Proceedings of the 9th International Conference on Applied Informatics
Eger, Hungary, January 29–February 1, 2014. Vol. 1. pp. 161–169
doi: 10.14794/ICAI.9.2014.1.161

Challenges in test system verification

Istvan Poser

NI Hungary Kft
thomas.brown@gmail.com
University of Debrecen, Faculty of Informatics
poser.istvan@gmail.com

Abstract

The Test Systems in manufacturing applications are highly complex measurement systems. The quality of the Test Systems is defined by the statistical properties of multiply repeated measurements. If the measurement results are „close" to each other and to a known good reference value, then the quality of the test system is considered „good". The most common reasons behind a „bad" quality measurement system can be faults, inaccuracies or their interactions with the environment, which can directly result in poor yields, false failures and quality risks. If the variation of the measurement results becomes too large, then they might even mask the variations in the manufacturing process.

The need is valid for automated or semi-automated system validation methods that can assure the quality of our measurement results. There are a number of methods used in the industry that can provide the required confidence in our measurements.

Keywords: measurement system, verification, Gage R&R

1. An overview of the test system

Test Systems using card modular platforms can provide the same functionality as rack and stack instruments to some extent on a much smaller footprint and a reduced cost. For these reasons the majority of the automatic test systems in manufacturing use card modular platforms. A typical card modular architecture usually includes modular instrumentation; an embedded controller and some type of control and trigger bus architecture [5].

The smaller space requirements also add an increased complexity to the test system. This introduces new type of requirements when it comes to verification of functionality or specification of the system or recertification of the instrumentation. Verification of the test system's accuracy has traditionally been done through the incorporation of external standards that were used to verify the accuracy of instrumentation used in the test system individually. This method can introduce problems in case of higher complexity systems, since just by removing modular instrumentation for verification or recertification purposes, we cause downtime to the test system and we might even cause connection or other different type of issues [5].

With wide usage of the card modular platform a new verification and certification strategy is required to be able to verify instrumentation inside the test system. A system based verification method in conjunction with application specific accuracy requirements can offer an improved method.

2. Process capability index

The process capability index is a number that compares the capabilities of a product and the processes used to manufacture that product to engineering specifications. A large process capability index indicates that the process is capable of producing products stably, that will meet or even exceed the published specifications. This index reduces all the complex information about the process to a simple number that can be used to monitor the quality of the assembled product or to compare changes in capabilities, when new processes are introduced to manufacturing [7].

2.1. Definition of capability indexes

The capability index is defined as the ratio of the distance from the process center to the nearest specified limit divided by the measure of the process variability. This is illustrated in Figure 1. The two most widely used capability indices are defined as:

$$C_{pk} = \min\left[\frac{USL-\mu}{3\sigma_{\overline{R}/d_2}}, \frac{\mu-LSL}{3\sigma_{\overline{R}/d_2}}\right]$$

$$P_{pk} = \min\left[\frac{USL-\mu}{3\sigma_s}, \frac{\mu-LSL}{3\sigma_s}\right]$$

Where USL and LSL represent the specified upper and lower limits and $\sigma_{\overline{R}/d_2}$ and σ_s and μ represent the process standard deviation and the mean. These values are usually calculated from the data collected from the process.

The estimate σ_s is the sample standard deviation $\sqrt{\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2 \Big/ (m-1)}$, whereas $\sigma_{\overline{R}/d_2} = \overline{R}/d_2$ is an estimate using the subgroup ranges. Where d_2 is an adjustment factor needed to estimate the process standard deviation from the average sample range. A large C_{pk} and P_{pk} value should represent a process that is capable of producing the vast majority of units within the specified limits [7].

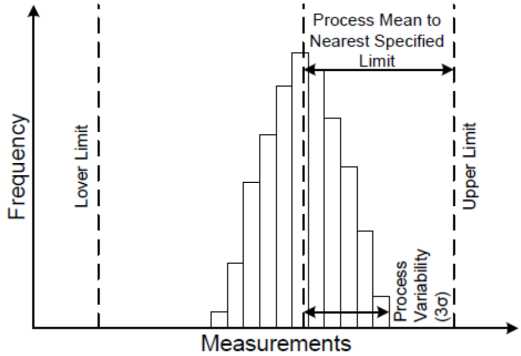

Figure 1: Graphical representation of the Capability Index

2.2. Problems with capability indexes

The measures C_{pk} and P_{pk} only defer in the estimation of the process standard deviation. In order to compare the two capability indexes, we need to compare the two process standard deviation estimation methods.

The range-based estimation $\sigma_{\overline{R}/d_2}$ only uses the variability within the subgroups to estimate the process standard deviation. While σ_s on the other hand combines the data together and uses both the within subgroup and between subgroup variability. As a result, σ_s estimates the standard deviation of the entire process, while C_{pk} can seriously can underestimate the total variation if the between subgroup variation is substantial [6].

Let's take an example from the electronics manufacturing industry. Two parameters were measured on a sample of 100 Units Under Test (from now on UUTs). The two parameters are the gain error and the offset error of the voltage signal produced by the UUTs compared to a reference value.

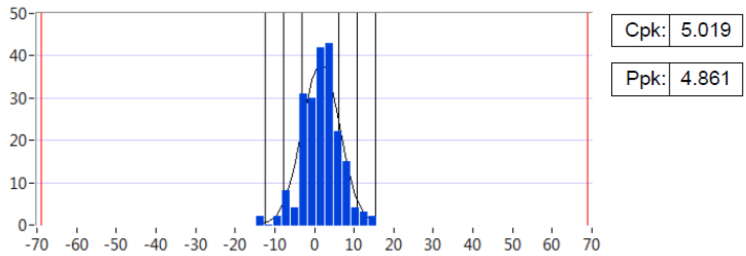

Figure 2: Voltage Accuracy Verification, Gain Error

As it can be seen from the above results too P_{pk} provides more pessimistic information about the capabilities of the manufacturing process which can give more confidence in the results.

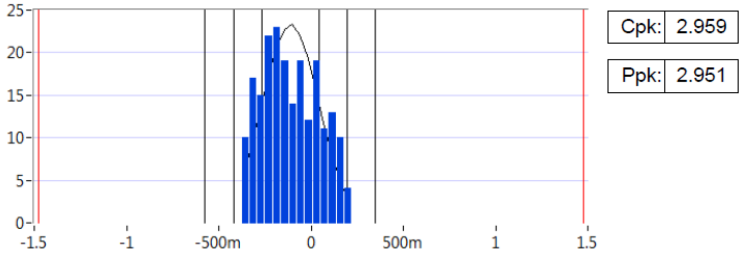

Figure 3: Voltage Accuracy Verification, Offset Error

Although the Process Capability Analysis represents the process capabilities with a simple, in electronic manufacturing there can be several reasons behind poor capabilities. The results of a simple voltage measurement can be affected by several hundred components taking part in generating that voltage, or the process that assembles the circuitry, or even the measurement equipment measuring the voltage. The process capability index does not give any information about the source of poor capabilities.

3. Gage repeatability and reproducibility

In process control the goal is to improve the quality of the manufactured products through the reduction of process variability. This process relies primarily on measurement and test data as input. To be able to address process variability problems, the variation due to the measurement system must be separated from the process variability.

In a measurement system the possible sources of variation are the gages (repeatability), the appraisers (reproducibility) and the variation within the sample (part-to-part variation). Repeatability and reproducibility together are the components of the measurement system variation. The GR&R study quantifies the variation relative to the total variation and to the specification range [2].

$$\sigma^2_{total} = \sigma^2_{part} + \sigma^2_{repeatability} + \sigma^2_{reproducibility}.$$

3.1. Repeatability and reproducibility

Repeatability is the variation in measurements, when the same operator measures the same parameters on the samples using the same measurement equipment, the same gage. This is represented of the left of Figure 4.

Whereas reproducibility in our case represents the variation in the measurements, when the same operator measures the same parameters on the samples using multiply measurement equipment of the same type. This is represented on the right of Figure 4. Reproducibility is the variation of the average measurements taken with different gages. Although their variability is the same independently, due to the differences in their bias, the total variability of the measurement systems increase.

The objective of a GR&R study is to determine if a measurement system is capable of monitoring a process. If the measurement system error is small relative to the total variation, then it is adequate. The number of samples selected should belong to the same process. The more is the number of samples, measurement systems and repeats, the grater the confidence can be in our results, but the higher numbers also have a higher cost impact and more time is required to conduct the study [2].

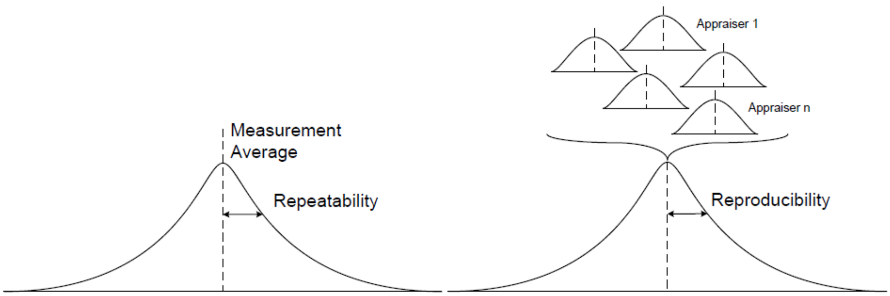

Figure 4: Repeatability and Reproducibility

3.2. Evaluating the measurement data

The variation caused by the equipment during replication of the measurement is the equipment variation (EV) and it is an estimate of the repeatability attribute of the GR&R study (Step 2. in Table 1). The variation caused by the differences between the measurement systems used is the appraiser variation (AV) and it is an estimate of reproducibility (Step 3. in Table 1). Part variation is attributed by variations in the process of manufacturing of the part (Step 5. in Table 1).

Where:

- UCL – Upper Control Limit
- d_2, D_4 – Control Chart Constants
- n, a, p – Number of trials, appraisers, parts
- R_a – Range of appraiser averages
- R_p – Range of part averages

In order to make a decision about the measurement system based on the study a criteria index was set up. If the %GRR is less than 10%, then the measurement

Steps	Formula
1.	$UCL = D_4 \times \overline{R}$
2.	$EV = \overline{R}/d_2$
3.	$AV = \sqrt{R_a/d_2^* - EV^2/n \times p}$
4.	$GRR = \sqrt{EV^2 + AV^2}$
5.	$PV = R_p/d_2^*$
6.	$TV = \sqrt{EV^2 + AV^2 + PV^2}$
7.	$\%EV = 100 \times [EV/TV]$, $\%AV = 100 \times [AV/TV]$, $\%GRR = 100 \times [GRR/TV]$, $\%PV = 100 \times [PV/TV]$

Table 1: Steps of GR&R

is acceptable. If the %GRR is greater than 30%, then the system is unacceptable and it needs improvement. Between the two percentage values the system might be acceptable dependent of the application. But of course the acceptance should not be decided based on a single set of measurement data and the long term performance of the system should also be monitored with different graphical analysis methods [1, 2].

3.3. Considerations regarding the GR&R study

Implementing GR&R in an electronic manufacturing environment has its challenges. In order to have a high confidence in the results a higher number of samples and a higher number of repeats are required. This can be problematic for products that have several hours long test time and also cost a lot to manufacture. It is always a case-by-case decision of the balance between the level of confidence we would like to have in our results and the amount of time and money we can spend on the study.

4. Paired t-test

The paired t-test is used to compare two population means, where we have two samples in which observations in one sample can be paired with observations in the other samples. This method can be used to:

- Compare before-and-after observations on a process (e.g. comparing measurement results before and after conducting an experiment on the samples)

- Compare two measurement systems with measuring parameters of the same sample set.

4.1. Conducting a paired t-test

Samples of n products were used to test a new process and based on the results decide if the new process has any significant affect on the quality of the products. Initially we need to test the null hypothesis that the mean of the differences between the two set of measurement results is zero. In order to achieve this, the procedure is as follows [6]:

- Calculate the difference between the two observations on each pair ($d_i = y_i - x_i$), making sure to distinguish between positive and negative results.
- Calculate the mean difference ($\overline{d} = \sum_i d_i/n$).
- Calculate the standard deviation of the differences

$$s_d = \sqrt{\sum_{i=1}^{n} \left(d_i - \overline{d}\right)^2 \Big/ (n-1)}$$

  and use it to calculate the standard error of the mean difference ($SE(\overline{d}) = s_d/\sqrt{n}$).
- Calculate the t-statistics ($T = \overline{d}/SE(\overline{d})$) which follows a t-distribution with n-1 degrees of freedom under the null hypothesis.
- Use tables of the t-distribution to compare the calculated T value to the $t_{\alpha/2,n-1}$ value where α is the probability and n-1 is the degrees of freedom. If $-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}$ is true, then our null hypothesis is acceptable.
- If $\overline{d} - t_{\alpha/2,n-1} \times SE(\overline{d}) \leq \mu_d \leq \overline{d} + t_{\alpha/2,n-1} \times SE(\overline{d})$is also true, then we can be confident with a level of 100(1-α)%, that our measurement results between the two observations do not have differences.

Table 2 shows a concrete example for a paired t-test, that is used to identify if a new manufacturing process has significant impact on the capabilities of the products manufactured with it, or not. Based on this, the null hypothesis is, that the process has no impact on the capabilities of the products.

With using a 95% of confidence and the degrees of freedom of 14, $t_{\alpha/2,n-1} = 2.145$. This means, that our null hypothesis is acceptable. And with

$$\overline{d} - t_{\alpha/2,n-1} \times SE(\overline{d}) \leq \mu_d \leq \overline{d} + t_{\alpha/2,n-1} \times SE(\overline{d})$$

$$6.82 - 2.145 \times 4.7 = -3.26 \leq \mu_d \leq 6.62 + 2.145 \times 4.7 = 16.91$$

also being true, we can be 95% confident, that the new process will not have an impact on our product capabilities.

Samples	Before Process	After Process	y_i-x_i	$(d_i-\hat{d})^2$
1	-34.695	-32.227	2.468	19.00495
2	1.488	1.166	-0.322	51.11487
3	-38.395	-35.452	2.943	15.08908
4	-18.485	33.021	51.506	1996.171
5	-158.194	-163.615	-5.421	150.0249
6	-84.81	-72.796	12.014	26.90013
7	-85.302	-89.389	-4.087	119.1256
8	-125.554	-84.706	40.848	1157.397
9	-75.189	-54.044	21.145	204.9918
10	-15.877	-22.76	-6.883	187.9769
11	-20.028	-20.058	-0.03	47.02485
12	-11.183	-10.759	0.424	41.00439
13	-7.737	-8.604	-0.867	59.20482
14	36.879	42.648	5.769	1.120352
15	-98.888	-115.983	-17.095	572.2844
$\hat{d}$			6.827466667	
s_d			18.22171987	
SE($\hat{d}$)			4.70482784	
T			1.451161849	

Table 2: Paired t-test

4.2. Consideration regarding the paired t-test

Paired t-tests have advantages when thinking of comparing several test systems to each other. It is simple to implement and provides detailed information on confidence. It can be expanded to multiply test systems and with more data the estimation could be improved too. So it is certainly a promising method for test system verification purposes that needs further research.

References

[1] An Honest Gauge R&R Study, Donald J. Wheeler, 2006 ASQ/ASA Fall Technical Conference

[2] Measurement System Analysis, Reference Manual, First Edition, October 1990 Second Edition, February 1995; Second Printing, June 1998 Third Edition, March 2002 Copyright ©1990, ©1995, ©2002 DaimlerChrysler Corporation, Ford Motor Company, General Motors Corporation

[3] Statistical measurement analysis of automated test systems, Flynn, M. ; Naval Air Syst. Command, Patuxent River, Autotestcon, 2007 IEEE

[4] Propagating Uncertainty in Instrumentation Systems, B. D. Hall, IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 54, NO. 6, DECEMBER 2005

[5] Instrument certification as part of a modular test platform architecture, Manor, D. ; Geotest -Marvin Test Syst., Irvine ; Dewey, M, Autotestcon, 2007 IEEE

[6] Applied Statistics and Probability for Engineers, Fifth Edition, Douglas C. Montgomery, George C. Runger, ISBN–13: 978-0-470-05304-1

[7] Understanding Process Capability Indices, Stefan Steiner, Bovas Abraham and Jock MacKay, Institute for Improvement of Quality and Productivity, Department of Statistics and Actuarial Science, University of Waterloo