Results of Proficiency Test
LIQUEURS
November 2012

Organised by: Institute for Interlaboratory Studies (iis)
Spijkenisse, the Netherlands

Authors: ing. R.J. Starink
Correctors: dr. R.G. Visser & ing. N. Boelhouwer

Report: iis12C13

February 2013

CONTENTS

1. INTRODUCTION

Since 2009, a proficiency test for liqueurs is organised every year by the Institute for Interlaboratory Studies. During the planning of the annual proficiency testing program 2012/2013, it was decided to continue the proficiency test for the analysis of liqueurs. In this interlaboratory study, 11 laboratories in 4 different countries have participated. See appendix 2 for a list of number of participants per country. In this report, the results of the 2012 proficiency test are presented and discussed.

2 SET-UP

The Institute for Interlaboratory Studies (iis) in Spijkenisse, the Netherlands, was the organiser of this proficiency test. Analysis for fit-for-use and homogeneity testing were subcontracted. It was decided to send two different samples of liqueur (1* 0.5 L of Spirit, labelled #12153 and 1* 0.5 L of chocolate liqueur, labelled #12154). Participants were requested to report rounded and unrounded results. The unrounded results were preferably used for statistical evaluation.

2.1 QUALITY SYSTEM

The Institute for Interlaboratory Studies in Spijkenisse, the Netherlands, has implemented a quality system on ISO guide 43, ILAC-G13:2007 and ISO17043:2010. This ensures strict adherence to protocols for sample preparation and statistical evaluation and 100% confidentially of participant's data. Also customer's satisfaction is measured on regular basis by sending out questionnaires.

2.2 PROTOCOL

The protocol followed in the organisation of this proficiency test was the one as described for proficiency testing in the report 'iis Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation' (iis-protocol, version 3.2) of January 2010.

2.3 CONFIDENTIALITY STATEMENT

All data presented in this report must be regarded as confidential and for use by the participating companies only. Disclosure of the information in this report is only allowed by means of the entire report. Use of the contents of this report for third parties is only allowed by written permission of the Institute for Interlaboratory Studies. Disclosure of the identity of one or more of the participating companies will be done only after receipt of a written agreement of the companies involved.

2.4 SAMPLES

The necessary bulk material (Spirit) for sample #12153 was obtained from a local producer. The approximately 10 litre bulk sample was, after homogenisation in a precleaned can, divided over 20 amber glass bottles of 0.5 L and labelled #12153. The homogeneity of these subsamples was checked by determination of Density in accordance with ASTM D4052:02e1 on 8 stratified random selected samples.

Sample	*Density @ 15ºC in kg/L*
Sample #12153-1	0.96231
Sample #12153-2	0.96230
Sample #12153-3	0.96230
Sample #12153-4	0.96231
Sample #12153-5	0.96227
Sample #12153-6	0.96229
Sample #12153-7	0.96232
Sample #12153-8	0.96232

table 1: Homogeneity test results of subsamples #12153

From the test results of table 1, the repeatability was calculated and compared with 0.3 times the corresponding target reproducibility in agreement with the procedure of ISO 13528, Annex B2 in the next table:

	Density @ 15ºC in kg/L
r (Observed)	0.00005
reference method	D4052:02e1
0.3 * R (ref. method)	0.00015

table 2: Repeatability of subsamples #12153

The necessary bulk material (Chocolate Liqueur) for sample #12154 was obtained from a local producer. The approximately 25 litre bulk sample was, after homogenisation in a precleaned can, divided over 30 amber glass bottles of 0.5 L and labelled #12154. The homogeneity of these subsamples was checked by determination of Density in accordance with ASTM D4052:02e1 on 8 stratified random selected samples.

Sample	*Density @ 15ºC in kg/L*
Sample #12154-1	1.11209
Sample #12154-2	1.11209
Sample #12154-3	1.11209
Sample #12154-4	1.11210
Sample #12154-5	1.11210
Sample #12154-6	1.11209
Sample #12154-7	1.11210
Sample #12154-8	1.11210

table 3: Homogeneity test results of subsamples #12154

From the test results of table 3, the repeatability was calculated and compared with 0.3 times the corresponding target reproducibility in agreement with the procedure of ISO 13528, Annex B2 in the next table:

	Density @ 15ºC in kg/L
r (Observed)	0.00001
reference method	D4052:02e1
0.3 * R (ref. method)	0.00015

table 4: Repeatability of subsamples #12154

The repeatabilities of the results from the homogeneity test were in agreement with the requirements of the respective standards. Therefore, homogeneity of all the prepared subsamples was assumed.

To each of the participating laboratories 1*0.5 L bottle of sample #12153 and 1*0.5 L bottle of sample #12154 were sent on October 31, 2012.

2.5 ANALYSES

The participants were asked to determine on sample #12153 and #12154: Density @ 20^{o}C, pH and Strength (in %V/V).
To get comparable results a detailed report form, on which the units were prescribed as well as some of the required standards and a letter of instructions were prepared and made available for download on the iis website (www.iisnl.com).
A SDS and a form to confirm receipt of the samples were added to the sample package.

3 RESULTS

During four weeks after sample despatch, the results of the individual laboratories were received. The original reported results are tabulated per determination in appendix 1 of this report. The laboratories are presented by their code numbers.

Directly after deadline, a reminder fax was sent to those laboratories that had not yet reported any results at that moment.
Shortly after the deadline, the available results were screened for suspect data. A result was called suspect in case the Huber Elimination Rule (a robust outlier test) found it to be an outlier. The laboratories that produced these suspect data were asked to check the results. Additional or corrected results are used for data analysis and original results are placed under 'Remarks' in the result tables in appendix 1.

3.1 STATISTICS

Statistical calculations were performed as described in the report ‘iis Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation’ (iis.-protocol, version 3.2) of January 2010.

For the statistical evaluation the *unrounded* (when available) figures were used instead of the rounded results. Results reported as '<…' or '>…' were not used in the statistical evaluation. First the normality of the distribution of the various data sets per determination was checked by means of the Lilliefors-test. After removal of outliers, this check was repeated. Not all data sets proved to have a normal distribution, in which cases the statistical evaluation should be used with due care.

In accordance with ISO 5725 (1986 and 1994) the original results per determination were submitted subsequently to Dixon and Grubbs outlier tests. Outliers are marked by D(0.01) for the Dixon test, by G(0.01) or DG(0.01) for the Grubbs test. Stragglers are marked by D(0.05) for the Dixon test, by G(0.05) or DG(0.05) for the Grubbs test. Both outliers and stragglers were not included in the calculations of averages and standard deviations.

For each assigned value, the uncertainty was determined in accordance with ISO13528. Subsequently the calculated uncertainty was evaluated against the respective requirement based on the target reproducibility in accordance with ISO13528. When the uncertainty passed the evaluation, no remarks are made in the report. However, when the uncertainty failed the evaluation it is mentioned in the report and it will have consequences for the evaluation of the test results.

Finally, the reproducibilities were calculated from the standard deviations by multiplying these with a factor of 2.8.

3.2 GRAPHICS

In order to visualize the data against the reproducibilities from literature, Gauss plots were made, using the sorted data for one determination (see appendix 1). On the Y-axis the reported analysis results are plotted. The corresponding laboratory numbers are under the X-axis.
The straight horizontal line presents the consensus value (a trimmed mean). The four striped lines, parallel to the consensus value line, are the +3s, +2s, -2s and -3s target reproducibility limits of the selected standard. Outliers and other data, which were excluded from the calculations, are represented as a cross. Accepted data are represented as a triangle.
Furthermore, Kernel Density Graphs were made. This is a method for producing a smooth density approximation to a set of data that avoids some problems associated with histograms (see appendix 4, nos.13-14).

3.3 Z-SCORES

To evaluate the performance of the participating laboratories the z-scores were calculated. As it was decided to evaluate the performance of the participants in this proficiency test (PT) against the literature requirements, e.g. ASTM reproducibilities, the z-scores were calculated using a target standard deviation. This results in an evaluation independent of

the spread of this interlaboratory study. The target standard deviation was calculated from the literature reproducibility by division with 2.8.
In case no literature reproducibility was available, other target values were used. In some cases, literature repeatability is available; in other cases, a reproducibility of a former iis proficiency test could be used and the Horwitz equation can be used to estimate target reproducibility.
The z-scores were calculated according to:

$$z_{(target)} = (\text{result} - \text{average of PT}) / \text{target standard deviation}$$

Absolute values for z<2 are very common and absolute values for z>3 are very rare. The usual interpretation of z-scores is as follows:

| z | < 1	good
1 < | z | < 2	satisfactory
2 < | z | < 3	questionable
3 < | z |	unsatisfactory

When a laboratory did use a test method with a reproducibility that is significantly different from the reproducibility of the reference test method used in this report, it is strongly advised to recalculate the z-score, while using the reproducibility of the actual test method used this in order to evaluate the fit-for-useness of the reported test result.

4. EVALUATION

In this proficiency test no problems were encountered with despatch of the samples. Only one participant reported the results after the final reporting date. Not all laboratories were able to perform all analysis requested. Finally, the 11 reporting laboratories did send in 46 (numerical) results. Observed were 4 outlying results, which is 8.7%. In proficiency studies, outlier percentages of 3% - 7.5% are normal.

4.1 EVALUATION PER TEST

In this section, the results are discussed per test.
The methods, which are used by the various laboratories, are taken into account for explaining the observed differences when possible and applicable. These methods are also in the tables together with the original data. The abbreviations, used in these tables, are listed in appendix 3.

On the report form the participants were asked to fill out the analytical details regarding the strength determination. Seven laboratories answered the questions fully or partially. (See Appendix 2).
For all tests normal distributions were found.

Density: This determination was not problematic for the tested samples. One statistical outlier was observed and both calculated reproducibilities are in agreement with the requirements of ASTM D4052:02e1.

pH: This determination was not problematic for the tested samples. One statistical outlier was observed and both calculated reproducibilities are in agreement with the requirements of EN15490:07.

Strength (%V/V): Regretfully, no standard test method with precision data exists for this determination.
From the analytical details it is clear all participants did perform a distillation before the strength determination.
In total two statistical outliers were observed. When compared with the calculated reproducibilities of the previous proficiency test iis11C13b, the spread found for sample #12153 is very large (0.258 vs 0.150), while the spread for sample #12154 is small (0.088 vs 0.110) in comparison with the spread on a similar sample in the previous PT.

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES

A comparison has been made between the reproducibility as declared by the relevant standard and the reproducibility as found for the group of participating laboratories. The average results per sample, calculated reproducibilities and reproducibilities derived from literature standards (in casu ASTM, EN standards) or previous proficiency tests are compared in the next table.

Parameter	Unit	n	average	2.8 *sd_R	R (lit)
Density @ 20°C	kg/L	11	0.9593	0.0001	0.0005
pH		6	7.28	0.72	0.69
Strength	%V/V	6	34.92	0.26	(0.15)

Table 5: Reproducibilities of sample #12153

Parameter	Unit	n	average	2.8 *sd_R	R (lit)
Density @ 20°C	kg/L	9	1.1097	0.0004	0.0005
pH		5	7.01	0.24	0.69
Strength	%V/V	5	13.86	0.09	(0.11)

Table 6: Reproducibilities of sample #12154

results between brackets are compared with the observed reproducibility in the previous proficiency test.

4.3 COMPARISON OF THE PROFICIENCY TEST OF NOVEMBER 2012 WITH PREVIOUS PT'S

	November 2012	*November 2011*	*November 2010*	*December 2009*
Number of reporting labs	11	13	17	23
Number of results reported	46	67	71	92
Statistical outliers	4	5	8	11
Percentage outliers	8.7%	7.5%	11.3%	12.0%

Table 7: comparison with previous proficiency tests.

In proficiency tests, outlier percentages of 3% - 7.5% are quite normal.

The performance of the determinations of the proficiency tests was compared against the requirements of the respective standards. The conclusions are given the following table:

Parameter	*November 2012*	*November 2011*	*November 2010*	*December 2009*
Herbal liqueur /Spirit				
Density @ 20^{o}C	++	++	--	++
pH	+/-	++	++	++
Strength	(--)	++	--	--
Chocolate liqueur				
Density @ 20^{o}C	++	--	++	--
pH	++	++	++	+/-
Strength	(+)	+	+/-	--

Table 8: comparison determinations against the standard
results between brackets are compared with the spread of the previous proficiency test or estimated from target reproducibility.

The performance of the determinations against the requirements of the respective standards is listed in the above table. The following performance categories were used:

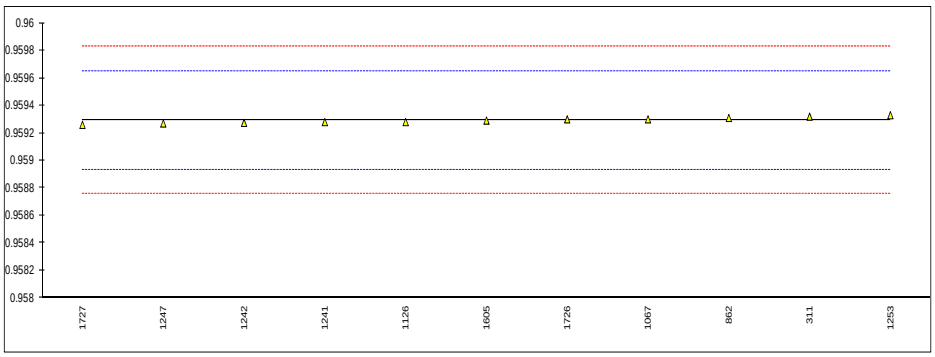
- ++: group performed much better than the standard
- + : group performed better than the standard
- +/-: group performance equals the standard
- - : group performed worse than the standard
- -- : group performed much worse than the standard

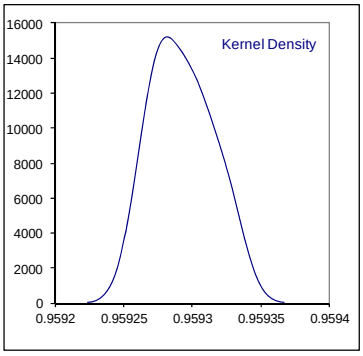
APPENDIX 1

Determination of Density @ 20°C on sample #12153 and #12154; results in kg/L.

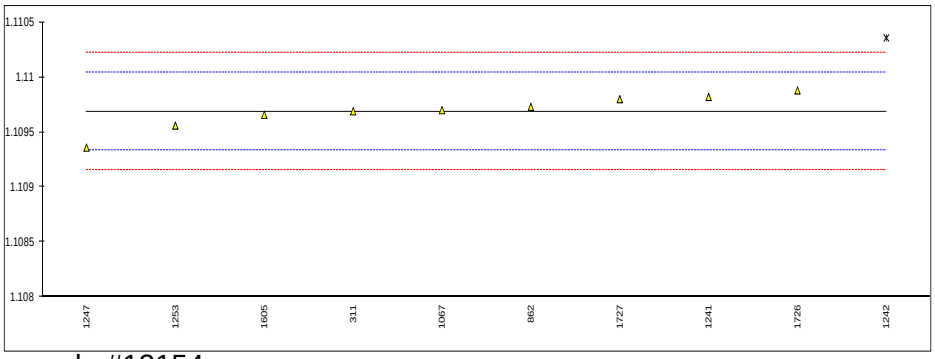
lab	method	#12153	mark	z(targ)	#12154	mark	z(targ)	remarks
311	D4052	0.95932		0.16	1.10969		0.01	
862	D4052	0.95931		0.10	1.10973		0.23	
1067	D4052	0.9593		0.04	1.1097		0.06	
1126	ISO12185	0.95928	C	-0.07	-----		-----	
1241	INH-50	0.95928		-0.07	1.109820		0.74	
1242	D4052	0.959274		-0.10	1.110357	G(0.05)	3.74	
1247	INH-4500	0.95927		-0.12	1.10936		-1.84	
1253	D4052	0.95933		0.21	1.10956	C	-0.72	
1605	D4052	0.959291		-0.01	1.109658		-0.17	
1726	D4052	0.95930		0.04	1.10988		1.07	
1727	D4052	0.95926		-0.18	1.10980	C	0.62	
	normality	OK			OK			
	n	11			9			
	outliers	0			1			
	mean (n)	0.95929			1.10969			
	st.dev. (n)	0.000022			0.000156			
	R(calc.)	0.00006			0.00044			
	R(D4052:02e1)	0.00050			0.00050			

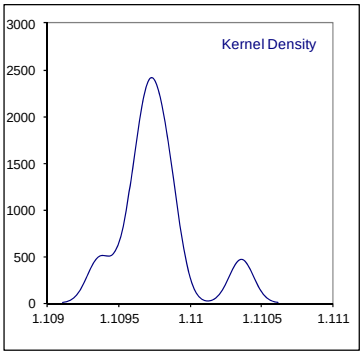
Corrected (C):
Lab 1126 reported 959.28
Lab 1253 reported 1.01956
Lab 1727 reported 1.10862





sample #12153

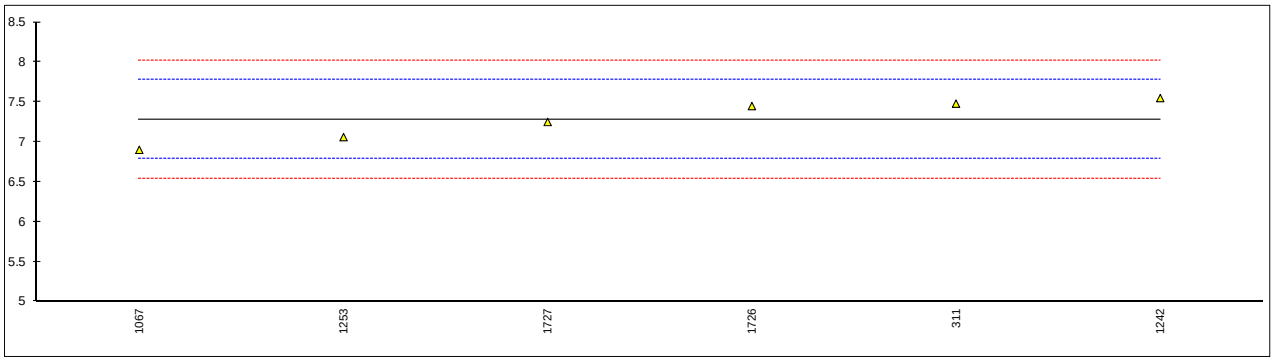




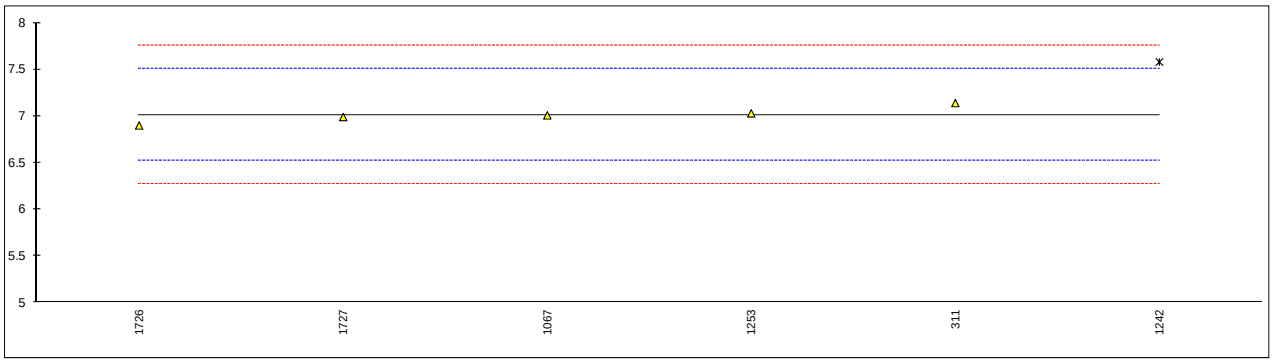
sample #12154

Determination of pH on sample #12153 and #12154

lab	method	#12153	mark	z(targ)	#12154	mark	z(targ)	remarks
311	EN15490	7.48		0.80	7.14		0.51	
862		-----		-----	-----		-----	
1067	EN15490	6.901		-1.55	7.008		-0.02	
1126		-----		-----	-----		-----	
1241		-----		-----	-----		-----	
1242	EN15490	7.55		1.09	7.58	G(0.05)	2.30	
1247		-----		-----	-----		-----	
1253	EN15490	7.06		-0.90	7.03		0.07	
1605		-----		-----	-----		-----	
1726	EN15490	7.45		0.68	6.90		-0.46	
1727	EN15490	7.25		-0.13	6.99		-0.10	
	normality	OK			OK			
	n	6			5			
	outliers	0			1			
	mean (n)	7.282			7.014			
	st.dev. (n)	0.2587			0.0862			
	R(calc.)	0.724			0.241			
	R(EN15490:07)	0.690			0.690			



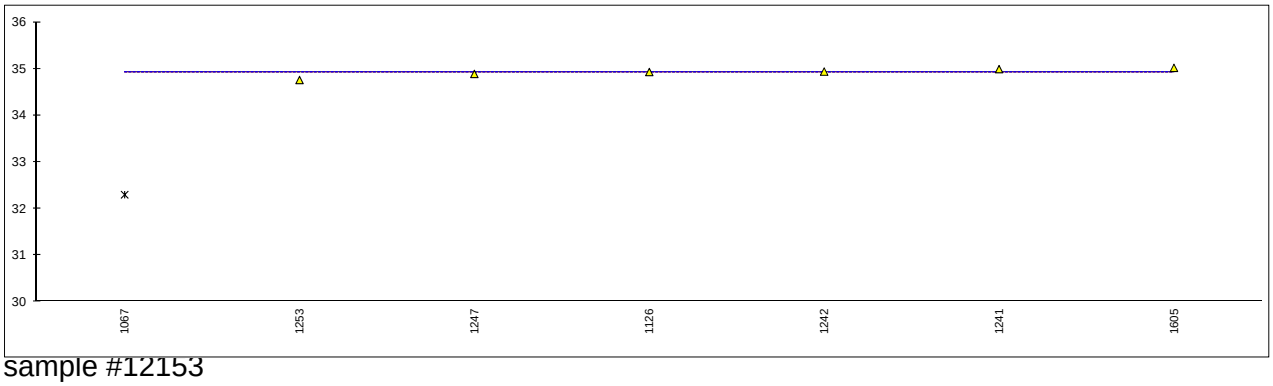
Sample #12153



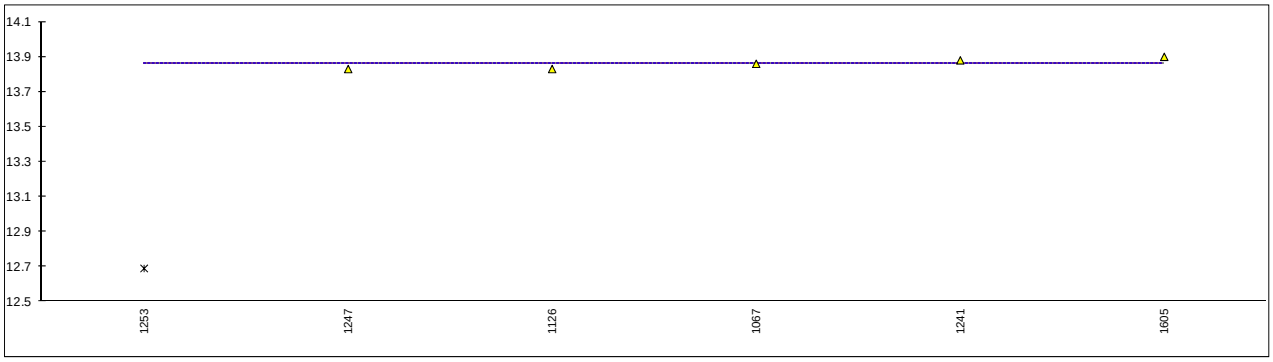
Sample #12154

Determination of Strength on sample #12153 and #12154; results in %V/V

lab	method	#12153	mark	z(targ)	#12154	mark	z(targ)	remarks
311		-----		-----	-----		-----	
862		-----		-----	-----		-----	
1067		32.30	G(0.01)	-----	13.86		-----	
1126		34.93		-----	13.83		-----	
1241		34.993		-----	13.880		-----	
1242		34.94		-----	-----		-----	
1247		34.89		-----	13.83		-----	
1253		34.76		-----	12.69	G(0.01)	-----	
1605		35.02		-----	13.90		-----	
1726		-----		-----	-----		-----	
1727		-----		-----	-----		-----	
	normality	OK			OK			
	n	6			5			
	outliers	1			1			
	mean (n)	34.922			13.860			
	st.dev. (n)	0.0920			0.0308			
	R(calc.)	0.258			0.086			
	R(lit)	unknown			unknown			
	R(iis11C13b)	0.150			0.110			



sample #12153



sample #12154

APPENDIX 2

Analytical details regarding Strength determination

Lab			#12153		#12154		Other details
	Distillation	Equipment	used sample in ml	how much distillate was obtained in ml	used sample in ml	how much distillate was obtained in ml	
311	--	--	--	--	--	--	
862	--	--	--	--	--	--	
1067	No / yes	--	--	--	100	30	
1126	yes	Densitymeter	100	200	200	200+100	
1241	yes	Densitymeter	100	210	125	210	
1242	yes	--	1	100	--	--	
1247	yes	Densitymeter	50 g	100	50 g	100	
1253	yes	Densitymeter	105	198	107	205	
1605	yes	Densitymeter	200	200	175	175	
1726	--	--	--	--	--	--	
1727	--	--	--	--	--	--	

APPENDIX 3

Number of participants per country

1 lab in BELGIUM
1 lab in P.R. of CHINA
2 labs in SPAIN
7 labs in THE NETHERLANDS

APPENDIX 4

Abbreviations:

C	= final result after checking of first reported suspect result
D(0.01)	= outlier in Dixon's outlier test
D(0.05)	= straggler in Dixon's outlier test
G(0.01)	= outlier in Grubbs' outlier test
G(0.05)	= straggler in Grubbs' outlier test
DG(0.01)	= outlier in Double Grubbs' outlier test
DG(0.05)	= straggler in Double Grubbs' outlier test
E	= error in calculations
ex	= excluded from calculations
n.a.	= not applicable
U	= unit error
SDS	= safety data sheet

Literature:

1 i.i.s. Interlaboratory Studies, Protocol for the Organisation, Statistics & Evaluation, January 2010

2 W. Horwitz and R. Albert, J. AOAC Int., Vol. 79, 3, p. 589, (1996)

3 ASTM E178-02

4 ASTM E1301-03

5 ISO13528-05

5 ISO 5725-86

6 ISO 5725, parts 1-6, 1994

7 M. Thompson and R. Wood, J. AOAC Int, 76, 926, (1993)

8 W.J. Youden and E.H. Steiner, Statistical Manual of the AOAC, (1975)

9 IP 367/84

10 DIN 38402 T41/42

11 P.L. Davies, Fr. Z. Anal. Chem, 331, 513, (1988)

12 J.N. Miller, Analyst, 118, 455, (1993)

13 Analytical Methods Committee Technical brief, No 4.January 2001

14 The Royal Society of Chemistry 2002, Analyst, 2002, 127, page 1359-1364, P.J. Lowthian and M. Thompson. (see http://www.rsc.org/suppdata/an/b2/b205600n/)