Comparison of Approaches to Gauge Repeatability and Reproducibility Analysis

Pavlína Mikulová[1,*], and *Jiří* Plura[2]

[1]VSB - Technical University of Ostrava, FMME, Department of Quality Management, Ostrava-Poruba, 17.listopadu 15, 70833 Czech Republic
[2]VSB - Technical University of Ostrava, FMME, Department of Quality Management, Ostrava-Poruba, 17.listopadu 15, 70833 Czech Republic

Abstract. The paper presents approaches to Gauge Repeatability and Reproducibility (GRR) analysis. The aim of the paper is to compare three methods of GRR analysis – A&R method, ANOVA and EMP III. The first part gives insight to use of Measurement System Analysis (MSA) since properly planned and executed measurement systems can provide a good basis for decision making in companies. One of the new approaches will be presented that is different from the others described by AIAG or VDA guide. Compared to known A&R method and ANOVA, EMP III interpret the results in more detail. Beyond the information about variability of measurement process, the approach provides other useful parameters as attenuation of process signals, change of detecting error shift and an ability to track process improvements. Wider use of the results is illustrated by an example of measuring the dimensions of fasteners.

1 Introduction

Decision-making based on facts is one of the fundamental principles of contemporary quality management. These facts are in most cases obtained from measurements. The proposed or already used measurement system requires attention, since an inadequate measurement system can provide distorted information that can lead to wrong decisions. The suitability of a measurement system is an important part of quality planning and quality improvement [1] and, for example, in automotive industry, it is strictly required [2].

An ineffective measurement system can allow bad parts to be accepted and good parts to be rejected. MSA could have prevented the problem. Thus, it is a collection of experiments and analyzes performed to evaluate a measurement system capability, performance and amount of uncertainty regarding the values measured [3, 4]. The knowledge of measurement system properties is important in assessing the conformity of products. The conformity/nonconformity of a product can be clearly confirmed only if the measurement uncertainty interval lies within or outside the tolerance [5]. Due to the tolerance limits given by the specifications, this means that the possibility of making

---

[*] Corresponding author: pavlina.mikulova.st@vsb.cz

a wrong decision about the part exists only if the measurement system error showed out of the limits. This leads to three distinct areas (see Fig. 1).
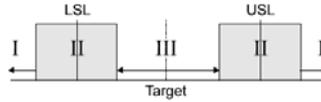

Fig. 1. Measurement uncertainty areas [3].

Area I shows bad parts will always be marked bad, area II shows a potential wrong decision can be made and area III represents good parts will always called good [3, 5]. There are many different factors that affect the measurement systems and influence the measured values including the causes related to the properties of measuring device itself, the measurement procedure, the operator performing measurements with a measured object, and a number of other factors [6]. The actions of these causes can be detected by analyzing the properties. The most used analysis of measurement system, which makes possible to evaluate repeatability and reproducibility of measurement system is GRR analysis.

2 Analysis of repeatability and reproducibility

Generally, it is not possible to ensure changeless measurement conditions in realization of the measurement. The real measurement conditions usually change, while it is mostly caused by the change of operator performing the measurement. Two major sources of measurement system variability are repeatability and reproducibility [7].

2.1 Methods of GRR analysis

Comprehensive GRR study can be performed by using one of the following techniques:
1. A&R method (Average and Range method)
2. ANOVA (Analysis of Variance)
3. EMP (Evaluating the Measurement Process).

A&R method uses subgroups of data to determine the variance due to the various sources of variation. It does not take into account the operator-part interaction. ANOVA is a technique that examines what sources of variation have a significant impact on the results. This approach actually adds another source of variation to the mix and it is able to identify the operator-part interactions [3]. Recently, the attention is focused on the third method known as EMP III approach characterized below.

2.2 Introduction of EMP III approach

D. J. Wheeler, an expert in statistical process control, introduces a new approach and points out the statistical errors (the use of less appropriate criteria) of AIAG approach, the most significant of which is the use of ratios based on the standard deviation. Wheeler introduces a number of notable differences considering: metrics based on the variance instead of standard deviation; Intraclass Correlation Coefficient (ICC); probable error providing an absolute characterization of measurement error; process signal attenuation; monitor classes classifying the measurement system [8]. A proper understanding of probable error and intraclass correlation coefficient will allow to quantify the uncertainty in the measurement. D. J. Wheeler classifies the measurement system as a first, second, third or fourth class monitor based on the ICC. It is the ratio of the part variance to the total variance that also can be simply calculated from the results of A&R or ANOVA method:

$$ICC=(PV)^2/(TV)^2 \tag{1}$$

In case the samples represent the production range and part variance can be determined, it is worth to calculate ICC. This ratio describes that proportion of the variation in the product measurements that is directly attributable to the variation in the product stream. The EMP results are interpreted by four categories which give insight into the measurement system (see Table 1): Firstly, how it can reduce the strength of a signal (out of control points) on a control chart. Secondly, give the chance of the measurement system detecting a large shift. Thirdly, are able to track process improvements. D. J. Wheeler claims those criteria give a good understanding of relative usefulness of the measurement system telling that the first, second and third class monitors include acceptable measurement systems [9].

Table 1. Four class monitors according to D. J. Wheeler [9].

ICC	Type of class monitor	Attenuation of process signal	Chance of detecting ±3σ std. error shift	Ability to track process improvements
0.8 to 1.0	First class	Less than 10%	More than 99% with rule 1	Up to Cp80
0.5 to 0.8	Second class	From 10% to 30%	More than 88% with rule 1	Up to Cp50
0.2 to 0.5	Third class	From 30% to 55%	More than 91% with rules 1, 2, 3, 4 (see [9])	Up toCp20
0.0 to 0.2	Fourth class	More than 55%	Rapidly vanishing	Unable to track

In Table 2, the acceptance criteria of measurement system are stated. Those are criteria prescribed by AIAG and compared with new evaluative criterion introduced by D.J. Wheeler, ICC. As seen, related to the acceptance of measurement system, the value of ICC corresponding to %GRR is allowed to be in the small range (0.91-1) if considering the AIAG criteria. However, observing the Table 1 Wheeler's criteria are significantly less strict (ICCϵ(0,5-1)).

Table 2. Comparison of evaluative crtiteria.

%GRR	Corresponding ICC	Classification
0 to 10	0.99 to 1	Acceptable
10 to 30	0.91 to 0.99	Conditionally acceptable
> 30	<0.91	Unacceptable

2.3 Application of EMP III approach

Measurement system which is used to measure the height of the nuts is subjected to analysis. Measured values of ten nuts (each was measured twice by an operator), representing the entire production range obtained by three operators using the digital calliper (data from [1]). The procedure of EMP III approach is identical with the one of A&R method with the difference of calculating the values considering the variance instead of the standard deviation. The procedure of the approach and equations for the all calculations are summarized in [9].

2.3.1 Characterizing the relative usefulness of measurement system

The individual proportions of variance were calculated: proportion of repeatability, proportion of reproducibility, proportion of repeatability and reproducibility, proportion of part variation to total variance which is, at the same, the value of intraclass correlation coefficient (see Equation 1). After multiplying the ICC by hundreds, it simply represents

percentage of total variance caused by part variance. The important data for calculating ICC were obtained: firstly, the height of ten nuts were measured twice by each operator (PV is calculated), secondly, single standard sample of nut is measured twice by each operator (TV caused by the measurement system can be obtained).

$$ICC=(PV)^2/(TV)^2 =[(TV)^2-(GRR)^2]/(TV)^2=1-(GRR)^2/(TV)^2 \qquad (2)$$

The value of ICC tells the following: if only the random causes of variability affect the process, %ICC is the variability between measured parts and, on the contrary, the %(1-ICC) is the variability of measurement system. According to Wheeler's criteria (see Table 1) and from the value of ICC=0.968 (see Table 3), our measurement system belongs into first class monitor. As seen in Table 3, the the results obtained by A&R method and EMP III are identical and very similar to ANOVA results. While according to AIAG criteria measurement system is classified as conditionaly acceptable, in accordance to EMP III approach the measurement system belongs into first class monitor.

Table 3. Comparison of results obtained by three methods of GRR analysis.

Source	EV	AV	GRR	PV	TV	
A&R method						**ndc**
Standard deviation	0.02511	0.02151	0.03307	0.18244	0.18542	
Variance	0.00063	0.00046	0.00109	0.03329	0.03438	7
% of total std.	13.54	11.60	17.83	98.40	100	
ANOVA						**ndc**
Standard deviation	0.02443	0.02102	0.03223	0.18589	0.18866	
Variance	0.00060	0.00044	0.00104	0.03455	0.03559	8
% of total std.	12.95	11.14	17.08	98.53	100	
EMP III						**ICC**
Standard deviation	0.02511	0.02151	0.03307	0.18244	0.18542	
Variance	0.00063	0.00046	0.00109	0.03329	0.03438	0.968
% of total variance	1.83	1.35	3.18	96.82	100	

2.3.2 Interpretation of results

Characterization of measurement system use by classification in the first, second, third or fourth class monitors: after classifying the measurement system into one of the four categories, the evaluation of measurement system is assessed from several aspects:

a)   attenuation of process signals: any signal of change in production process will be reduced (not fully recorded in control chart). This attenuation is characterized by:

$$\textit{Attenuation of process changes signals} = 1 - \sqrt{ICC} \qquad (3)$$

Hence, this attenuation means how many percentages of process change signals will be reduced in control chart. In our case, these signals are reduced by 1.6%.

b)   Chance of detecting ±3σ std. error shift: the chance of detecting a significant changes during process in control chart for the elected position (control chart for individual values or averages) by ±3σ is more than 99% (considering the first class monitor) using rule 1 (Western Electric rules apply if 10 samples are measured).

c)   Ability of measurement system to track process improvements that means to reduce production variability up to the corresponding C_p level. Recognizing significant process changes is very important because if this does not happen, there is a problem concerning the control chart detection of incorrect measurements. Either the control chart should have

detected incorrect measurement, but did not do so, or the measurement was correct, but control chart identified it as an incorrect one. After calculating value of C_p80 (see the equation in [9]), we have got value 4.75. The value tells reduction in variability of the process, whose C_p index would prove to be more than 4.75, also indicates as for the assessment of measurement system, that the value of intraclass correlation coefficient will be less than 0.80. It is proven by moving the measurement system from one class monitor downwards. Capability of measurement system to capture an improvement of the process reflects how great an improving or worsening the process have to be.

2.4 Comparison of approaches to GRR analysis for simulated data

For a better comparison and understanding the differences between EMP III and AIAG approaches, we did a simulation of the data where following situation occurred – a measurement system is unacceptable as the results of A&R method and ANOVA, but acceptable by Wheeler's method.

Table 4. Comparison of three approaches to GRR analysis for simulated data

Source	EV	AV	GRR	PV	TV	
A&R						**ndc**
Standard deviation	0.08242	0.05919	0.10147	0.18244	0.20876	
Variance	0.00679	0.00350	0.01030	0.03329	0.04358	2
% of total std.	39.48	28.35	48.61	87.39	100	
ANOVA						**ndc**
Standard deviation	0.073137	0.057764	0.093197	0.180378	0.203032	
Variance	0.005349	0.0033366	0.0086857	0.0325363	0.041222	2
% of total std.	36.02	28.45	45.9	88.84	100	
EMP III						**ICC**
Standard deviation	0.08242	0.05919	0.10147	0.18244	0.20876	
Variance	0.00679	0.00350	0.01030	0.03329	0.04358	0.76
% of total variance	15.59	8.04	23.62	76.38	100	

After the data simulation, A&R method and ANOVA evaluated the measurement system (see Table 3) as unacceptable for next using. In compliance with EMP III study, the measurement system belongs to the second class monitor (due to ICC value equal to 0.76). For the illustration of excessively high variability of measurement system (see Fig. 2), both situations are compared (with original and simulated data), where the difference of variability can be observed in Fig. 2.
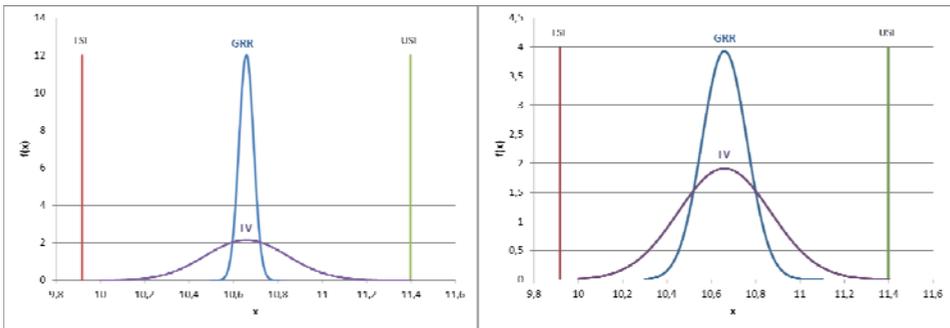

Fig. 2. Comparison of variability of measurement system (GRR) and total variability (TV) for a) original data, b) simulated data

These figures include the calculated tolerance limits that correspond to the centred process with C_{pk} = 1.33. In particular, Figure 3 shows that the measurement uncertainty interval for simulated data is very wide and creates almost half (48.61%) of the production range width. Related to the tolerance width, it means 36.5%. Due to the fact we have to calculate the uncertainty interval at each of the tolerance limits, the given measurement system would only be able to evaluate the parts as conforming if the measured value would lie in the area constituting 27% of tolerance width around the tolerance centre. It means for product conformity assessment this measurement system cannot be considered as acceptable.

3 Conclusion

Due to the above results of acceptance or rejection the measurement systems obtained by the three methods of GRR analysis, we have come to the following conclusion. Even though, methodology recommended by AIAG strictly refuses measurement system with %GRR = 48.61% and ndc = 2 (see Table 3), EMP III allows to classify the measurement system to the second class monitor, i.e. accepts it with any doubts. According to Wheeler this measurement system is still able to monitor the process changes (to detect the out of control points in control chart and to indicate process improvement), i.e. it is applicable only for cases when measurement system is used for process monitoring. For the measurement systems, which are used for decision making about products conformity, EMP III criteria cannot be applied.

References

1. J. Plura, Quality Planning and Continuous Quality Improvement, **244** (2001)

2. AIAG, *Production Part Approval Processs (PPAP)*, (Chrysler Gp. LLC, Ford Motor Company, General Motors Corp., 2006)

3. MSA Work Gp., *Measurement Systems analysis*, *Reference Manual fourth edition*, (2010)

4. C. D. Montgomery, *Introduction to Statistical Quality Control*, 754 (2009)

5. ISO 10576-1, *Statistical methods - Guidelines for the evaluation of conformity with specified requirements - Part 1: General principles*, (2003)

6. J. Plura, P. Klaput, *The Possibilities of Confidence Improvement of Measurement Systems Analyses Results*, *20th Anniversary International Conference on Metallurgy and Materials "METAL 2011"*, (Brno, Czech Republic, 2011)

7. J. Plura, *Quality Planning II,* 172 (2012)

8. Miner. *Intro to MSA of Continuous Data-Part 7* (Quality Forum Online, 2016)

9. D. J. Wheeler, *EMP III: Evaluating the Measurement Process & Using Imperfect Data*, 316 (2006)