

DETAILED ANALYSIS OF GRR STUDY RESULTS AND THEIR VISUALIZATION

doi: 10.2478/cqpi-2019-0065

Date of submission of the article to the Editor: 11/105/2019

Date of acceptance of the article by the Editor: 30/05/2019

Pavλίna Mikulová¹ - *orcid id: 0000-0002-2035-029X*

Jiří Plura¹ - *orcid id: 0000-0002-4739-0208*

¹VSB - Technical University of Ostrava, Faculty of Materials Science and Technology, Department of Quality Management, **Czech Republic**

Abstract: The paper presents the possibilities of analyzing the measurement system repeatability and reproducibility (GRR) analysis results in more detail and their visualization. Based on real data, the influence of appraisers' behavior on GRR analysis results is evaluated. The measured data obtained in the GRR study by three operators are analyzed in more detail for cases involving only two operators. Comparison of the behavior of individual operators is performed using Gaussian curves, which allow to evaluate graphically the repeatability of individual operators and the reproducibility of measurement system. This approach makes possible to visualize the GRR analysis results with regard to behavior of individual operators. Obtained results make possible to improve the interpretation of GRR analysis results and are useful for measurement system improvement.

Keywords: repeatability, reproducibility, factors, appraiser, visualization

1. INTRODUCTION

The benefit of using data-based quality management methods is largely dependent on quality of measured data. These are used to determine whether and how to control the production process, hence their high quality is necessary. With the current potential and rapid progress in automation, it is desirable to prove the suitability of the measurement system and the credibility of its results. The quality of the measured data is defined by the statistical properties of multiple measurements obtained from a measurement system operating under stable conditions. One of the most common causes of low-quality data is too much variation. The statistical property such as stability representing the change in bias over time is necessary to meet in order to have a stable measurement process (process in statistical control with respect to position). Consistency represents the degree of repeatability change over time (the consistent measurement process is in statistical control with respect to variability). Another property to be analyzed is the measurement system repeatability and reproducibility. Repeatability is the variation in successive trials under fixed and defined conditions of measurement - changeless parts, instrument, method, appraiser, environment, etc.

Reproducibility is the variation in the averages of the measurements made by different appraisers using the same gage when measuring a characteristic on one part (MSA Work Group, 2010).

The aim of the paper is to introduce the way how to visualize GRR analysis results in order to improve the measurement system. The visualization will be demonstrated by Gaussian curves which allow presenting the variation due to position and variability at once what standardly used tools do not provide and MSA manual does not use.

2. GAGE REPEATABILITY AND REPRODUCIBILITY ANALYSIS (GRR)

In real measurement, the conditions are usually changeable. Most often it is a change of the appraiser carrying out the measurement. In this situation the GRR analysis is used whose purpose is to determine if the measurement system variability is sufficiently low in relation to the variability of monitored process (Burdick et al., 2003).

2.1. Methods used for GRR analysis

The most commonly used methods for GRR analysis are the Average and Range (A&R) method, analysis of variance (ANOVA) and Evaluating the Measurement Process (EMP). The A&R method is an approach which provides an estimate of both repeatability and reproducibility for a measurement system. It allows studying the variation between the measured parts and variation between the appraisers. However, variation due to the interaction between the appraiser and the part is not accounted for this method. The MSA manual prefers ANOVA as it can provide a more exact estimation of variance and extracts the interaction between part and appraiser from the experimental data (MSA Work Group, 2010). The approach called EMP corresponds to the A&R method. Wheeler (Wheeler, 2006) introduces the metrics based on variance instead of the standard deviation which are used as the criteria for evaluating the measurement system suitability. Moreover, there are additional criteria used for evaluating the measurement process (Mikulová and Plura, 2018).

2.2. Assumptions of proper GRR analysis application

Although, the MSA methodology deals with verifying the assumptions inadequately, it is necessary to verify the stability and consistency of measurement system, the stability of measurement process in terms of variation caused by trials, and data normality. Analysis of measurement system stability and consistency should precede GRR analysis. In order to verify the stability and consistency of the measurement system, it is necessary to obtain data from trials of the changeless parts at appropriately elected time intervals. As part of GRR analysis the stability of measurement system with respect to variation caused by trials must be verified where the range control chart is used as standard to detect unstable measurement conditions. Further for most measurement processes, it is assumed that the distribution of measured values corresponds to the normal distribution (Montgomery, 2009). In case of GRR analysis data normality is important assumption for correct application A&R and ANOVA methods.

2.3. Graphical tools used within GRR analysis

As mentioned, one of the assumptions and first steps of GRR analysis is to verify if the measurement process is in statistical control in terms of variability of repeated

measurements performing by individual appraisers. For such evaluation the range control chart is used (Fig. 1). All the range values plotted in the chart should be within the control limits. If it is not so, the measurement system should not be evaluated before removing nonrandom variability causes and collection of new data.

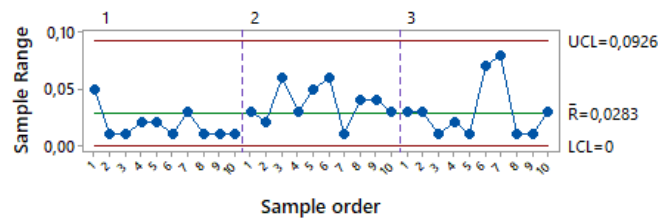


Fig. 1. R control chart

The assumption that the measurement system is appropriate for assessing the variation between the measured parts can be evaluated by control chart for averages of repeated measurements of individual product samples taken by individual appraisers (Fig. 2). This compares the variation within repeated measurements of the same parts with the variation within the production process (only in the case when measured samples cover entire production range).

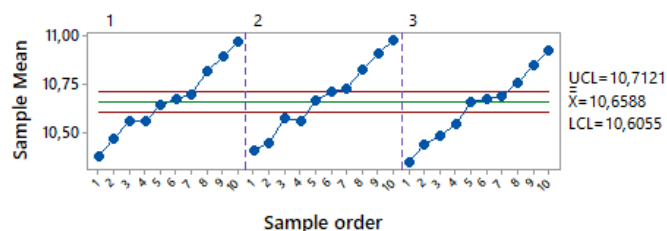


Fig. 2. \bar{X} control chart

Another standardly used graphical tool of GRR analysis is the diagram of interactions among the appraisers and the measured parts. Although, interactions are evaluated numerically within the framework of GRR analysis using ANOVA, the occurrence of these interactions can be estimated using this diagram (Fig. 3). Such diagram allows comparing averages of single samples measured by individual appraisers much clearer than Xbar chart.

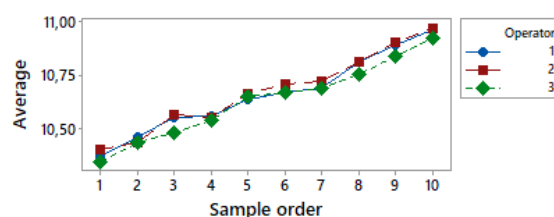


Fig. 3. Diagram of interactions among the appraiser and measured parts

Many other useful diagrams which allow analyzing the measurement system deeper were proposed by Klapat and Plura (Klapat and Plura, 2011). Stated diagrams are considered to be clear-read due to the samples displayed in the chart are ordered by the value of measured characteristic instead of the random order. This allows assessing the measurement system uniformity also and gives the clarity of diagrams.

Graphical tools of GRR analysis should be processed before numerical analysis, because they can identify various suspicious measurements and their causes. Then GRR analysis can continue with evaluating the repeatability (EV), reproducibility (AV), combined repeatability and reproducibility (GRR) and with the evaluation of variation between the measured parts (PV). On the basis of these variability sources the percentages of total variation (TV) are calculated (MSA Work Group, 2010), further, the criterion interdependent on percentage of GRR, number of distinct categories (ndc), is evaluated (García, 2013).

3. FACTORS AFFECTING GRR ANALYSIS RESULTS

GRR analysis results can be affected by various factors which can be divided into two groups (Plura and Klaput, 2011):

- Factors affecting the values of GRR analysis results,
- Factors affecting the confidence of GRR analysis results.

Significant factors affecting the GRR analysis results are the coverage of production range by measured samples, the way of total variation expressing, meeting the measurement conditions, selection of appraisers etc. In the cases when total variation is calculated by the PV and GRR, selected samples need to cover the entire production range of the process. If this is not fulfilled, the results of GRR analysis may be significantly distorted. It is caused by a decrease of the variation between the measured parts, and thus, by the decrease of total variation.

Results of GRR analysis can be considerably influenced by selection of appraisers involved in measurement system.

4. IMPACT OF THE APPRAISERS ON GRR ANALYSIS RESULTS

Based on real data the influence of appraiser behavior on GRR analysis results has been evaluated. The measurement system that is used for measuring the height of the nuts was used for analysis. The values of ten nuts, which represented entire production range, were measured by each of three appraisers using the digital caliper. The measured values are given in Table 1.

Table 1
Measured data of nuts height (in mm)

Operator	Trials	Measured sample									
		1	2	3	4	5	6	7	8	9	10
1	1	10,96	10,90	10,67	10,35	10,71	10,82	10,55	10,65	10,46	10,55
	2	10,97	10,89	10,68	10,40	10,68	10,81	10,56	10,63	10,47	10,57
2	1	10,99	10,89	10,68	10,42	10,73	10,80	10,54	10,64	10,45	10,57
	2	10,96	10,93	10,74	10,39	10,72	10,84	10,60	10,69	10,43	10,54
3	1	10,94	10,85	10,71	10,36	10,73	10,75	10,48	10,65	10,42	10,55
	2	10,91	10,84	10,64	10,33	10,65	10,76	10,49	10,66	10,45	10,53

Source: (Plura, 2001).

The case of three operators performing the measurements is given, each measures ten samples twice. Standardly used graphical tools of GRR analysis of this measurement system are presented in Fig. 1, Fig. 2 and Fig. 3. According to recommendation of

Klaput and Plura (Klaput and Plura, 2011) samples on the x axes are ordered from lowest to highest value of height.

On the basis of real data, the similar results were reached for the %GRR using the A&R method and ANOVA. The variability caused by interactions between operators and parts was evaluated as statistically insignificant and the measurement system has been assessed as conditionally acceptable on the basis of %GRR value equal to 17.83% (17.08%) and ndc equal to 7 (8) (see Table 2 – part Operator 1, 2, 3).

Table 2

Resultant values of GRR analysis in case of operators 1, 2, 3 and operators 2, 3

Operators	1, 2, 3				2, 3			
Methods	A&R		ANOVA		A&R		ANOVA	
EV	0,0251	13.54%	0,0244	12.95%	0,0294	15,78%	0,0275	14.49%
AV	0,0215	11.60%	0,0210	11.14%	0,0293	15.81%	0,0294	15.49%
GRR	0,0331	17.83%	0,0322	17.08%	0,0414	22.33%	0,0403	21.21%
PV	0,1824	98,40%	0,1859	98,53%	0,1809	97,47%	0,1856	97,73%
TV	0,1854	100%	0,1887	100%	0,1856	100%	0,1899	100%
ndc	7		8		6		6	

In the case when only operators 2 and 3 are involved in measurement system, the increasing in EV, AV and GRR can be seen. This is due to the higher average variation of trials, thus, the higher repeatability percentage (%EV), and due to the higher estimate of the reproducibility standard deviation. However, the resultant value of TV remains similar in both cases. This is just because TV takes into account both GRR value and PV value where, even though, GRR value increases, but on the other hand, the PV value slightly decreases.

The Table 2 shows the results are similar according to %GRR from used methods' point of view. The measurement system is assessed as conditionally acceptable on the basis of the conditions – %GRR lower than 30% and ndc value higher than 5.

Results of GRR analysis for other combinations of two operators are given in Table 3. These resultant values of measurements taken by two operators (including operator 1 – the one reaching the mean of all parts measurements) shows the decrease of %GRR compared to Table 2. If ANOVA is used, the results slightly vary what is the cause of different method used for reproducibility evaluation.

Table 3

Resultant values of GRR analysis in case of operators 1, 2 and operators 1, 3

Operators	1, 2				1, 3			
Methods	A&R		ANOVA		A&R		ANOVA	
EV	0,0240	13,05%	0,0216	11,41%	0,0210	11,26%	0,0238	12,74%
AV	0,0079	4,28%	0,0082	4,36%	0,0200	10,72%	0,0198	10,59%
GRR	0,0253	13,73%	0,0231	12,21%	0,0290	15,54%	0,0310	16,57%
PV	0,1824	99,05%	0,1877	99,25%	0,1840	98,78%	0,1843	98,63%
TV	0,1842	100%	0,1891	100%	0,1863	100%	0,1869	100%
ndc	10		11		8		8	

On the basis of these result, it is concluded that the best results of GRR analysis can be obtained by measurement system consisting of operator 1 and 2 and can be

recommended the training in measurement for operator 3. Of course, also other actions for measurement system improvement are needed.

4.1. Presentation of GRR results by Gaussian curves

One of the important parts in evaluating GRR analysis is to visualize its results, which can serve an useful way of results interpretation. Within the MSA methodology, the GRR study evaluates only the average repeatability and does not evaluate the repeatability reached by individual appraisers. Moreover, it is not clear how individual appraisers contribute to this variability when evaluating reproducibility. It is desirable to visualize the results of combined repeatability and reproducibility. This will be presented on the real case whose numerical results are stated in the previous section and will be visualized by using the Gaussian curves. The assumption which must be met before using Gaussian curve is that repeated measurements of quality characteristic have the normal distribution. This assumption is met for given data. Attention is focused on the way how to construct Gaussian curves, which can be original way of visualizing GRR results. The means correspond to averages of all measurements taken by each operator and the standard deviation is estimated by pooled standard deviation of repeated measurements of individual operator.

This might be an appropriate way of visualization for the cases with no interactions. Nevertheless, the variability induced by the interactions assigns to reproducibility, it does not reflect in the variability between averages of measurements.

Starting with the following situation – measurements carried out by three operators is illustrated by Gaussian curves in Fig. 4.

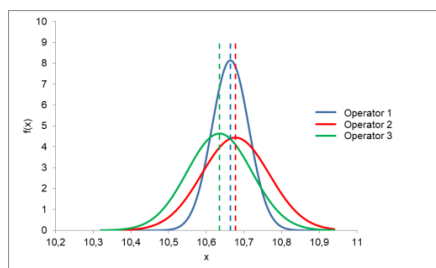


Fig. 4. Gaussian curves of GRR results in case of operators 1, 2, 3

From the Fig. 4 it is obvious the operator 1 has reached the lowest variability of repeated measurement. The variability reached by operator 2 and 3 gets worse and is mutually comparable. Regarding the variability of the means of measurements taken by individual operators, the averages of measurements taken by operator 1 and 2 are slightly similar. However, average of operator 3's measurement significantly differs. Hence, operator 2 measured on average the sort of higher values than operator 1, on the contrary, operator 3 measured considerably lower values than operator 1. The variation can be read due to position as well as variability simultaneously.

Moving to the cases of measurements performed by two operators where one of those was omitted. At first, Fig. 5 shows the measurements of operator 2 and 3. Operator 2 has reached the highest value of average of all measurements and, on the contrary, operator 3 has reached the lowest value of average of all measurements.

This combination of operators leads to the worst %GRR result, mainly due to the high sort of repeatability and reproducibility as well. The high repeatability value is related to

the high variability of repeated measurements of these operators. The high reproducibility value is related to the large difference between averages of measurements taken by these operators.

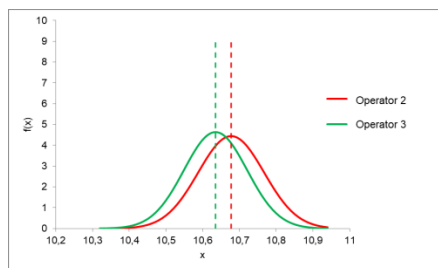


Fig. 5. Gaussian curves of GRR results in case of operators 2, 3

Another combination of measurements taken by two remaining operators and its influence is shown in Figure 6 and 7. Figure 6 shows the measurements taken by operator 1 and 2, which create mean and upper limit average of all parts measurement.

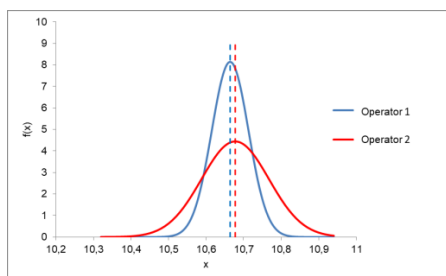


Fig. 6. Gaussian curves of GRR results in case of operators 1, 2

This combination of operators leads to the best %GRR results, mainly due to the low reproducibility (i.e. similar values of all measurements averages). Owing to low GRR value, the ndc value is propitious too.

Figure 7 shows the measurements taken by operators 1 and 3. This combination of operators leads to worse results compared to the previous case stated in Fig. 6, mostly due to higher reproducibility (greater difference between the values of all measurements averages).

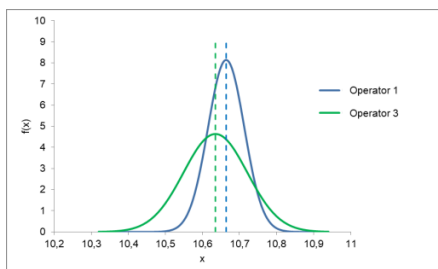


Fig. 7. Gaussian curves of GRR results in case of operators 1, 3

Comparing the results obtained by each of three operators with the cases, where one of those was skipped, leads to the following decision – need to reach the low variability of repeated measurements and low sort of reproducibility (i.e. values of all measurements averages are alike, the differences between the averages are negligible).

5. CONCLUSION

Detailed analysis of data collected for GRR analysis and visualization of results can provide valuable tools for better understanding of measurement system and its improvement. Detailed analysis is focused on the changes of repeatability and reproducibility results in the cases of including only some of the operators into measurement system. This analysis provides useful information on finding opportunities for measurement system improvement. Very important is also visualization of GRR analysis results. Standardly used graphical tools do not provide sufficient information about behavior of individual operators. The way how to visualize the results of GRR analysis has been proposed in this paper. It offers an appropriate way regarding the illustration of the GRR results what MSA manual does not use. The visualization of the results is performed by constructing the Gaussian curves which clearly presents the variation due to both position and variability simultaneously.

ACKNOWLEDGEMENTS

The work was supported by the specific university research of the Ministry of Education, Youth and Sports of the Czech Republic No. SP2019/62 and from ERDF „A Research Platform focused on Industry 4.0 and Robotics in Ostrava“, No. CZ.02.1.01/0.0/0.0/17_049/0008425. The paper has been elaborated in the framework of the grant program „Support for Science and Research in the Moravia-Silesia Region 2018“ (RRC/10/2018), financed from its budget.

REFERENCES

- Burdick, R.K., Borror, C.M., Montgomery, D.C., 2003. *A review of methods for measurement systems capability analysis*, Journal of Quality Technology, 35, 342–354. DOI: 10.1080/00224065.2003.11980232
- García, A.C., Ríó, A.G., 2013. *Number of distinct data categories and gage repeatability and reproducibility, A double (but single) requirement*. Measurement, 46 (8), 2514–2518, DOI: 10.1016/j.measurement.2013.04.065
- Klaput, P., Plura, J., 2011. *The importance of graphical tools for measurement systems analysis*, Proceedings of QMOD Conference on Quality and Service Sciences 2011, Pamplona: Universidad Navarra, 1010-1028.
- Mikulová, P., Plura, J., 2018. *Comparison of approaches to gauge repeatability and reproducibility analysis*, 12th International Conference Quality Production Improvement QPI 2018, Zaborze, 183-189.
- Montgomery, D.C., 2009. *Introduction to statistical quality control*. 6th ed. Hoboken, N. J.: Wiley, New Jersey.
- MSA Work Group, 2010. *Measurement System Analysis, Reference Manual*, fourth ed. Chrysler Group LLC.
- Plura, J., 2001. *Quality Planning and Continuous Quality Improvement (In Czech)*, first ed. Computer Press, Prague, 244 pp., ISBN 80-7226-543-1.
- Plura, J., Klaput P., 2011. *The possibilities of confidence improvement of measurement systems analyses results*, 21th International Conference on metallurgy and Materials - METAL, Brno, Czech republic.
- Wheeler, D. J., 2006. *EMP III: Evaluating the Measurement Process & Using Imperfect Data*, Third ed. SPC Press, Knoxville, Tennessee.