Interlaboratory Comparison Measurement EstOil-2

Final Report

03.11.2006

Report Compiled by

Lauri Jalukse

Ivo Leito

Tiina Kukk

Testing Centre, University of Tartu

Werol Tehased AS

Tartu 2006

Table of Contents

1 The Aim of the Intercomparison

The aim of the EstOil-2 intercomparison was to allow the participating laboratories to assess their performance in determining four edible oil parameters: moisture content, free fatty acids (below FFA) content, peroxide value (below PV) in refined rapeseed oil and phosphorus (below P) content in crude rapeseed oil. This is the second intercomparison of this series. The previous one (EstOil-1[1]) took place in 2005.

2 Organization

2.1 General

The intercomparison measurement was organized jointly by University of Tartu Testing Centre (below UT) and Werol Tehased AS (below WT). See for the detailed contact information of organizers in Table 1.

Table 1. Detailed Contact Information of Organizers.

University of Tartu, Testing Centre http://www.ut.ee/katsekoda Jakobi 2, 51014 Tartu, Estonia Phone: +372 51 84 176 Fax: +372 737 5264 E-mail: lauri.jalukse@ut.ee	Werol Tehased AS Painküla 48331, Jõgevamaa, Estonia Phone: +372 77 68 234 Fax: +372 77 68 220 E-mail: tiina.kukk@werol.ee

The Final report was compiled jointly by UT and WT and will be made public (on the website of UT). The participants are listed in the final report but the results are presented in random order, so that the results cannot be traced back to the participants. Every participant will receive a private letter revealing his/her result number and permitting assessment of performance.

2.2 The Samples

The oil samples were prepared and distributed by WT. The samples were refined rapeseed oil and crude rapeseed oil samples of approximately 120 ml in sealed glass ampoules. The samples were prepared from a single bulk of oil that was well mixed before sealing into ampoules. The ampoules were filled and closed during a short time (around three minutes per ampoule) after that the ampoules were immediately wrapped into dark polyethylene film. The laboratories got random ampoules from the pool of ampoules. The first and last ampoules were not distributed.

2.3 Data Treatment

The evaluation of participant data was done at UT according to the ISO Guide 43-1.[2] The z-score approach was used. The z-score for a particular measured value of a participant is calculated according to the following equation:

[1] http://www.ut.ee/katsekoda/ILC/Estoil/Estoil_1_rep_Final.pdf

[2] ISO Guide 43-1 *Proficiency Testing by Interlaboratory Comparisons. Part 1: Development and Operation of Proficiency Testing Schemes*, ISO/IEC 1997.

$$z = \frac{x - x_c}{s}, \quad (1)$$

where x is the participant's value, x_c is the consensus value and s is the target standard deviation. The consensus values are found as the mean values (after elimination of possible outlier results using the Grubbs test). The target standard deviation in our case is found as the real standard deviation of the participant values (after elimination of possible outliers).

Absolute values of z-scores (|z| values) are used for assessing the acceptability of the results as described in Table 2.

Table 2. Assessemnt of Acceptability of the Results Using z-Scores.

|z| Value	**Acceptability of the Result**	**Required Action**
$\|z\| \leq 2$	Acceptable result	No action is required
$2 < \|z\| < 3$	Doubtful result	Preventive action is required
$\|z\| \geq 3$	Unacceptable result	Corrective action is required

3 Participants

Invitations were sent to a number of laboratories in Estonia and other countries from Eastern Europe. The participants are listed in Table 3.

Table 3. Participants to EstOil-2.

Institution	**Country**
University of Tartu, testing Centre	Estonia
Werol Tehased AS	Estonia
Tallinn Veterinary and Food Laboratory	Estonia
Tartu Veterinary and Food Laboratory	Estonia
Agricultural Research Centre	Estonia
Latvija Latvenergo, Riga Thermal Power Plants Chemical laboratory	Latvia
Petrol d.d., Ljubljana – Laboratory Petrol	Slovenia
Laboratory for Fermentation Technologies and Refrigeration in Food Industry	Romania
National Food Investigation Institute Analytical Division	Hungary
General Inspectorate for Consumer Protection Laboratory of Foodstuffs and Chemicals	Hungary
Department of Food Chemistry, Lipid Laboratory	Hungary

4 Results

4.1 Results of the Participants

Results of the participants are presented in Table 4. In most cases the results are presented the same way as provided by the participants. In certain cases participants were asked to convert their values, so that uniformity of units was achieved.

Table 4. Participant Results together with the Expanded Uncertainties and the Derived Consensus Values.

	Moisture Content			Free Fatty Acid Content		
Lab number[a]	Result[b] ppm[d]	Uncertainty ppm	k[c]	Result[b] %	Uncertainty %	k[c]
1				0.0145	0.005	2
2				0.0130	0.001	2
3				0.0440	0.024	2
4	96	33	2	0.0290	0.006	2
5	76.3	8.1	2			
6	106	10.1	2			
7	60	20	2	0.0300	-	
8	100	50	2	0.0277	0.003	2
9				0.0300	-	
10				0.0350	0.001	2
11				0.0440	0.004	2
Consensus value	**87.7**			**0.030**		

	Peroxide value			Phosphorus Content in crude rapeseed oil		
Lab number[a]	Result[b] meq/kg	Uncertainty meq/kg	k[c]	Result[b] ppm[d]	Uncertainty ppm	k[c]
1	5.7385	0.0571	2			
2	7.25	-				
3	6.44	0.7	2			
4	5.182	2.848	2	129.7	16.5	2
5						
6						
7	62.5	-		153	-	
8	8.5	0.85	2			
9	5.71	-		90	-	
10	10.65	0.53	2			
11	5.745	0.5745	2			
Consensus value	**6.90**			**124.2**		

[a] The participating laboratories are given numbers in random order that is different from the order given in Table 3. [b] The results are presented the same way as they were presented by the participants. [c] Coverage factor. [d] Although ppm is not an SI unit, it was decided to present the results in this unit respecting the established practice. ppm is identical to mg/kg.

Before calculating the consensus values and target standard deviations for calculating the z-scores Grubbs tests were performed to check for outlying laboratories. One laboratory was outlying in PV measurement (see Table 6). The consensus values were derived from the participant data as average values: 87.7 ppm (moisture content) and 0.030 % (FFA content), 6.90 meq/kg (PV) and 124.2 ppm (P content). The standard deviations of the participant data were used as the target standard deviations: 19.1 ppm (moisture content), 0.01 % (FFA content), 1.9 meq/kg (PV) and 31.9 ppm (P content). Full information about consensus values and target standard deviations is presented in Table 5 (the respective data of the EstOil-1 intercomparison is also given for reference).

Table 5. Information about Consensus Values and Target Standard Deviations of Interlaboratory Comparison Measurements EstOil-2 and EstOil-1.

Moisture content	FFA content	PV	P content	
ppm	%	meq/kg	ppm	unit
		EstOil-2		
87.7	0.03	6.9	124.2	consensus value
19.1	0.01	1.9	31.9	target standard deviation
6.4	0.0003	0.03	1.9	between-sample variability
22%	37%	27%	26%	relative target standard deviation
7.3%	1.1%	0.4%	1.5%	relative between-sample variability
		EstOil-1		
362.5	0.07			consensus value
32.9	0.01			target standard deviation
6.0	0.0004			between-sample variability
9%	19%			relative target standard deviation
1.7%	0.5%			relative between-sample variability

The z-scores are calculated according to equation 1 and are presented in Table 6.

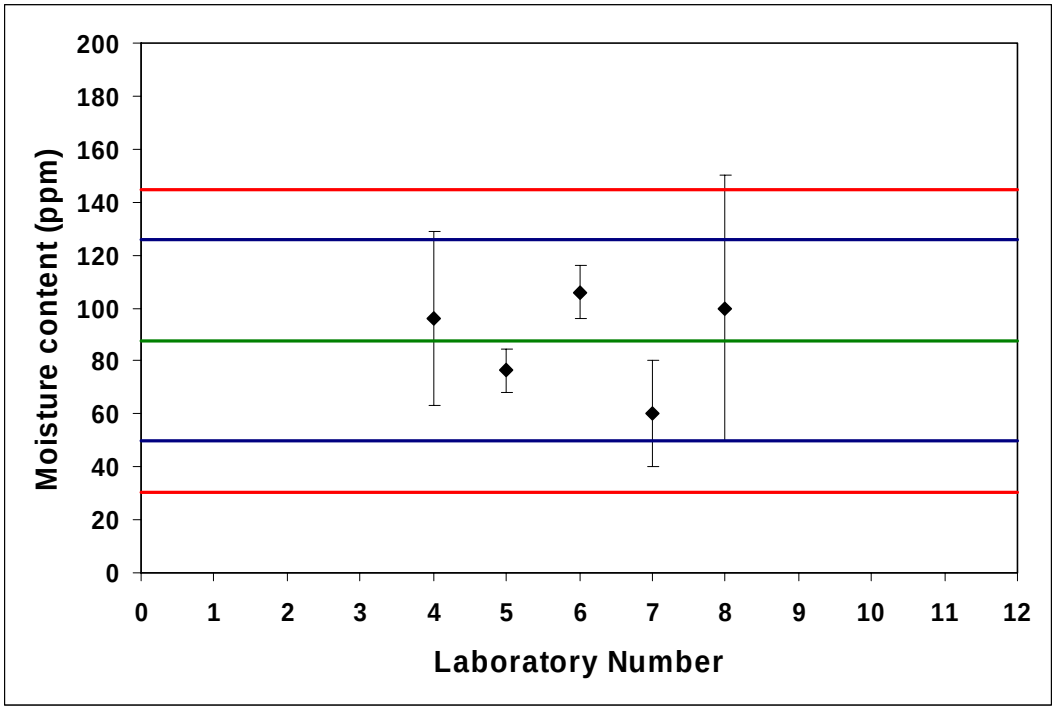
Table 6. Participant z-Scores.

	z scores			
Lab number[a]	Moisture content	FFA content	PV	P content
1		-1.4	-0.6	
2		-1.5	0.2	
3		1.3	-0.2	
4	0.4	-0.1	-0.9	0.2
5	-0.6			
6	1.0			
7	-1.5	0.0	30.0[b]	0.9
8	0.6	-0.2	0.9	
9		0.0	-0.6	-1.1
10		0.5	2.0	
11		1.3	-0.6	

[a] The participating laboratories are given in random order that is different from the order given in Table 3 but is identical to the order given in Table 4. [b] Outlying result according to the Grubbs test.

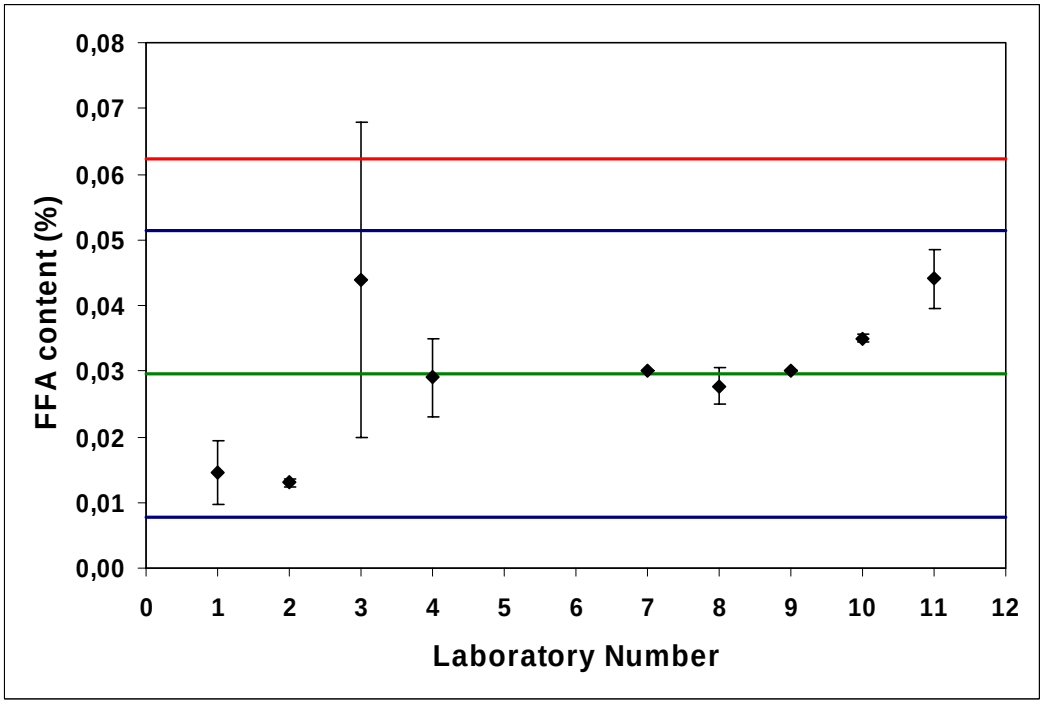
The results are presented in graphical form in the Figures below:

Figure 1. Results of Participants with the z-Score Boundaries.[a] Moisture Content Measurement.



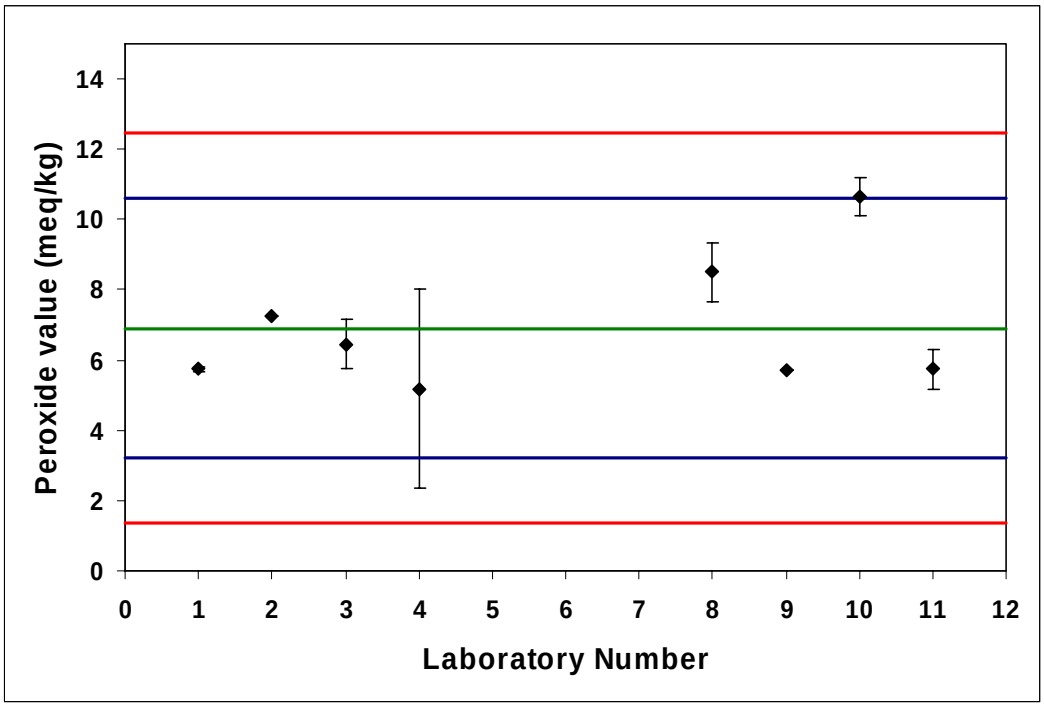
[a] The consensus value is denoted by a green line. The 2*s* boundaries are denoted by blue lines. The 3*s* boundaries are denoted by red lines.

Figure 2. Results of Participants with the z-Score Boundaries.[a] FFA Content Measurement.



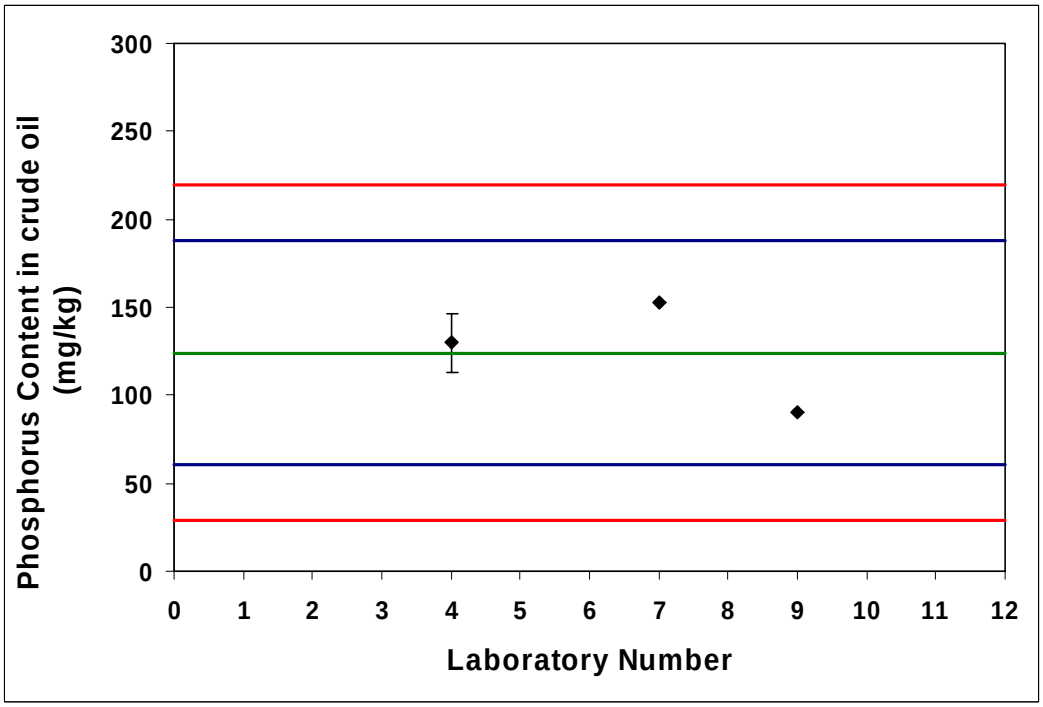
[a] The consensus value is denoted by a green line. The 2*s* boundaries are denoted by blue lines. The 3*s* boundaries are denoted by red lines.

Figure 3. Results of Participants with the z-Score Boundaries.[a] PV Measurement[b].



[a] The consensus value is denoted by a green line. The 2*s* boundaries are denoted by blue lines. The 3*s* boundaries are denoted by red lines. [b] There is one outlier in the PV measurement. This result is out of the figure range (the result is 62.5 meq/kg).

Figure 4. Results of Participants with the z-Score Boundaries.[a] P Content Measurement.



[a] The consensus value is denoted by a green line. The 2*s* boundaries are denoted by blue lines. The 3*s* boundaries are denoted by red lines.

From Table 6 and the Figures it can be concluded that based on the z-score approach only one participant in one parameter determination (PV) has not performed satisfactorily.

4.2 Between-Sample Variability

Between-sample variability was determined by Werol Tehased AS (moisture content, FFA content, PV and P content) and by University of Tartu (moisture content only) under the repeatability conditions. The data were treated using the ANOVA technique to separate the effects of between- and within-sample variability.[3]

For the FFA content the between-sample variability standard uncertainty was found 0.0003% (four ampoules), which is 33 times lower than the between-participant standard deviation. For the PV the between-sample variability standard uncertainty was found 0.03 meq/kg (five ampoules), which is 62 times lower than the between-participant standard deviation. For the phosphorus content the between-sample variability standard uncertainty was found 1.9 ppm (three ampoules), which is 17 times lower than the between-participant standard deviation. For the Humidity content the between-sample variability standard uncertainty found by Werol Tehased AS was 7.2 ppm (four ampoules). For the results obtained at UT Testing Centre the same is 2.7 ppm (two ampoules). From these two datasets the pooled standard deviation is found as 6.4 ppm. This is 3 times lower than the between-participant standard deviation.

We can conclude that in phosphorus content, FFA content and PV measurement the between-sample variability has negligible effect on the between-participant variability. The same difference in moisture content measurement is 3 times. Therefore with this parameter the between-sample variability is taken into account in E_n score calculation (see below). It is also necessary to note that the difference between the mean moisture contents obtained at UT Testing Centre and Werol Tehased AS was around 20 ppm indicating significant laboratory bias also between the organizers.

5 Discussion

Most of the participant results have z scores between -2 and 2. However, this is primarily caused by using the relatively large actual standard deviations as target standard deviations in z score calculation. The spread of the participant values is significant. As indicated by the between-ampoule variability determinations. The spread is first of all caused by large laboratory biases. These could be due to the following factors:

1. The contents of moisture and FFA in the samples were low. This is probably one of the main reasons for the large spread of FFA content measurement (measuring near of LoQ). The target standard deviation is larger than in EstOil-1 (see Table 5).
2. It is clear that PV is very unstable measurand. In spite of this, the spread of PV measurement results is quite normal (compared to the other parameters, see Table 5). One result in PV measurement was identified as an outlier (according to the Grubbs test). The extremely high value 62.5 meq/kg (nine times higher than the consensus value) obtained by the participant cannot be caused by rancidity of the sample. Independent monitoring in our laboratory has established that even in an ordinary plastic bottle (not hermetically sealed ampoule) the PV value increases by only around 3 meq/kg per year. It is impossible that the PV value in a sealed ampoule increases by more than 50 meq/kg during 1-2 months. Even non-observance of the latest day (June 17, 2006) for PV measurement cannot lead to such dramatic PV increase.

[3] The treatment was carried out as described in A.M.H., van der Veen, J. Pauwels, *Accred. Qual. Assur.*, **2000**, *5*, 464-469.

3. Different methods were used by different participants. This can have an effect on the agreement of the participant results. However, this year we did not observe any trends or systematic differences between results of laboratories comparing these methods. The results of participants using identical methods often differed very much. In particular, results of moisture determination by loss on heating are notorious in not agreeing with KF titration. Out of the 5 laboratories who determined moisture content one used loss on heating and their result agrees very well with the remaining four results (obtained by KF titration).
4. Sample-to-sample variability plays probably insignificant role (see previous section), with 3 out of 4 parameters (the exception is the moisture content) however, it cannot be ruled out. Also, several participants reported that it was difficult to open the sample ampoules and the ampoules could not be conveniently closed once they were opened. Difficulties in handling the ampoules can cause some moisture uptake from the atmosphere.

Although according to the z-score approach all the participants performed satisfactorily, it is of interest to compare the self-declared uncertainties of the participant results to the agreement between the results. **It can easily be seen from such a comparison that most probably the majority of the participants have underestimated their uncertainties in FFA and PV determination. Some uncertainty underestimation is probably present also in moisture determination.** The lacking uncertainty data does not permit to make any conclusions on phosphorus content.

For example, in FFA content measurements the uncertainty intervals of laboratories 1, 8, 10 and 11 do not overlap in any of the two-lab pairs (!). This means that at least three of the four laboratories have underestimated the uncertainties of their results!

Very illuminating in this respect are comparisons of the results of the participants in pairs. This is done using the E_n numbers as described in ISO Guide 43-1:[2]

$$E_n = \frac{C_{lab1} - C_{lab2}}{\sqrt{U_{lab1}^2 + U_{lab2}^2}} \ . \qquad (2)$$

where C_{lab1} and C_{lab2} are the results of the two laboratories that are compared and U_{lab1} and U_{lab2} are their expanded uncertainties. Agreement between two results is considered acceptable if $|E_n| \leq 1$. The paired comparisons are presented in Table 7, Table 8 and Table 9. No reasonable comparison is possible with phosphorus content measurement, because only two (!) participants have supplied uncertainties for the results.

As we see in Figure 2 most of the laboratories have seriously underestimated the uncertainties of their FFA content results. Contrary to FFA content measurement, the laboratories have given rather realistic uncertainties of their moisture content results. Comparing the present results to those of EstOil-1 (see Table 5) one can see that the uncertainty estimates are generally larger this time indicating improved understanding of the measurement uncertainty.

Table 7. Comparison of the Results of Participants in Pairs: FFA Content Measurement.

	E_n value						
Lab No	1	2	3	4	8	10	11
1	0	-0.3	1.2	1.9	2.3	4.2	4.5
2	0.3	0	1.3	2.7	5.2	26.3	7.0
3	-1.2	-1.3	0	-0.6	-0.7	-0.4	0.0
4	-1.9	-2.7	0.6	0	-0.2	1.0	2.0
8	-2.3	-5.2	0.7	0.2	0	2.6	3.1
10	-4.2	-26.3	0.4	-1.0	-2.6	0	2.0
11	-4.5	-7.0	0.0	-2.0	-3.1	-2.0	0

The numbers of participants are the same as in Table 4.

Table 8. Comparison of the Results of Participants in Pairs: PV Measurement.

	E_n value					
Lab No	1	3	4	8	10	11
1	0	1.0	-0.2	3.2	9.2	0.0
3	-1.0	0	-0.4	1.9	4.8	-0.8
4	0.2	0.4	0	1.1	1.9	0.2
8	-3.2	-1.9	-1.1	0	2.1	-2.7
10	-9.2	-4.8	-1.9	-2.1	0	-6.3
11	0.0	0.8	-0.2	2.7	6.3	0

The numbers of participants are the same as in Table 4.

Because the between sample variability uncertainty of moisture content measurement is relatively high, it is necessary to use a modified equation (that allows to take the additional variability into account) to calculating E_n values for moisture content:

$$E_n = \frac{C_{lab1} - C_{lab2}}{\sqrt{U_{lab1}^2 + U_{lab2}^2 + (t_{95}(4) \cdot s_s)^2}} \quad . \tag{3}$$

where s_s – is the pooled standard uncertainty of between-sample variability of moisture content and $t_{95}(4)$ is the student coefficient at 95% confidence level with 4 degrees of freedom.

Table 9. Comparison of the Results of Participants in Pairs: Moisture Content Measurement.

	E_n value				
Lab No	4	5	6	7	8
4	0	-0.5	0.3	-0.8	0.1
5	0.5	0	1.4	-0.6	0.4
6	-0.3	-1.4	0	-1.6	-0.1
7	0.8	0.6	1.6	0	0.7
8	-0.1	-0.4	0.1	-0.7	0

The numbers of participants are the same as in Table 4.

As can be seen from the E_n values 8 out of 10 pairs display satisfactory agreement in moisture content measurement. At the same time only 7 out of 21 in FFA content measurement and also only 6 out of 15 in PV measurement are in agreement. Clearly, most of the participants should take a close look at their uncertainty estimates of FFA and PV measurements. The abovementioned factors provide some guidelines.

An alternative and more rigorous way to assess the performance of the participants would be to compare their results to independent (metrological) reference values of the measurands. This is, however, impossible in the current intercomparison, because the consensus values of the measurands obtained from the participant data are not trustworthy enough.

The general conclusions from the results can thus be worded as follows:

1. **Given the low values of the measured parameters in the samples the agreement of the obtained values, leaving aside their claimed uncertainties, can be considered satisfactory.**

2. **The uncertainties of the results of most participants have been underestimated.**