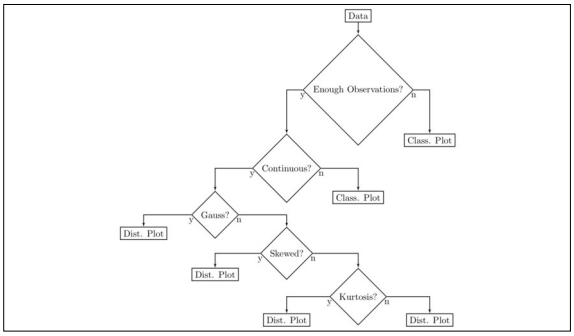

Alexander van Schie

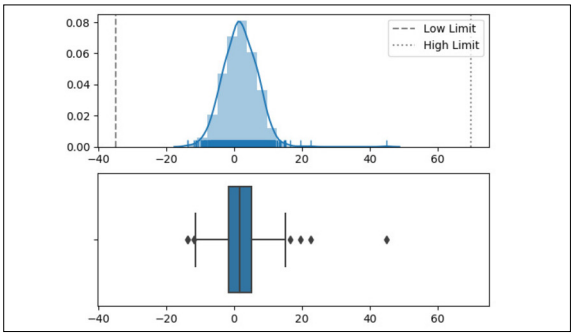

Nicolas Manuel Ganz

Graduate Candidates	Alexander van Schie, Nicolas Manuel Ganz
Examiner	Prof. Dr. Lin Himmelmann
Co-Examiner	Jürg Dietrich, AXA, Winterthur, ZH
Subject Area	Software
Project Partner	ams International AG, Rapperswil, SG

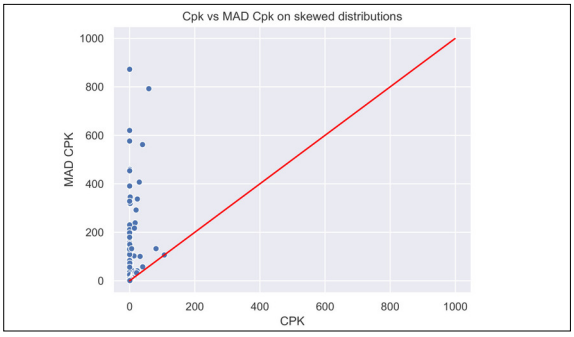
Statistical Analysis of Test Data in Microelectronics Industry



The decision tree to determine the type of distribution
Own presentment



An example of a distribution plot with random data points
Own presentment



Comparison of the MAD Cpk to the CPK for skewed distributions
Own presentment

Problem: The international company ams AG is a producer of semiconductors. In order to meet the customer's expectations ams AG needs to guarantee to only ship parts which meet the specifications. Therefore, several tests will be executed on the parts. The results need to stay within the specification limits. If a part is systematically out of range, a tester needs to analyse the test data. The goal of this bachelor thesis is to simplify the process of analysing the cause of faults. ams AG is looking for a tool meeting the following requirements:

1. Scan test data logs and calculate robust measures based on the type of distribution.
2. Create an algorithm which detects and removes outliers. The identified parts are likely defective devices.
3. Build an interactive GUI for analysis and manipulation purposes. This task is optional.

As an asset for development, ams AG provided real project data.

Approach / Technology: An initial data analysis helped to understand the structure of the data and which information could be omitted to improve the performance.
With the reduced complexity, multiple tests were performed to identify whether a measure is robust or not.
In parallel, different outlier detection algorithms were compared with each other. Metrics were examined to decide whether or not omitting those outliers have a significant influence on the data.
Finally, appropriate visualisations of the data were developed to optimise the display of the obtained results. These are combined into a report, highlighting the relevant metrics.

Result: A decision tree was built to classify the distributions. Depending on the type of distribution, a matching quality measure was applied. For example using the MAD Cpk is more significant than using the Cpk for skewed and tailed distributions.
The outlier detection was done by applying the squared Mahalanobis distance. The significance of omitting the outliers was measured by the Cook's squared distance. Using these metrics, negatively influential data points can be removed and result in more stable distributions.
To visualise the results, an HTML report was built. This offers an overview of the data and helps to efficiently determine the quality of the tested parts.