{ boxes} / 3 \text{ boxes} = 1,67$$

$$C_{pkL} = \frac{\text{OSG} - \text{mean}}{3s}$$

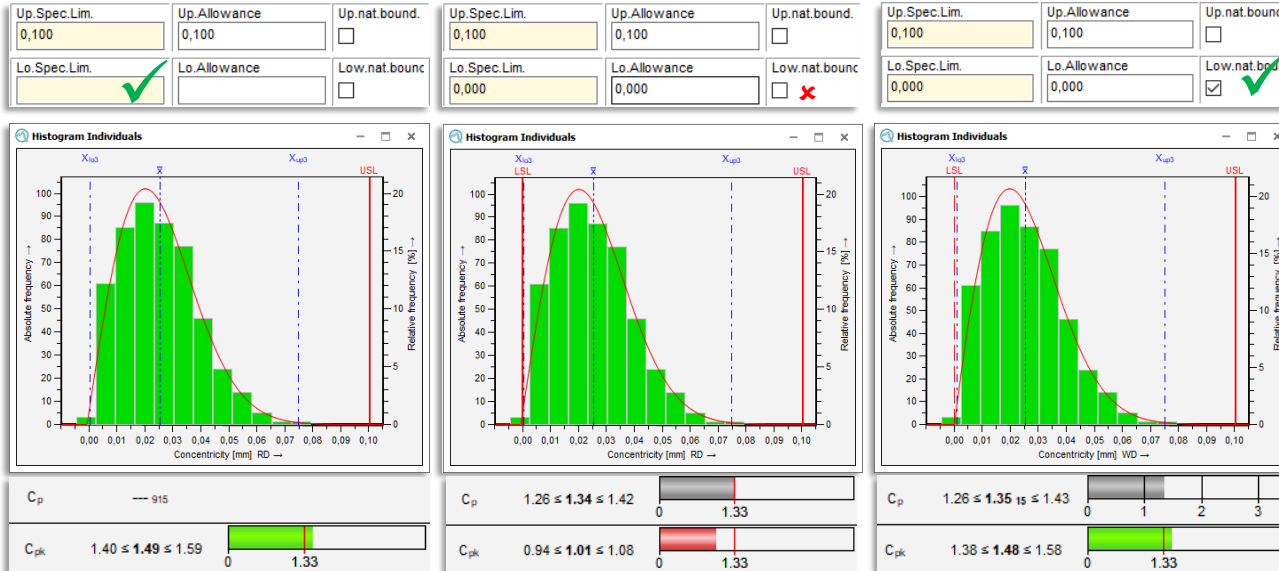
$$3 \text{ boxes} / 3 \text{ boxes} = \underline{1}$$



Notes:

Special case: one sided tolerances

- Used in all shape and position tolerances
- The Cpk will only be calculated to one direction (to the „critical“ limit)
- Alternatively the lower specification is omitted
- If measurement devices write 0 but do not set a natural border the result may be bad



Notes:

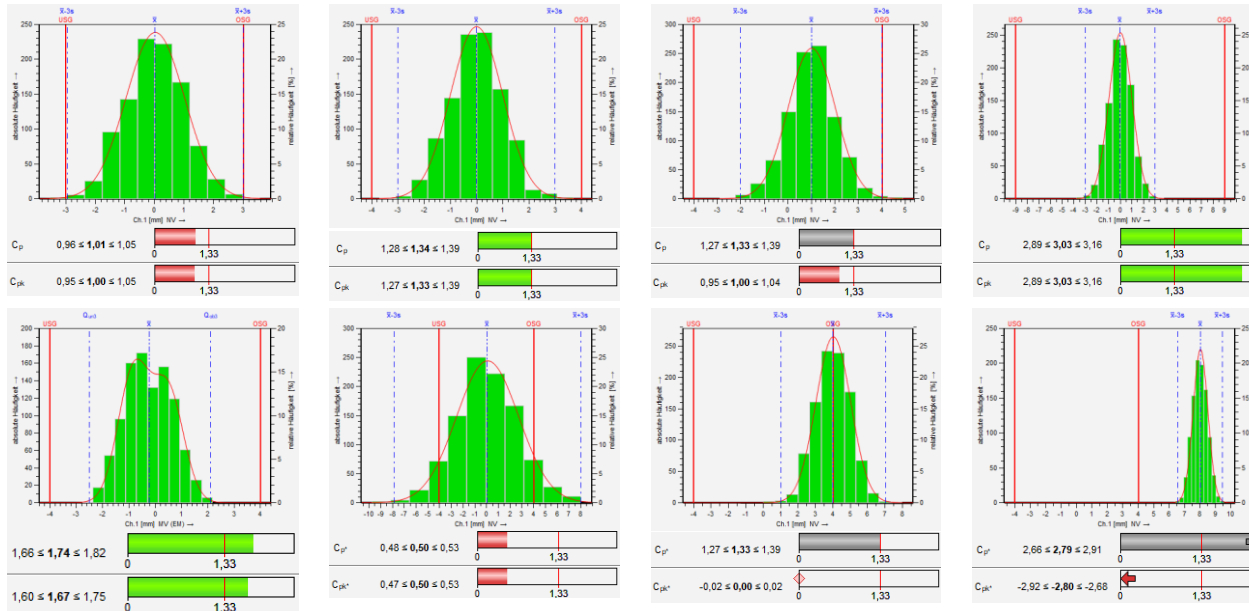
Nice to know:

- If $Q_{l03} = L_{sl}$ or $Q_{u03} = U_{sl}$ the C_{pk} in this direction is 1
- If the mean/median is on a specification limit $C_{pk} = 0$
- If the mean/median is outside the specification C_{pk} is negative
- In one sided specifications C_{pk} is crucial and C_p may be suppressed
- In non normal distributed characteristics normally the quantile limits (X_3) will be taken for calculation (percentile method)
- If the distribution is changed C_p/C_{pk} will most likely also change(!)

Exercise

- Guess Cp/Cpk without calculation

Notes:



Notes:

Conclusion

- Cp and Cpk are metrics
- The figures are an objective possibility to compare characteristics
- They are depending on the evaluation strategy
- It is necessary to work with one strategy to compare statistics