QC – Yield Statistics (YST) S35

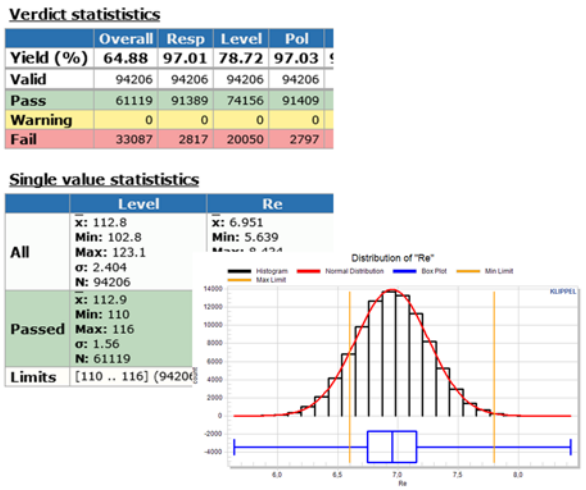
Module of the KLIPPEL ANALYZER SYSTEM (QC Version 7.0, dB-Lab 212; Document Revision 1.6)

FEATURES	BENEFITS
• Simple overview of production statistics • Filter date, time, operator, and SN • Automatic data collection • Open output format (CSV) • Statistics of verdicts and single values • Use for regular statistics • Observe distribution of single values	• Supervision of production • Overview of verdicts and single value results • Simple interface to 3rd party software • Generate customized reports of yield and single values • Keep track and identify drifts of single values (e.g. coil shift versus time) • Generate customized reports

This complementary module delivers a quick statistical overview on test results of production lines. Log files, generated by *Klippel QC* software are processed and the results are shown as *Verdict statistics* and *Single value statistics*.

The data can be filtered by date/time, operator, SN and single value names. Different production lines can be parsed, by copying the log files to a network folder that is accessible for the statistic module.

Third party statistics software can be used by exporting a CSV file of the parsed data.



Verdict statistics

	Overall	Resp	Level	Pol
Yield (%)	64.88	97.01	78.72	97.03
Valid	94206	94206	94206	94206
Pass	61119	91389	74156	91409
Warning	0	0	0	0
Fail	33087	2817	20050	2797

Single value statistics

	Level	Re
All	x: 112.8 Min: 102.8 Max: 123.1 σ: 2.404 N: 94206	x: 6.951 Min: 5.639
Passed	x: 112.9 Min: 110 Max: 116 σ: 1.56 N: 61119	
Limits	[110 .. 116] (9420	

Application:

- Overview of yield
- Regular (daily, weekly, …) statistics
- Overview of verdicts
- Overview of single value results

CONTENT:

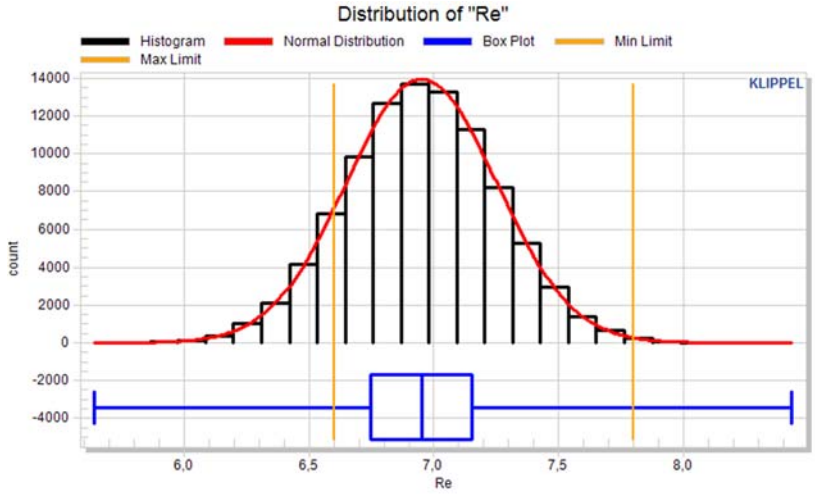
1 Summary

The *Yield Statistics* module processes log files produced by the *Klippel QC* software. No hardware is required for executing this module.

The *Yield Statistics* can be used to generate regular statistics. Refer to application note *AN 48 Yield Statistics* for examples.

2 Definition of Statistics

Valid Measurements	The verdict counters are fed with data from the parsed log files. The following counters are available and represent valid measurements: • Pass • Warning • Fail The total number of valid measurements is shown additional to these counters. Please note that the verdict Warning equals a Pass from the yield point of view.
Yield	The yield *Y* is defined as the ratio of number of passed DUTs N_{pass} (*Pass* and *Warning*) to total number of all valid measurements N_{total} (*Pass*, *Warning* and *Fail*) in percent $\mathrm{Y}=\frac{N_{\text{pass}}}{N_{\text{total}}}\cdot 100\%.$ The yield is calculated for all available individual verdicts. The overall yield is highlighted.
Invalid Measurements	Invalid measurements (verdicts *Invalid*, *Noise* and *Void*) do not contribute to the calculation of the yield. The invalid measurements are shown below the valid measurements. Please note, that the invalid measurement verdicts are not shown if they are all zero.
SINGLE VALUE STATISTICS	
All / Pass	Each of the following statistical values is calculated both for the overall passed single values and for all valid single values.
Mean ($\bar{x}$)	The mean $\bar{x}$ of a single value is defined as the ratio of the sum of the single value $\sum_{n=1}^{N} x(n)$ to the number of samples N $\bar{x} = \frac{\sum_{n=1}^{N} x(n)}{N}.$ **Note:** The responsibility to interpret the single value statistics lies with the user. For example: a linear averaged level differs from terms like "averaged SPL". The total number of samples equals the number of valid measurements (verdicts Pass, Warning and Fail).
Min/Max	The smallest/largest element found in the parsed data $x_{\min} = \min_{n} x(n),$ $x_{\max} = \max_{n} x(n).$

Deviation (σ)	The standard deviation σ of the single value is defined as follows: $\sigma = \sqrt{\left(\sum_{n=1}^{N} x(n)^2\right) - \frac{1}{N}\left(\sum_{n=1}^{N} x(n)\right)^2}$ $= \sqrt{\left(\sum_{n=1}^{N} x(n)^2\right) - N \cdot \bar{x}^2}.$
Samples (N)	The number of valid measurements due to the overall verdict.
Limits	Minimum and maximum value of the limit with absolute number of occurrences of this combination of minimum and maximum limit. If the limits change (e.g. in different log files), the different limits are listed one below the other. If no limit is available for a single value "-" is shown.
DISTRIBUTION ANALYSIS	
Histogram	Distribution analysis of single values as a histogram plot:  Note: The histogram cells are defined as right closed (left open) intervals. Only the very first cell is a right closed (left closed) interval to ensure that all single values are taken into account.

3 Input Parameters

Location of log files	Directory containing the log-files to parse. Log-files may be in sub folders, files are searched recursively. Absolute and relative (to the database's location) paths are allowed. If *Folder* is empty, the database's location is used.
Name of statistic	Name of test to be shown in HTML output window. If not specified the project's name is extracted from the database's parent folder.
Date Filter	Date range of statistics. • YYYYMMDD or YYMMDD: The single date notation specifies the start date of the statistics. The end date is set to now. • YYYYMMDD-YYYYMMDD or YYMMDD-YYMMDD: The date range notation specifies the start date and the end date of the statistics. Use either 2-digit or 4-digit year notation!

	The short notation can be used for flexible and fast date/time range input. It is explained with examples in section *Placeholder Tokens for Date/Time Input* below. If the short notation is used, the *Time Filter* will be ignored. Please note, that the date range is displayed in the configuration table in the window *HTML out*. If *Date Filter* is empty, all available DUTs are considered, until now.
Time Filter	Time range of statistics. This parameter specifies the exact time at the beginning and end of the date range. **HH:MM:SS:DS** or **HH:MM:DD:DS-HH:MM:SS:DS** The time range is not interpreted if the short notation for date is used. The time notation can be a range or a single time specifier. If not specified as range the end time is set to 23:59:59:9. The time specifier can be coarse (e.g. HH:MM, HH:MM:SS, HH-HH, HH:MM:SS-HH:MM:SS) and asymmetric (e.g. HH-HH:MM:SS). Please note, that the time range is also displayed in the configuration table in window *HTML out*.
Operator	Consider only specified operator. You could also separate multiple operators by line feed. This parameter is case-insensitive.
define SN prefix	Check this checkbox if you want to explicitly define a prefix for filtering the serial numbers. If not checked, any prefix of the serial numbers is ignored.
SN prefix	Define your serial number prefix here (Example: “ABC_”). If Prefix is empty only serial numbers without any prefix are shown. Note: Prefixes and serial numbers are separated at the last non-digit (e.g, whole SN = “g7h_38” → prefix = “g7h_”; number = “38”)
SN start value	Start value of range of serial numbers to be filtered.
SN end value	End value of range of serial numbers to be filtered.
Ignore SV	List of single values that should not be considered in the verdict and single value statistics. These single values are not exported to the CSV file, too. This parameter is case sensitive. Multiple single value names could be separated by line feed.
CSV-File	Writes all parsed data into a merged CSV file. The file is located in the specified folder. Absolute and relative (to the database’s location) paths are allowed. Note: If specified file exists already, it will be overwritten without any warning.

Edit plots	Configuration for distribution analysis of single values. The configuration is done in a table with each row representing a separate plot. Single rows could easily be commented by using "//" at the beginning of the parameter **SV**. There is no limitation on the number of plots. The first line is the heading of the configuration table..
	Available parameters are:
	• **SV:** specifies the name of the single value to be analyzed. (Example: 'Fs')
	• **Type:** specifies the chart type, currently only **dist** type is available.
	• **Filter:** (optional) configures the range of considered single value. This can be a string **Outlier** (for automatic determination of outliers with GRUBBS' algorithm), **Limits** (everything within the QC limit range for this single value) or a range **[start,end]** (definition of absolute range).
	• **Bins:** (optional, only for distribution plots) number of bins for histogram plot. Individual start and stop values for histogram bins may be defined with a vector. This parameter is recessive against the parameter **Filter**.
	Example:
	`'//SV' 'TYPE' 'FILTER' 'BINS' 'UNIT' 'INTERVAL'`
	`'//Re' 'dist' '[1,10]' '30' '' ''`
	`'Re' 'dist' 'limit' '' '' ''`

4 Placeholder Tokens for Date/Time Input

YYYYMMDD or YYMMDD	The single date notation specifies the start date of the statistics. The end date is set to current time. Use either 2-digit or 4-digit year notation!
YYYYMMDD-YYYYMMDD or YYMMDD-YYMMDD	The date range notation specifies the start date and the end date of the statistics. Use either 2-digit or 4-digit year notation!
Short notation of statistic's date/time range [Y|Q|M|W|D|H|{-|n}]	Note: Since the short notation defines the time range as well, the time notation is not considered if the short notation is used. The short notation specification can relate to either: • Year: **Y**, • Quarter of year (Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec): **Q**, • Month of year: **M**, • Week of year (starting with Sunday): **W**, • Day of month: **D** or • Hour of day: **H**. Only one specifier for a period of time can be used: • **Alone** to specify current period of time, • **With minus sign** to specify last period of time or

- **With a positive number** to specify an arbitrary period of time

Examples (assuming now is Wednesday 13th July 2011, 11:15:09:2, (2011-07-13 11:15:09:2)):

Example	Interpreted date/time range
Y	2011-01-01 00:00:00:0 - 2011-07-13 11:15:09:2
Q	2011-07-01 00:00:00:0 - 2011-07-13 11:15:09:2
M	2011-07-01 00:00:00:0 - 2011-07-13 11:15:09:2
W	2011-07-10 00:00:00:0 - 2011-07-13 11:15:09:2
D	2011-07-13 00:00:00:0 - 2011-07-13 11:15:09:2
H	2011-07-13 11:00:00:0 - 2011-07-13 11:15:09:2
Y-	2010-07-13 11:15:09:2 - 2011-07-13 11:15:09:2
Q-	2011-04-13 11:15:09:2 - 2011-07-13 11:15:09:2
M-	2011-06-13 11:15:09:2 - 2011-07-13 11:15:09:2
W-	2011-07-06 11:15:09:2 - 2011-07-13 11:15:09:2
D-	2011-07-12 11:15:09:2 - 2011-07-13 11:15:09:2
H-	2011-07-13 10:15:09:2 - 2011-07-13 11:15:09:2
Y9, Y2009	2009-01-01 00:00:00:0 - 2009-12-31 23:59:59:9
Q4	2010-10-01 00:00:00:0 - 2010-12-31 23:59:59:9
M7	2011-07-01 00:00:00:0 - 2011-07-13 11:15:09:2
W3	2011-01-09 00:00:00:0 - 2011-01-15 23:59:59:9
D31	2011-06-30 00:00:00:0 - 2011-06-30 23:59:59:9
H12	2011-07-12 12:00:00:0 - 2011-07-12 12:59:59:9

5 Handling Large Amounts of Data

In general, there is no limitation of the size of data to be analyzed. To avoid trouble with plotting of very huge data sets there is a threshold of 100.000 data sets. If the requested single value has more data sets, it will be quantized to reduce the data. If the number of samples is below this threshold there is no change of the data.

In case of quantization of the single values a relative error of <0.1% is guaranteed.

Find explanations for symbols at:

http://www.klippel.de/know-how/literature.html

Last updated: May 27, 2021