XVIII IMEKO WORLD CONGRESS
Metrology for a Sustainable Development
September, 17 – 22, 2006, Rio de Janeiro, Brazil

FIRST ROUND OF THE PROFICIENCY TEST SCHEME ON BRAZILIAN *CACHAÇA*

[1]*Akie K. Ávila*, [2]*Renata M. H. Borges*, [1]*Thiago de O. Araujo*, [1]*Maria Cristina B. Quaresma*, [1]*Fabio S. Bazilio*, [1]*Lígia C. C. Oliveira*

[1,2]INMETRO, [1]Chemical Metrology Division of the Directorate of Scientific and Industrial Metrology, [2] General Coordination of Accreditation. toaraujo@inmetro.gov.br

Abstract: *Cachaça* is the Brazilian most popular spirit and the quality of analytical results is essential for protection of consumers and to expand the international acceptance of the product. The first round of the proficiency-testing scheme (interlaboratory study), Brazilian *cachaça*, was carried out between November 2004 and March 2005. The blended sample was prepared for this round containing several components, including copper, methanol and ethyl carbamate. The aim of this paper is to evaluate the laboratory performance on the determination of the copper concentration in Brazilian *cachaça*. Eleven laboratories reported results for copper; these results were compared to the reference value attributed by Inmetro in partnership with The Laboratory of the Government Chemist – LGC from the United Kingdom and with the consensus value. A comparison between these values was made to determine which approach could guarantee better reliability to the conclusions taken from the PT scheme. The samples used have shown good homogeneity and stability, the reference value attributed by Inmetro and LGC is SI traceable and showed to be more appropriate to be used as the conventional true value than the consensus value calculated with the participants' results.

Keywords: *Cachaça*, Proficiency Test Scheme, copper

1. INTRODUCTION

The Brazilian *cachaça* is the most popular spirit, involving an internal financial market of more than 2 billion dollars per year, which reaches all the social classes with a growing exportation index [1]. Due to this large consumption, it is necessary to monitor and to compare the analytical results of different Brazilian laboratories. Concerned about the health of consumers, the Brazilian Ministry of Agriculture issued a regulation establishing the levels allowed of the most common contaminants in *cachaça* [2], and to help laboratories to comply with this regulation, Inmetro is developing a CRM and organizing PT schemes. This paper reports the evaluation of the methodologies applied for copper determination, thereby contributing to the improvement of the beverage quality.

The comparison of different laboratories that perform comparable analyses with their own individual method was established. The first round of the proficiency-testing (PT) scheme, Brazilian *cachaça*, was carried out between November 2004 and March 2005. One of the objectives of the program is to organize one round per year for evaluating the laboratory performance on the determination of the copper concentration in Brazilian *cachaça*.

The blended sample was prepared for this round containing several components, including copper, methanol and methyl carbamate.

The evaluation of the results, when compared to the reference value, was carried out with the fixed coefficient of variance (CV) of 2%, most laboratories are satisfied with these limits. Laboratories that do not need to comply with these limits, or need more strict limits, could recalculate their z-scores values as indicated in the final report submitted to all participating laboratories.

2. EXPERIMENTAL

A batch of 180 bottles containing 100 mL of Brazilian *cachaça* was prepared. From this batch, 20 bottles were selected and subjected to a homogeneity assessment.

The analytical determinations were carried out by Atomic Absorption Spectrometry (AAS), the equipment used was an Aanalyst 800 equipped with an air-acetylene burner and hollow-cathode lamp. For copper determination an external calibration curve was used and the standards were prepared from NIST calibration certified reference material (CRM) SRM 3114 in an ethanol solution (40%) prepared from Merk HPLC grade ethanol and type one water obtained from a Millipore system, to reproduce the *cachaça's* matrix.

The analysis performed at LGC were made using Isotope Dilution Inductively Coupled Plasma Mass Spectrometry (ID-ICPMS), this methodology is considered a primary method of analysis by the *Comité consultatif pour la quantité de matière* (CCQM). The CCQM is the BIPM committee responsible for the chemical metrology in the highest level of hierarchy in the world, and a primary method is considered to have the highest metrological quality and a complete uncertainty sources description.

3. RESULTS AND DISCUSSION

The objective of the homogeneity test assessment was to determine the "between-bottle" homogeneity standard deviation (s_{bb}) to evaluate whether the samples of the batch can be considered statistically equal. The results of the homogeneity test are presented in Table 1 and 2. For evaluation of the results for homogeneity, unifactorial analysis of variance (one-way ANOVA) was applied.

Table 1. Results obtained for homogeneity test for copper.

Samples	Mean (µg/kg)	SD* (µg/kg)	RSD** (%)
24C	2132,41	22,54	1,1
16E	2090,55	81,81	3,9
11C	2110,14	13,49	0,6
30A	2110,20	12,37	0,6
01D	2124,34	21,54	1,0
09B	2122,85	6,41	0,3
14D	2116,52	10,52	0,5
04F	2120,53	13,46	0,6
30D	2122,67	12,34	0,6
30F	2128,84	15,25	0,7
17C	2137,55	18,89	0,9
12E	2124,40	17,25	0,8
01A	2121,28	8,39	0,4
22C	2118,96	9,17	0,4
11F	2119,18	8,40	0,4
05B	2118,22	16,34	0,8
14B	2125,84	12,33	0,6
25A	2125,22	20,88	1,0
13F	2133,55	8,66	0,4
10E	2121,57	5,70	0,3

*SD: Standard deviation; **RSD: Relative Standard Deviation.

Table 2: ANOVA to determine the homogeneity of *cachaça* samples.

Source of variation	SQ	df	MQ	F	P value	Critical F
Between Groups	9448,5	19	497,3	**0,94**	0,54	**1,72**
Within Groups	42236,1	80	528,0			
Total	51684,6	99				

The calculated *F* is smaller than the critical one confirming the batch is homogeneous.

For characterization, the samples were analyzed at Inmetro by AAS and also at LGC by ID-ICPMS in order to validate the methodology used at Inmetro. The results obtained by LGC and Inmetro are presented in Table 2 and 3, respectively.

Table 2. Values obtained at LGC.

Sample	Cu Concen. (mg . kg^{-1})
03A	2.0428
03A	2.0514
03A	2.1281
07F	2.0779
07F	2.0503
07F	2.0091
02E	2.1402
02E	2.0863
02E	2.0768
Mean (mg/kg)	2.0737
Standard deviation (mg/kg)	0.0414
RSD* (%)	2.0

*RSD: Relative Standard Deviation.

Table 3. Values obtained at Inmetro used for characterization.

Sample	Cu Concen. (mg . kg^{-1})
1 A	2.0725
30 F	2.1038
11 C	2.0039
18 D	2.0097
03 C	1.9760
10 E	1.9920
08 A	1.9380
14 B	1.9250
09 F	2.0140
22 C	2.0310
05 B	2.0060
24 F	1.9920
18 C	2.0320
29 E	2.0140
25 A	2.0490
28 B	2.0550
03 D	1.9530
13 F	1.9330
27 B	2.3010
04 F	2.3110
20 A	2.0530
12 D	2.0710
17 C	2.0720
06 E	2.0490
01 D	1.9980
11 F	2.0140
16 E	2.0390
08 F	2.0450
20 B	2.0280
15 C	2.0430
30 A	2.0710
22 E	2.0850
09 B	2.0960
14 D	2.0740
30 D	2.0410
01 F	2.0650
24 C	2.1460
12 E	2.1520
12C	2.0380
Mean (mg/kg)	2.0485
Standard deviation (mg/kg)	0.0790
RSD* (%)	3.9

*RSD: Relative Standard Deviation.

An *F* test was applied to assess the equivalence of the variances of the results and determine whether they are comparable. The results are shown in Table 4.

Table 4. *F* test: two samples of Brazilian *Cachaça* for variance.

	Variable 1	*Variable 2*
Mean	2.07	2.05
Variance	0.00	0.01
Observations	9.00	39.00
df	8.00	38.00
F	**0.28**	
P(F<=f) uni-tailed	0.03	
critical f uni-tailed	**0.33**	

As the critical *F* is greater than the calculated one, therefore it can be assumed the variances obtained at the laboratories are comparable.

The *t* test was performed to assess whether the means are statistically equal and it can be seen in Table 5.

Table 5. *t* test: two samples of *Cachaça* assuming equivalent variances.

	Variable 1	*Variable 2*
Mean	2.07	2.05
Variance	0.00	0.01
Observations	9.00	39.00
Grouped variance	0.01	
Mean diference hypothesis	0.00	
df	46.00	
Stat t	**0.92**	
P(T<=t) uni-tailed	0.18	
crítico t uni-tailed	1.68	
P(T<=t) bi-tailed	0.36	
critical t bi-tailed	**2.01**	

As the calculated *t* is smaller than the critical one, it can be assumed the means are statistically equal.

In the recent years a great efforts is being made regarding the improvement of stability study designs as well as the estimation of the contribution of short- and long-term (in) stability to the uncertainty budget of reference materials. Inmetro has made a stability study for short-term stability assessment through a residues analysis of the concentration as a function of time.Table 6 shows the analyses of variance of the regression due to the period of the study [3, 4].

Table 6. Analyses of Variance – multiple regressions.

	SS	df	MS	F value	P value
Regression	894.8024	1	894.8024	1.225	**0.35**
Residual	2191.438	3	730.4795		
Total	3086.241				

In case P was smaller than 0.05 (95%), such regression would present a tendency and therefore it would characterize the material as unstable. For the sample used, the *P* value is much greater than 0.05 indicating that the regression is insignificant and that the material is stable.

Eleven laboratories signed up for participation in the PT scheme. And it was decided that the results reported by the participating laboratories would be compared with Inmetro's and LGC's results and they would also be compared to the consensus value obtained from their own results.

In order to establish the performance of the laboratories, the results were compared according to ISO 5725 [5].

The determination of the consensus value was carried out according to an adapted procedure that is described in the subsequent paragraphs. Classical, as well as robust estimates are determined.

The data was treated analogous to the procedures described in ISO 5725.1 and .2. First, from the complete set of the reported data the grand mean (m), the repeatability standard deviation (s_r), the between-laboratory standard deviation (s_l) and the reproducibility standard deviation (s_R) are calculated.

The consensus value is defined according to:

$$Y = \frac{\sum_{i=1}^{p} n_i . y_i}{\sum_{i=1}^{p} n_i}$$

where *n* equals the number of results reported by laboratory i, y_i the average result of this laboratory and *p* is the total number of laboratories. The repeatability standard deviation is follows from:

$$s_r^2 = \frac{\sum_{i=1}^{p} (n_i - 1).s_i^2}{\sum_{i=1}^{p} (n_i - 1)}$$

where s_r is the repeatability standard deviation of the results of laboratory *i.*

The between-laboratory standard deviation is calculated according to:

$$s_l^2 = \frac{s_d^2 - s_l^2}{\eta}$$

The reproducibility standard deviation is calculated according to:

$$s_R^2 = s_l^2 + s_r^2$$

The complete data set is also used to calculate a number of robust estimates: the median (y_{med}), the median of absolute differences (MAD) and the average absolute deviation (AAD).

For qualification of the laboratory results, z-scores were calculated.

The z-score was defined as:

$$z_i = \frac{y_i - y_{ref}}{y_{ref} . CV}$$

where y_{ref} is the reference value obtained by Inmetro, y_i the result of laboratory *i*. The evaluation of the comparison takes place at a level of 2 % [6] and the results are shown in Table 7.

Table 7. Performance of participants.

Code	Mean	Stand. Dev.	z-score Inmetro	z-score Consensus
PEP1.1/01	2.502	0.160	11.07	1.18
PEP1.1/02	2.097	0.000	1.18	-0.30
PEP1.1/03	3.250	0.000	29.33	4.12
PEP1.1/05	2.412	0.000	8.87	0.81
PEP1.1/06	2.086	0.011	0.92	-0.35
PEP1.1/07	1.837	0.002	-5.16	-1.27
PEP1.1/09	2.027	0.061	-0.52	-0.55
PEP1.1/11	1.999	0.003	-1.21	-0.67
PEP1.1/12	2.405	0.032	8.70	0.82
PEP1.1/13	1.029	0.003	-24.88	-4.25
PEP1.1/16	2.363	0.030	7.68	0.67

The Figure 1 shows the dispersion of the results in comparison with Inmetro's assigned value.

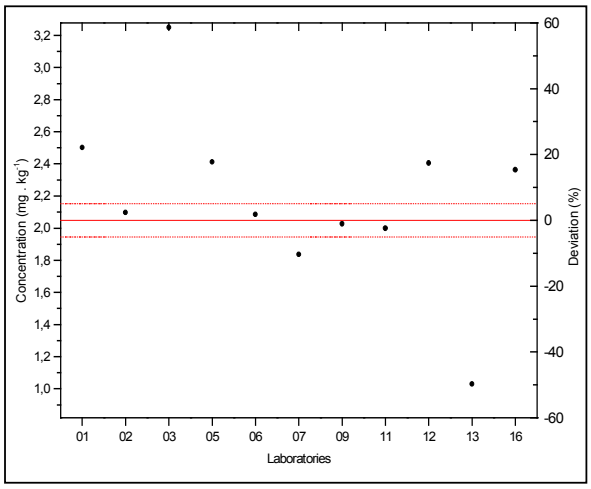

Figure 1 – Results of participants and Inmetro's assigned value.

Figure 2 shows the z-score of participants relative to Inmetro's reference value using a CV of 2%.

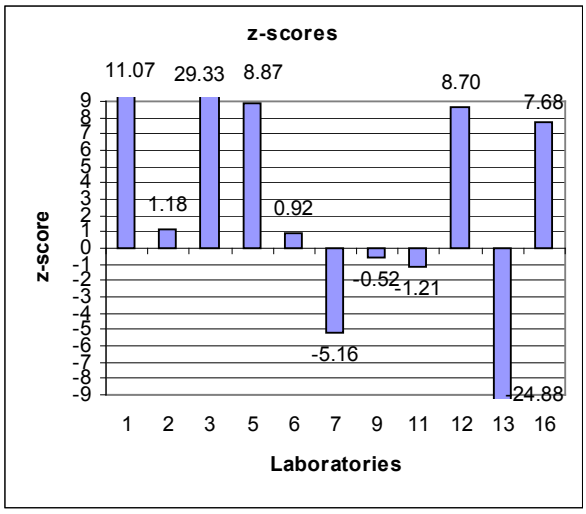

Figure 2 – z-scores of participants relative to Inmetro's reference value.

After removing the highest and the lowest values, detected to be outliers, the mean and the standard deviation of the results provided the participating laboratories were calculated. Table 8 shows the new values. The analysis for determination of outliers is presented in Figure 3.

Table 8. Mean and standard deviation of participant's results.

Consensus value	Reproductibility	RSD
2.1814	0.2713	12.4

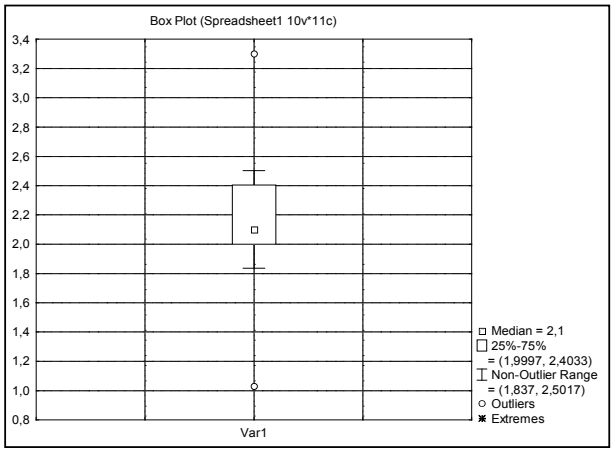

Figure 3 – Analysis of results for determination of outliers.

The z-scores of participants related to the consensus value are much better if compared to the one related to Inmetro's reference value. It is so because the dispersion of the results was used as the CV, therefore allowing a greater deviation of the participant's value from the consensus value. It is important to notice that the CV is about 12% in this case; this implies that a laboratory is allow to deviate 24% from the consensus value and still be considered satisfactory. It can be denoted in Figure 4.

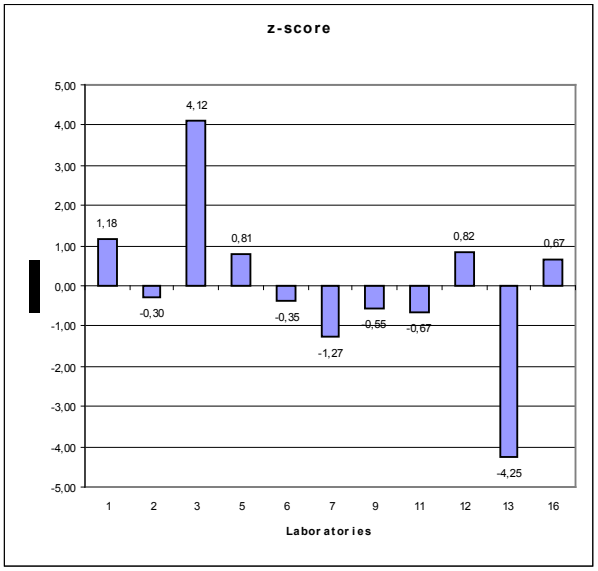

Figure 4 – z-scores of participants relative to consensus value.

A *t* test was applied to compare the consensus value with the reference value and, although the variances are equivalent, the values cannot be considered statistically equal. The results are displayed in Table 9 and 10.

Table 9. *F* test: Inmetro's and participant's values for variance.

	Variable 1	*Variable 2*
Mean	2.049	2.192
Variance	0.006	0.053
Observations	39.000	9.000
df	38.000	8.000
F	**0.118**	
P(F<=f) uni-tailed	0.000	
critical f uni-tailed	**0.456**	

Table 10. *t* test: Inmetro's and consensus values.

	Variable 1	*Variable 2*
Mean	2.05	2.19
Variance	0.01	0.05
Observations	39.00	9.00
Grouped variance	0.01	
Mean diference hypothesis	0.00	
df	46.00	
Stat t	**-3.24**	
P(T<=t) uni-tailed	0.00	
crítico t uni-tailed	1.68	
P(T<=t) bi-tailed	0.00	
critical t bi-tailed	**2.01**	

4. CONCLUSION

The Brazilian *Cachaça* proficiency-testing scheme with SI-traceable reference values provides participants the possibility to evaluate both their international comparability and their laboratory bias in anlysis of copper.

The homogeneity study indicated no significant contribution due to between bottle variations. In this case, the combined standard uncertainty from the determination of the composition and that of the homogeneity test was considerably lower than the evaluation level of the proficiency test.

This comparison showed the importance of having a reference value to compare the participant's results. The consensus value would lead to untrue conclusions, for the mean of result disregarding the ouliers is different from the conventional true value. Besides, the CV calculated with the participant's values is too large and allow many result to be considered satisfactory generating no corrective actions in laboratories that may be in need of them.

Inmetro has also validated its methodologies of analysis of copper in cachaça through the comparison of its results with LGC results, and will be able to attribute reference values for the contaminants in the future rounds of this study.

5. ACKNOWLEDGEMENTS

The authors would like to thank MCT that has granted Inmetro with financial support through FINEP/CNPq.

And also would like to thank Dr Ruth Hearn from LGC for the partnership, receiving our researcher and analyzing the samples.

6. REFERENCES

[1] Nascimento, R. F. et al., "Mineral profile of Brazilian *Cachaças* and other international spirits", J. Food Composition and Analysis, vol. 12, pp. 17-25, 1999.

[2] Instrução Normativa Nº 13, de 29 de junho de 2005.

[3] ISO Guide 35: Reference Materials – General and statistical principles for certification, 2006.

[4] Linsinger TPJ, Pauwels J, Lamberty A, Schimmel HG, van der Veen AMH, Siekmann L **"**Estimating the uncertainty of stability for matrix CRMs", Fresenius J Anal Chem, vol 370, pp. 183-188, 2001.

[5] ISO 5725-1994 (E), "Accuracy (trueness and precision) of measurement methods and results".

[6] Van der Veen AMH, Linsinger TPJ, Lamberty A, Pauwels J, "Uncertainty calculations in the certification of reference materials 3. Stability study", Accred Qual Assur, vol. 6, pp. 257-263, 2001.