Int. J. Metrol. Qual. Eng. **4**, 23–28 (2013)

DOI: 10.1051/ijmqe/2012033

Evaluation of the R&R and the compatibility index for non-independent measurements

A. Garcia-Benadí*, E. Molino-Minero-Re, J. del Río-Fernández, and A. Mànuel-Lázaro

SARTI Research Group. Electronics Dept. UniversitatPolitècnica de Catalunya (UPC), Rambla Exposició 24, 08800, Vilanova i la Geltrú, Barcelona, Spain

Received: 22 October 2012 / Accepted: 22 November 2012

Abstract. This paper describes the methodology of the compatibility criteria E_n and the methodology of repeatability and reproducibility (R&R) throught the average and range method. With this paper we will use the methodology R&R for the evaluation of the compatibility criteria between the staff of the laboratories, where independent measurements are not insured.

Keywords: Compatibility; uncertainty; R&R; Monte Carlo's Method

1 Introduction

Currently in an increasingly competitive world, the differential values between companies or institutions are becoming more critical to the client, who has to choose one or the other. One important factor is quality. In this paper, a requirement found in the standard ISO 17025 [1] applicable for calibration or testing laboratories is presented. As stated in the ISO 17025, the centers must perform various tasks to ensure the quality of their measurements. Among other tests, there are two which are very important: the intercomparison exercises between laboratories and the repetitions between the staff of the same laboratory. These two tests are precisely the factor to be analyzed in this article.

The intercomparison exercises between laboratories, as stated in the ISO 17043 [2], describes different types of studies, being one of the most used the compatibility index, shown in (1).

$$E_n = \frac{|A - B|}{\sqrt{U_A^2 + U_B^2}} \tag{1}$$

where A and B are the average corrections, and U_A and U_B are the expanded uncertainty of A and B. The criteria of evaluation of the compatibility index for independent laboratories is if $E_n < 1$ the laboratories are compatible.

This index is perfectly valid as long as we can say that the measurements have always been independent of each other, a claim that can be valid if the measurements are performed in different laboratories, and also if the laboratories do not know the values found in other laboratories. The problem arises when the repeatability study is performed between workers in the same laboratory, since the independence of the measurements, although intended to be independent, can always be dependent because the measurements are done in the same technical facilities, and also because technicians can talk to each other about the tests. Changes must be made for this kind of cases (1). Given the non-independence of measurements, we obtain a new index E'_n, equation (2). This equation will be mathematically derived as the article develops.

$$E'_n = \frac{|A - B|}{\sqrt{U_A^2 + U_B^2 - 2U_{A,B}}} \tag{2}$$

where $U_{A,\,B}$ is the contribution to the uncertainty of the correlation between A and B.

The aim of this paper is to show that the methodology of repeatability and reproducibility (R&R) for the repeatability and reproducibility evaluation with technicians within the same laboratory is better than the use of compatibility index E_n. In Section 2 we show the compatibility index, E'_n, calculation with the contribution of the non-independence of measurements. In Section 3 we show the R&R method application for repeatability and reproducibility evaluation. In Section 4 we develop the uncertainty budget. In Section 5 we show the results for all cases and conclusions are presented in the last section.

2 Compatibility index (E'_n) calculation

To determine the correlation term, we will show the process from the beginning.

$$Y = \overline{X_1} - \overline{X_2}. \tag{3}$$

The numerator of (1) likes (3), and this fact is very important to start working with (3). Once we have defined the function, we proceed by calculating the associated uncertainty, which in this case we do by following the criteria, set by the GUM guide [3] and develop the Taylor series

* Correspondence: albert.garcia-benadi@upc.edu

Article published by EDP Sciences

$$u(Y)=\sqrt{\left(\frac{\partial Y}{\partial X_1}u_{X_1}\right)^2+\left(\frac{\partial Y}{\partial X_2}u_{X_2}\right)^2+2\frac{\partial Y}{\partial X_1}\frac{\partial Y}{\partial X_2}\rho(X_1,X_2)u_{X_1}u_{X_2}} \tag{5}$$

up to the second order, with the purpose of observing the correlation term. The result is shown in equation (4).

$$u(Y) = \sqrt{\left(\frac{\partial Y}{\partial X_1}u_{X_1}\right)^2+\left(\frac{\partial Y}{\partial X_2}u_{X_2}\right)^2+2\frac{\partial Y}{\partial X_1}\frac{\partial Y}{\partial X_2}u_{X_1,X_2}}. \tag{4}$$

Following the criteria of the GUM guide, we introduce the correlation coefficient, shown in equation (5).

See equation (5) above

where the correlation index ρ is defined. It can take values between $-1 \leqslant \rho \leqslant +1$. The value ρ can be calculated by the standard deviations of the measurements as indicated in the GUM guide, show in equation (6).

$$\rho(X_1,X_2)=\frac{s(X_1,X_2)}{s_{X_1}s_{X_2}} \tag{6}$$

where the value of $s(X_1,X_2)$ is calculated by equation (7).

$$s(X_1,X_2)=\frac{1}{n(n-1)}\sum(X_1-\overline{X_1})(X_2-\overline{X_2}). \tag{7}$$

Another option is to take the maximum value, +1, which is the worst case we can have because it implies a strong correlation. If ρ is equal to 0, it implies no correlation between the variables X_1 and X_2. Taking the hypothesis of the worst case, we obtain equation (8),

$$u(Y)=\sqrt{(1u_{X_1})^2+(-1u_{X_2})^2+2\times 1(-1)1u_{X_1}u_{X_2}}. \tag{8}$$

This leads to equation (9), which corresponds to the denominator of equation (2).

$$u(Y)=\sqrt{u_{X_1}^2+u_{X_2}^2-2u_{X_1}u_{X_2}}. \tag{9}$$

Through equation (9), we can calculate the combined uncertainty for cases where the measurements are correlated. With equations (3) and (9) we can compare between the function and its uncertainty value, where we want the result of the quotient to be less than or equal to 1. We see that this result is given in equation (2).

3 Application of the R&R method

For the applications of the R&R method, the procedure detailed in reference [4] has been followed.

First, a process where the R&R methodology is going to be used is defined: three technicians will perform six measurements and on every measurement three trials. From these we have three evaluations, six parts, and three trials. Measurements are generated automatically following expression (15).

Table 1. Results of the R&R test.

Parameter	Value		% of TV
EV	1.63		86
AV	0.00		0
GRR	1.63		86
PV	0.98		51
TV	1.90		0
		ndc	1

The values of the R&R study are represented through the *GRR* index, which are expressed in %. An example of an R&R method application can be shown in Table 1, where arbitrary values are used to explain the method variables.

The term *EV* evaluates the contribution of the repeatability and the term *AV* evaluates the reproducibility of the test.

It is considered that the appraisals are compatible as long as the *GRR* value is less than 30%, and the *ndc* ratio is greater than 5.

For the example in Table 1, we can see that GRR = 86% (>30%) and $ndc = 1$ (<5) that means that appraisals are not compatible, they can even be reproduced because of the value of *AV*.

4 Uncertainty budget

For every assessment, a typical contribution considers the repeatability, the resolution, and the master expanded uncertainty.

The repeatability contribution is $u_{s,\,APi}$ and the value is obtained with equations (10) and (11) for every appraisal, where $i = 1, 2, 3$.

$$\begin{aligned}u_{S,APi} &= \max\left(u_{S,APi;Pj}\right)\\ &= \max\left(u_{S,APi,P1};u_{S,APi,P2};u_{S,APi,P3};\right.\\ &\qquad \left.u_{S,APi,P4};u_{S,APi,P5};u_{S,APi,P6}\right)\end{aligned} \tag{10}$$

$$u_{S,APi,P1}=\sqrt{\frac{1}{3}\frac{\sum_{j=1}^{3}\left(x_{APi,j,P_1}-\overline{x_{Api,P_1}}\right)^2}{3-1}}. \tag{11}$$

Another contribution is the resolution of the device, which is shown in equation (12).

$$u_R=\sqrt{\frac{(Resolution)^2}{12}}. \tag{12}$$

The uncertainty budget from reference [3] has been simplified and the result is shown in equation (13).

$$u=\sqrt{u_S^2+u_R^2+u_P^2}. \tag{13}$$

Table 2. Results of the simulated experimental test with three appraisals showing the R&R results, the uncertainty budget and the evaluation of the compatibility index.

Appraisal	Parameter	P1	P2	P3	P4	P5	P6
AP1	N1	9.846	19.806	29.697	39.123	49.088	59.370
	N2	9.891	19.969	29.429	39.268	49.273	59.593
	N3	9.175	19.513	29.147	39.061	49.932	59.111
	Correction	0.363	0.238	0.576	0.849	0.569	0.642
	Range	0.716	0.456	0.550	0.207	0.844	0.482
	Std. deviation	0.401	0.231	0.275	0.106	0.444	0.241
AP2	N1	9.836	19.791	29.684	39.067	49.046	59.351
	N2	9.888	19.966	29.408	39.215	49.269	59.591
	N3	9.132	19.480	29.140	39.033	49.928	59.110
	Correction	0.382	0.254	0.589	0.895	0.586	0.649
	Range	0.756	0.486	0.544	0.182	0.882	0.481
	Std. deviation	0.422	0.246	0.272	0.097	0.459	0.241
AP3	N1	9.840	19.798	29.678	39.054	49.047	59.326
	N2	9.885	19.968	29.374	39.222	49.272	59.582
	N3	9.174	19.469	29.092	39.039	49.931	59.066
	Correction	0.367	0.255	0.619	0.895	0.584	0.675
	Range	0.711	0.499	0.586	0.183	0.884	0.516
	Std. deviation	0.398	0.254	0.293	0.102	0.459	0.258

And the expanded uncertainty of the measurement is calculated using (14), because the probability distribution is known, it is Gaussian.

$$U = 2.00\,u. \qquad (14)$$

5 Results

We analyze the results for various scenarios. In this work, we propose the structure order shown in (15) should be followed.

$$\begin{aligned} AP1 &= Random1 \\ AP2 &= AP1\left(1+\frac{Random2}{10}\right) \\ AP3 &= AP1\left(1+\frac{Random3}{10}\right) \end{aligned} \qquad (15)$$

where:

- The *Random1*, *Random2* and *Random3* are comprised from 0 to 1, with a rectangular distribution probability.
- The *Correction* is the real value minus the average value. The *Range* is the top value minus the lowest value, and the *Std. deviation* is the standard deviation calculated as (11) by square root of number of values, in our case three values.

We study three possible cases:

Case a: $E_n < 1$, $E'_n > 1$, $GRR > 30\%$.

We find that the compatibility index E_n is less than 1. Nevertheless, we know the relationship between operators and we can calculate E'_n obtaining values greater than 1. Likewise, we can identify that the study of R&R has given a result that shows that they are not compatible, because $GRR > 30\%$. The numerical values are shown in Tables 2–5.

Table 3. The R&R results of the simulated experimental test with three appraisals.

Parameter	Value		% of TV
EV	1.96		84
AV	0.00		0
GRR	1.96		84
PV	1.42		59
TV	2.42		0
		ndc	1

Table 4. The uncertainty budget of the simulated experimental test with three appraisals.

	AP1	AP2	AP3
u_S	0.256	0.265	0.265
u_R	0.0003	0.0003	0.0003
u_P	0.001	0.001	0.001
u	0.256	0.265	0.265
U	0.512	0.530	0.530

Table 5. The evaluation of the compatibility index of the simulated experimental test with three appraisals.

	E_n	E'_n
AP1-AP2	0.03	1.15
AP2-AP3	0.01	7.06
AP1-AP3	0.04	1.45

Case b: $E_n < 1$, $E'_n < 1$, $GRR < 30\%$.

This case indicates that the measurements are completely independent, which is not possible due the considerations presented in (15).

Case c: $E_n < 1$, $E'_n < 1$, $GRR > 30\%$.

This case indicates that the R&R method induces an error, see Tables 6–9. Making iterations we can find this case, as shown on Tables 6–9.

Table 6. Values found for case c.

Appraisal	Parameter	P1	P2	P3	P4	P5	P6
AP1	N1	9.030	19.262	29.211	39.237	49.064	59.751
	N2	9.758	19.639	29.103	39.820	49.505	59.853
	N3	9.127	19.832	29.751	39.043	49.916	59.412
	Correction	0.695	0.422	0.645	0.634	0.505	0.328
	Range	0.728	0.570	0.648	0.777	0.852	0.441
	Std. deviation	0.395	0.290	0.347	0.404	0.426	0.231
AP2	N1	9.000	19.259	29.135	39.206	48.978	59.751
	N2	9.741	19.631	29.013	39.814	49.489	59.850
	N3	9.077	19.831	29.739	38.983	49.908	59.358
	Correction	0.727	0.426	0.704	0.666	0.542	0.347
	Range	0.741	0.572	0.726	0.831	0.930	0.492
	Std. deviation	0.407	0.290	0.389	0.430	0.466	0.260
AP3	N1	9.030	19.198	29.193	39.221	49.001	59.743
	N2	9.753	19.633	29.027	39.819	49.467	59.852
	N3	9.047	19.824	29.736	38.951	49.910	59.400
	Correction	0.723	0.448	0.681	0.669	0.541	0.335
	Range	0.723	0.626	0.709	0.868	0.909	0.452
	Std. deviation	0.412	0.321	0.371	0.444	0.454	0.236

Table 7. The R&R results for the case c.

Parameter	Value	% of TV
EV	2.48	95
AV	0.00	0
GRR	2.48	95
PV	0.85	32
TV	2.62	0
	ndc	0

Table 8. The uncertainty budget for the case c.

	AP1	AP2	AP3
u_S	0.246	0.269	0.262
u_R	0.0003	0.0003	0.0003
u_P	0.001	0.001	0.001
u	0.246	0.269	0.262
U	0.492	0.538	0.525

Table 9. The evaluation of the compatibility for case c.

	E_n	E'_n
AP1-AP2	0.04	0.67
AP2-AP3	0.00	0.18
AP1-AP3	0.04	0.86

The analysis of this particular case can be done from the conditions expressed in (16).

$$E_n = \frac{\Delta}{\sqrt{U_i^2 + U_j^2}}$$
$$E_n^{'} = \frac{\Delta}{\sqrt{U_i^2 + U_j^2 - 2U_iU_j}}. \quad (16)$$

From these, we obtain the result shown in equation (17), with the initial assumption that $E'_n > E_n$ and that the squared value of E_n/E'_n is negligible for mathematical operations of addition or subtraction.

Table 10. Empirical data for the case AP1, AP2.

$E_n < 1,\ E'_n < 1,\ GRR > 30\%$			$E_n < 1,\ E'_n > 1,\ GRR > 30\%$		
u_{S1}	u_{S2}	u_{S2}/u_{S1}	u_{S1}	u_{S2}	u_{S2}/u_{S1}
0.234	0.249	1.06	0.184	0.188	1.02
0.257	0.276	1.07	0.222	0.228	1.03
0.216	0.226	1.05	0.194	0.197	1.02
0.310	0.332	1.07	0.210	0.218	1.04
0.214	0.232	1.08	0.252	0.258	1.02
0.262	0.282	1.08	0.196	0.198	1.01
0.245	0.261	1.07	0.256	0.270	1.05
0.250	0.268	1.07	0.189	0.190	1.01
0.234	0.255	1.09	0.210	0.216	1.03
0.280	0.296	1.06	0.198	0.201	1.02
	Average	1.07		Average	1.02

$$\frac{U_j}{U_i} = 1 \pm \sqrt{2}\left(\frac{E_n}{E'_n}\right). \quad (17)$$

Considering only the statistical contribution of the uncertainty, which is the main contribution, we perform an empirical study of the measured samples, show in Table 10.

With the data form Table 10 and equation (17), we obtain:

- $E_n < 1,\ E'_n < 1,\ GRR > 30\%$ the value of E_n/E'_n is 0.049;
- $E_n < 1,\ E'_n > 1,\ GRR > 30\%$ the value of E_n/E'_n is 0.014.

There is a limit value for the quotient u_{s2}/u_{s1}, of 1.05, which leads to a value of $E_n/E'_n = 0.035$.

From Table 10, we observe that for U_j/U_i lower than 1.05, the system detects that the variables have a certain level of correlation and this conditions the fact that E'_n is greater than 1, but for $U_j/\ U_i$ larger than 1.05, the variables show a small correlation and the calculus E'_n considers them as independent.

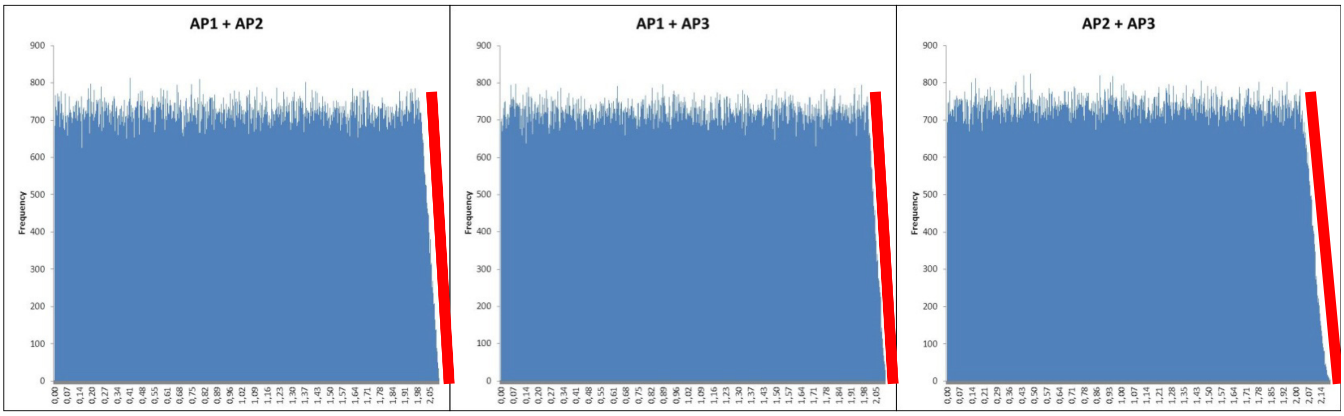

Fig. 1. Probability distributions for cases AP1 + AP2, AP1 + AP3 and AP2 + AP3, all with a denominator of 10.

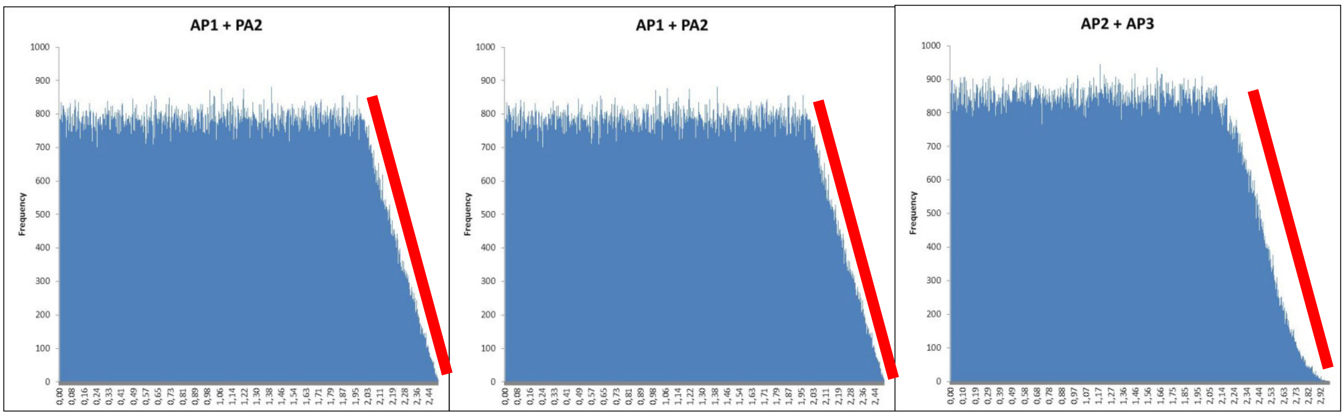

Fig. 2. Probability distributions for cases AP1 + AP2, AP1 + AP3 and AP2 + AP3, all with a denominator of 2.

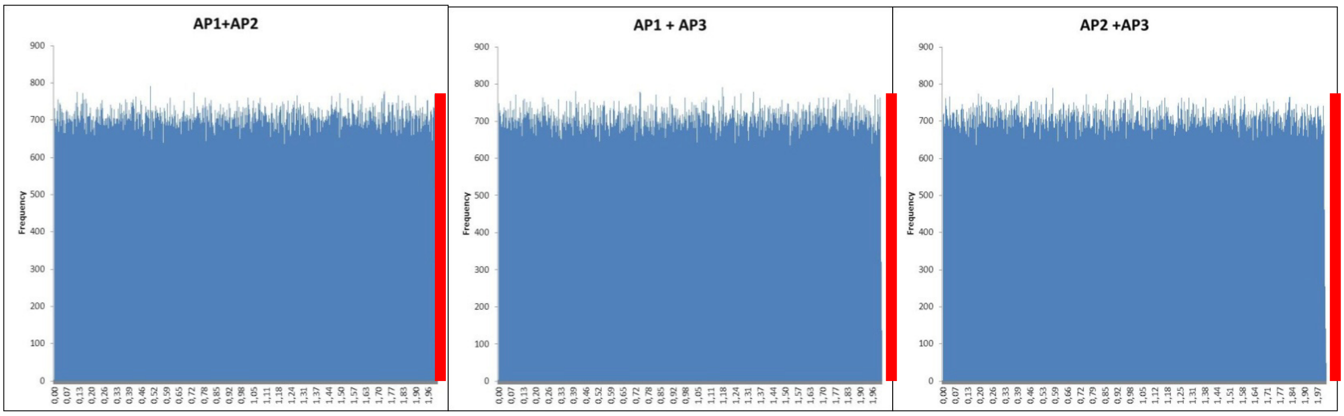

Fig. 3. Probability distributions for cases AP1 + AP2, AP1 + AP3 and AP2 + AP3, all with a denominator of 100.

In order to understand the reason of the duality between cases a and c better, we have to think about the correlation between the three variables set in (15). Each one of these variables has a rectangular probability distribution, and the term that links all variables is the quotient, which in the case of (15) has been set to 10. As this quotient increases the more delimited the probability distribution will be. Following the Monte Carlo method, 5×10^5 iterations are performed on the resulting function of the addition. The histogram is used to show the results for the representation of the probability distribution. In Figures 1–3, it is possible to observe the different frequency distributions obtained from the addition of AP2 with AP3, for a different denominator.

It is possible to observe graphically through the slope, marked in red, in Figure 2, that the resulting probability distribution is not rectangular. Likewise, with a quotient of 2, the relation between variables is of 50%. In Figure 3, we observe that the probability distribution is rectangular, and this is because the relations between variables are close to 100%, meaning that they are the same variables. In the case of Figure 1, we can see an intermediate relation, which can lead to a possible case c.

6 Conclusion

On this paper, we have shown a method to obtain R&R values and compare them with the compatibility criteria. The paper shows that we can use the R&R methodology because it presents better results than the compatibility with index E_n, since we have to evaluate the possible correlation between measurements. In the R&R method, the correlation is automatically detected. A particular case has been studied and the input variables are tied-up lineally, where the contribution to the uncertainty of the statistical term is the most important contribution. This fact implies that this exercise is not valid for cases where the contribution of the statistical term is comparable to the uncertainty of the reference or to the contribution of the scale range of the measurement instrument.

Future work could try to demonstrate that this methodology is valid for tied-up variables of other fields, for example tied-up to a triangular, Gaussian probability, etc.

References

1. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories (2005)
2. ISO/IEC 17043:2010 Conformity assessment – General requirements for proficiency testing (2010)
3. Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1), Evaluation of measurement data-Guide to expression of uncertainty in measurement (2008)
4. DaimlerChrysler Corporation, Ford Motor Company, General Motors Corporation (eds.), *Measurement System Analysis (MSA)*, 3rd edn. (2002)