IMEKO XVIII Word Congress
IV Congresso Brasileiro de Metrologia
17 a 22 de Setembro de 2006, Rio de Janeiro-RJ, Brasil

Methodology to evaluate calibrations: A case of study on the Interlaboratorial Comparison Program

Sousa, J.J.L., MSc
CHESF, Recife, Brasil, joel@chesf.gov.br

Leitão, L.T.S.
CHESF, Recife, Brasil, torres@chesf.gov.br

1. INTRODUCTION

The actual maximum error during the calibration of the PCI-Wh, beyond being directly related with the error and the uncertainty of the standards used in this calibration depends, as well, on the inherited uncertainty of the Calibration Network and also on the instability of the PCI-Wh itself.
A laboratory that owns a standard of energy that presents an uncertainty of 100µWh/Wh, in one determined point with power factor equal to 1, will be able to have its Better Capacity of Measurement modified for a high value when this standard is calibrated with another standard that has uncertainty of the same order.
Thus, although the error of the laboratory is small, the maximum error could be great in function of the uncertainty brought from Network to inside this laboratory. This can be observed in Table-1, what represents the participation of Chesf in the Interlaboratorial Comparison Program, during the period from 2000 to 2004. It fits to stand out that in the years of 2000 and 2001 we did not have tracked our standards with the RBC (Brazilian Calibration Network). Thus, in those years, the uncertainty was calculated considering only its class declared by the manufacturer, which is 100µWh/Wh (fp=1). Only after 2002 Chesf started to calibrate its standard in the RBC. Immediately, an increase of the uncertainty is evidenced, whose bigger parcel is inherited from the Network. On the other hand, a trend in the reduction of the error is observed, showing that now it is being corrected.

Chesf	error(µWh/Wh)	inc(µWh/Wh)	e(max)
2000	210,0	120,0	330,0
2001	200,0	120,0	320,0
2002	-20,0	180,0	200,0
2003	10,0	280,0	290,0
2004	-110,0	470,0	580,0

Table-1: Chesf data on PCI-Wh

The main objective of this Program of Interlaboratorial Comparison is to evaluate the results obtained in the unit of Wh, of the participants in relation to a reference, in our case, the Inmetro, as to:

- used procedures of measurement;
- metrology tracking of the standards;
- compatibility of the standards due to uncertainty levels;
- compatibility among the laboratories through the En pointer.

Key Words: volatileness (instability) - outliers - conformity – beta factors.

2. OBJECTIVE

The objective of this work is to consider a form to evaluate the results of calibrations of Wh performed by the laboratories that participate in the PCI-Wh Program using a methodology based on the linearization of the normal function distribution. This evaluation shows the behavior of each laboratory during the period of its participation in this Program.

3. PRELIMINARY CONSIDERATIONS

It has been taken as reference for this work the spread sheet of the calculations of errors and uncertainties of Standard-X: 120V, 0,5A, 60 Hz, p.f. 1, Tap 1A, relative to years from 2000 to 2004 of all the participants in the PCI-Wh Program. The variable is the maximum error [e (max)], that is, the module of the uncertainty added to the module of the error, taken year by year, in a sample size of five. The maximum errors will be treated as being a normal distribution. In Table-2 are the data already treated for the objective of this study. In this table Inmetro and Chesf are detached, standing out that the sequence of presentation of the companies in this table is randomly and does not correspond to the sequence of PCI-Wh report. Also is good to stand out that we only analyze the companies that had remained during all those years in the Program. For better visualization of the data it have opted to represent them in µWh/Wh.

	2000	2001	2002	2003	2004
company	e(max)	e(max)	e(max)	e(max)	e(max)
Inmetro-inicial	230,0	190,0	140,0	270,0	160,0
A	200,0	310,0	260,0	370,0	330,0
B	770,0	180,0	180,0	480,0	300,0
chesf	330,0	320,0	200,0	290,0	580,0
D	580,0	340,0	390,0	400,0	360,0
E	330,0	200,0	180,0	180,0	210,0
F	140,0	140,0	150,0	190,0	210,0
G	330,0	400,0	420,0	430,0	420,0
H	460,0	400,0	480,0	990,0	580,0
I	580,0	610,0	710,0	720,0	640,0
J	500,0	200,0	280,0	210,0	330,0
Inmetro-final	180,0	110,0	190,0	230,0	150,0

Table-2: Laboratory e(max)

It is convenient to say that the calculations of the maximum errors are based on the data contained in the Reports of the Program of Interlaboratorial Comparison (PCI-Wh), from 2000 to 2004, and that, in these documents, clearly is said that each participant is responsible for the calculation of the error and the uncertainty of the laboratory, being also recommended the use of the "Guide for the Expression of the Uncertainty of the Measurement", edited for the Inmetro, as reference of consultation for calculation of the uncertainty of the measurement.

4. TECHNIQUES CONSIDERATIONS

This work is based on the linearization of the normal function distribution, placing e(max) values in the axle-x, sorted from the minor to the greater, and in the axle-y we have placed the median-rank of order i´s. Tracing graph Y-X a straight line is to be showed in the case of the data in analysis have a normal behavior.
The methodology of linearization of the normal distribution can be read in books on goodness-of-fit techniques. Some are in the reference of this work.
In this process the graph supplies three measures: the first measure is obtained from the crossing of the straight line of regression of the data with the straight line F(emax)=0,5 and represents the average of the data, being a measure of the distance relative to the error zero; the second measure is the inclination of the straight line of regression that represents, in this work, volatileness (dispersion of the data around the average), being a measure of the instability of the process, and the third measure is related to the degree of proximity of the data in relation to the straight line of regression, represented for the r-square. These measures will supply the Beta Factors, that are the indices called beta-3 and beta-2, beta-1, objects of this study.

Beta-1 index (Interpretation): The origin of the axle of the abscissas represents, here, a perfect process where the error is zero with respect to the VVC (Conventional True Value). The point of intersection of the straight line of regression of the data with straight line F (emax) =0,5 supplies the average of these data. How much bigger is this average, much more far away from the origin it is presented, that is, much more distant from error zero this point will be. As this point of intersection represents the maximum error (module of the error plus the uncertainty) and the module of the error must be corrected during the calibration of the circulating standard, this point should be more related to only the uncertainty of the calibration, predominantly the one that is inherited from the calibration certificate.
To obtain the Beta-1 index we divide the e (max) of a company to the e (max) of Inmetro, year by year. Therefore, one expects that this index is bigger or at least equal to the value 1.
We did not stipulate limits to this index, what could be made in posterior studies.

Beta-2 index (Interpretation): The more inclined is the straight line of regression of the data, the less spread the data is around its average and the less is the volatileness of the process (less unstable). To obtain this index we divide the value of this inclination for Inmetro laboratory to value of the inclination of one determined company, year by year.
The way this index has been calculated, it presents values that, how much bigger it is, how much unstable the process is. Also we do not stipulate limits for this index, what could be made in posterior studies.

Beta-3 index (Interpretation): So that this study be characterized as adjusted it is necessary that the data present a certain level of approach with the straight line, that is, that the process be normal. Thus, points that are out of the curve (outliers) will tell us that there is not a conformity of the data relative to the straight line. To obtain this index we only need to get the r-square of the data. This index indicates, then, an unconformity of the data when its value is inferior to a certain limit, that we could adopt as being next to 80%, according to some authors.
Other boundary-values could be adopted in the case of the PCI-Wh Program. For the time being, laboratories with beta-3 lesser than 80% could have its data re-analyzed to detect problems of not conformity with their process.
We present, in Figure-1 below, the graph relative to the study of the data of calibration of the PCI-Wh for the Inmetro (initial calibration), that we have considered, in this study, as being our point or reference.

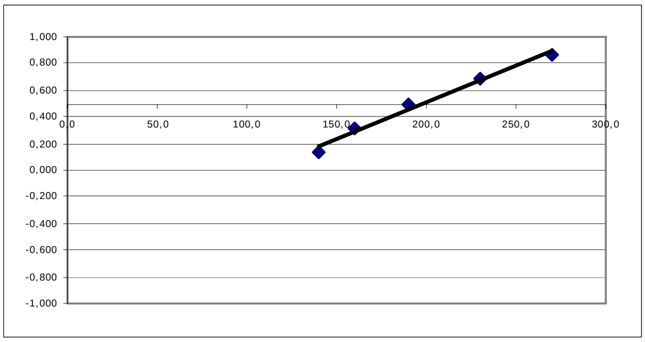

Figura-1: Initial calibration PCI-Wh (Inmetro)

We observe, in Table-3 below, a very good conformity when we study the data of the Inmetro, what it is represented by the raised value of the r-square (98.29%). The distance relative to error zero is represented by the average of 198 µWh/Wh, and the 0,55% inclination value, that has been obtained by the division of (0,5 + 0,6) /198, reflects the volatileness of the data (instability of the process).

média =>	198,0 µWh/Wh
inclinação =>	0,55%
r-quadrado =>	98,29%

Tabela-3: Result of the linearization

From this point on we defined our reference as being these values of Inmetro, obtained during the initial calibration of the PCI-Wh. Thus, the references for the Factors Beta 1, 2 and 3 obtained dividing the values of the Inmetro for themselves, which are presented in Table-4.

média => beta1=	1,00	(afastamento do VVC)
inclinação => beta2=	1,00	(volatilidade)
r-quadrado => beta3=	100,0%	(conformidade)

Tabela-4: Referential Values for Beta

The objective now is to construct, for each laboratory, a graph having the Inmetro as reference and, from this graph, calculate the values that will generate all the other Beta Factors when we divide theses values by the values of Inmetro. In the case of the Chesf, analyzing its graph showed in Figure-2, we have the following values:

Dados da análise	
Inclinação →	0,185%
Média →	344µWh/Wh
r-quadrado →	79,88%

Tabela-5: Calculated values for Chesf´s Lab

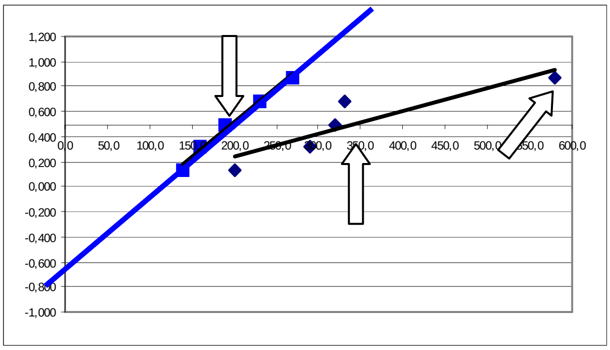

Figura-2: Comparative graph Chesf X Inmetro

As we expected there is an irregular behavior when we study the data of Chesf, what it is represented by the value of the r-square next to 70%. The relative distance between the data average and the error zero is represented by the value of 344µWh/Wh, what is sufficiently far away from the Inmetro referential average (198µWh/Wh). Low inclination of the line of about 0,185% represents a bigger scattering of the data compared with the referential (Inmetro-0,55%), reflecting a higher instability of the procedures of calibration of Chesf laboratory. We can observe, in the graph, that there exists a point out of curve (outliers) which reflects an unfavorable Beta Factors for Chesf. It has been through this analysis that we have the idea to use this methodology of data linearization to create Beta Factors to study the behavior of Chesf laboratory relative to the behavior of Inmetro laboratory, and relative to the behavior of the other laboratories, as well. This has helped us in the analysis of our conformity according to the procedures of calibration used during the period of analysis.
Studying this particular outlier we could get to the point that the inherited uncertainty from RBC was the greatest responsible for our unconformity.

Now we proceed on calculating the Beta Factors of Chesf, represented in Table-6 below:

$$\beta_1 = \frac{media\ Chesf}{media\ Inmetro} = \frac{344}{198} = 1{,}74$$

$$\beta_2 = \frac{Inclinacao\ Inmetro}{Inclinacao\ Chesf} = \frac{0{,}55}{0{,}185} = 2{,}98$$

$$\beta_3 = (r - quadrado) = 79{,}88\%$$

Tabela-6: Chesf Beta Factor

Obs: In the calculation of the beta-2 we have inverted the reference to indicate “how much bigger, more volatile”.

Below we present Table-7 with all the Beta Factors (beta-1, beta-2 and beta-3) of the laboratories which participated if the PCI-Wh Program from 2000 up to 2004.

Empresa	Inmetro Inicio	A	B	chesf	D	E
beta-1	1,00	1,48	1,93	1,74	2,09	1,11
beta-2	1,00	1,26	5,00	2,98	2,10	1,43
beta-3	98,29%	97,06%	88,15%	79,88%	73,64%	68,92%

Empresa	F	G	H	I	J	Inmetro fim
beta-1	0,84	2,02	2,94	3,29	1,54	0,87
beta-2	0,65	0,89	5,15	1,18	2,45	0,86
beta-3	87,62%	73,33%	75,15%	95,76%	87,39%	97,03%

Tabela-7: Beta Factors of all laboratories

Analyzing the beta values we observe that almost all laboratories have presented values superior than 1 for beta-1 and beta-2 factors, what has been expected. However company F has presented values inferior than 1 (with regard to the Inmetro reference) for both factors, and the company G has presented beta-2 index value lesser than the reference. This, in principle, is not in accordance with the methodology of analysis adopted in this work. However, such fact can be explained by similar situations such as those we have found in Chesf, or by the use of different procedures of calibration used by the Inmetro to calibrate the PCI-Wh and to calibrate the standards of the laboratory in question, leading to different uncertainties. It is important to stand out that the final calibration of the Inmetro also has presented factors values lesser than that presented by the reference. This analysis reflects, among others, that the circulating standard has a sufficiently steady behavior throughout the time.

We present, in Figure-3 below, the general graph that relates the normalized data of all the laboratories. We can observe, in this type of graph, an evident visual easiness to interpret the results of each laboratory, among themselves, and in relation to the Inmetro. In it we can see the Chesf (red line) and the Inmetro (initial calibration in blue line and final calibration in green line). We observe clearly, as well, that there exists a laboratory that presents Beta Factors better than that presented by reference (Inmetro), what, in principle, deserves a more refined study on the subject.

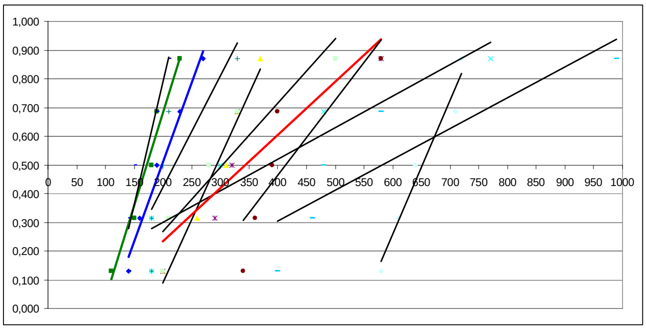

Figura-3: General graph of all the laboratories

5. CONCLUSION

This work shows a methodology of analysis of data that leads to elaboration of graphs of easy visualization to compare calibrations among the laboratories that participate in the PCI-Wh Program. It discloses, as well, if a laboratory has an unstable behavior in its process of calibration and, also, reveals unconformity in this process. Finally, however not less important, the methodology serves to analyze the behavior of the circulating standard in what its instability says respect. We remember that this analysis takes into consideration the whole period of participation of the laboratories in the PCI-Wh Program, not being, therefore, a prompt analysis (year by year analysis). With this analysis we can also check uniformity in the procedures of calibrations, in the procedures of calculations of uncertainties and in the procedures of correction of the systematic erros.

The Beta Factors (beta-1, beta-2 and beta-3) reflect the behavior of the calibration of the PCI-Wh standard, showing that the inherited uncertainty of the calibration excessively influences the Best Capacity of Measurement of a laboratory. As improvement for future work, we could suggest, that would be stipulated limits for each one of these factors, taking as reference the behavior of Inmetro´s calibration. One another aspect that is evidenced in this work, and that deserves a deeper evaluation, says respect to the use of Wh reference standards, by the metrology laboratories, with uncertainties of the same order of the ones used by Inmetro, compromising the consecrated concept of TUR equal or bigger than three.

We suggest that the group that composes the PCI-Wh Program evaluates the convenience of the use of this tool as an improvement of the analysis of the data of calibration of the circulating standard.

REFERENCES

I. PCI-Wh reports, from 2000 to 2004.
II. Goodness-of-fit Techniques, D'Agostino, Ralph B.,1986 Marcel Dekker, Inc.,pp.35-59
III. Reliability in Engineering Design, Kapur,K.C., 1977, John Wiley & Sons, pp. 300
IV. Probability and Statistics, DeGroot, M.H.,1989, Addison-Wesley Publishing Company, pp. 529-530
V. Probability, Statistics and Queuing Theory, with Computer Science Applications, Allen, A.O., 1990, Academic Press, Inc, pp. 524-534
VI. Procedimentos Estatísticos para o Desenvolvimento da Avaliação de Desempenho de Sistemas, Sousa, J.J.L., 1998, Dissertação de Mestrado na UFPE, pp. 47-68
VII. Probability Plotting for the Normal Distribution, sites: http://www.weibull.com http://www.statsoft.com