Results of the inter-laboratory comparison exercise for TC and EC measurements (ref.: OCEC-2016-2)

Summary

The European Centre for Aerosol Calibration (ECAC) under ACTRIS-2 completed (October 2016) an inter-laboratory comparison for the measurement of total carbon (TC), elemental carbon (EC) and organic carbon (OC) in particulate matter collected on filters. The aim of this comparison was to evaluate the performances of the measurement method (i.e. reproducibility and repeatability) and of individual laboratories (biases).

This exercise was based on ambient PM2.5 and PM10 aerosol samples collected on quartz fiber filters at regional background sites in Norway, Germany, and Spain, and at one urban background station in Greece. A solution of phthalic acid prepared at the JRC (the inter-laboratory comparison exercise coordinator) was also distributed.

Thirteen laboratories participated in this exercise running their usual thermal-optical EUSAAR_2 protocol with their usual analytical instrument. Among those, ten are responsible for the aerosol chemical speciation at the EMEP or ACTRIS stations located in their countries (i.e. Belgium, Czech Republic, France, United Kingdom, Slovenia, The Netherlands, Sweden, Switzerland, Denmark and Italy).

Measurement method performance: for TC determination, repeatability and reproducibility relative standard deviations ranged from 2% to 4% and from 3.5% to 13%, respectively. For the determination of the EC/TC ratio, repeatability and reproducibility relative standard deviations ranged from 3% to 14% and from 15% to 66%. Repeatability and reproducibility standard deviations show an inverse dependence on TC loadings and on EC/TC ratios becoming exponentially poorer toward lower TC contents i.e. <13 µgC / cm² and EC/TC ratio. i.e. <0.07, respectively.

Still in absence of a suitable certified reference material for atmospheric OC and EC, the assigned values for TC loadings and EC/TC ratios in the test samples were calculated as the robust average values among all participants. The assigned value for the concentration of phthalic acid was determined from primary gravimetric and volumetric measurements.

Laboratory performance: for both TC loadings and EC/TC ratios, laboratories' performances were assessed in terms of z-scores, calculating the *standard deviation for proficiency assessment* (σ*) *from the data obtained in the round of the proficiency testing scheme.*
For TC loadings, fifteen outliers and four stragglers were identified; and 78% of all entries were within 10% from the assigned TC concentration value.
Regarding EC/TC ratios, the majority of the twenty outliers and six stragglers were produced by four participants. Only 40% of all entries is within 10% of the assigned value and 73% were within the 25% of the assigned EC/TC ratio.

Although the contribution of localized sample heterogeneities and/or contaminations to biased data cannot be totally excluded, the random scheme adopted to distribute sub-samples was such that the recurrence of stragglers or outliers (more than two) for single laboratories most probably indicates an unsatisfactory laboratory performance as compared to the other participants. Laboratories showing unsatisfactory precision (both in terms of repeatability and reproducibility) or significant biases for several test samples shall carefully examine their operating procedures and instrumental set-up and identify appropriate corrective actions with the help of ECAC staff if needed.

Introduction

Total carbon (TC), including Organic Carbon (OC) and Elemental Carbon (EC) is a relevant constituent of the fine fraction of particulate matter (PM), both from the perspective of health risks due to inhalation and indication of air pollution sources. For these reasons requirements for measuring EC and OC in $PM_{2.5}$ at rural background locations have been included in Air Quality Directive 2008/50/EC.

The directive states that measurements should be made in a manner consistent with those of the cooperative programme for monitoring and evaluation of the long range transmission of air pollutants in Europe (EMEP). Thermal-optical analysis has been recognized as the most suitable method for the determination of EC and OC collected on filters (see Technical Specification by the CEN/TC265 WG35) and the thermal protocol EUSAAR-2 with a transmittance optical correction for pyrolysis -already adopted in the EMEP manual for sampling and analysis- has been recently selected as the European standard thermal protocol (FprEN16909).

The *European center for aerosol calibration* within the European project ACTRIS-2 has organized in September-October 2016 an inter-laboratory comparison exercise (ILCE) (ref. OCEC-2016-2) among thirteen applicants including, among others, laboratories in charge of OC and EC measurements at EMEP/ACTRIS station in Belgium, Czech Republic, France, United Kingdom, Slovenia, The Netherlands, Sweden, Switzerland, Denmark and Italy.

1 Organization

1.1 Samples, sub-samples and sub-sample homogeneity

In lack of certified reference material for atmospheric OC and EC, this ILCE made use of ambient (outdoor) PM aerosol collected with high-volume samplers on quartz fiber filters at 4 sites across Europe (Table 1). Upon receipt at JRC, filters were stored in a refrigerator.

Table 1: filter test samples used for the inter-laboratory comparison

Sampling location				Sample collection		
Station	Country	Symbol	Site type	Period	Size fraction	Filter type
Melpitz	Germany	D	rural	Winter 2015	PM2.5	Munktell MK360
Demokritos	Greece	G	urban	Winter 2015	PM2.5	MCV-QF1
Montseny	Spain	E	rural	Winter 2015	PM2.5	Whatman QMH
Birkenes	Norway	N	rural	Apr./May 2000	PM10	Whatman QMA

Rectangular filter punches of ca. 3.6 cm x 1.8 cm were randomly distributed to participants to allow them to triplicate measurements.

The homogeneity of these test samples was investigated by the JRC on distinct filter samples collected at each location. From each sample, ten subsamples of 1 cm^2 were taken along two perpendicular axes across the filter surface and analysed for their TC, OC and EC contents. The

filter homogeneity was assessed as the standard deviation of the average of the 10 replicate analyses. This leads to an upper limit for the filter homogeneity since it includes the repeatability of the JRC laboratory (< 3 and 6% for TC and EC, respectively). The homogeneity is better than 6 and 11% for TC and EC/TC, respectively (Table 2). If sampling at each location occurred under repeatable conditions, it can be assumed that the test samples had similar homogeneities.

Table 2: homogeneity of the deposits on filters collected with the samplers used to produce the eight test filters. Analyses were performed with the protocol EUSAAR_2 and charring correction by transmittance monitoring.

Test sample and origin	Homogeneity for TC (%)	Homogeneity for EC/TC (%)
48_ _ MEL (D)	2.5	4.0
DEM _ DEM(G)	5.7	7.6
T66_ _ Mont (E)	3.9	10.2
A2__ BIR (N)	4.7	11.0

An aqueous solution of phthalic acid was also distributed to the participants to assess the uncertainty of the instrument calibration constant determination. The solution was prepared by dissolving a precisely known mass of pure phthalic acid (≥ 99.5%) in a precisely known volume of ultra-pure water (resistivity ≥18.2 mΩ cm).

1.2 Participants

Participants were selected among applicants to ECAC choosing (in the interest for the scientific community) in a first place laboratories which submit TC and EC data to the EBAS database and laboratories which could most benefit from the outcome of this exercise in term of their expertise development.
The list of the thirteen participants is reported in Table 3. For brevity, the number assigned to each participant will be used in the remainder of the document.

1.3 Sample shipment and reporting of results

Test samples were shipped to all participants (except "local" participant 15) on 31st Aug. 2016 via postal mail at ambient temperature without temperature record in closed petri dishes. Participants were asked to report TC and EC concentration, in μg C cm^{-2} units with three decimal digits, from three replicates of test ambient PM samples, by the end of October 2016. In addition, participants were asked to report the OC content of 10 μl of a phthalic acid solution precisely prepared and traceable to primary measurements.

1.4 Thermal-optical analysis

The thermal protocol EUSAAR-2 [Cavalli et al., 2010] with a transmittance optical correction for pyrolysis has been recently selected as the European standard thermal protocol for the measurements of TC, OC and EC in PM samples (FprEN16909).
In this exercise all laboratories applied the EUSAAR_2 protocol (Table 5) with transmittance-based correction.

Table 3: List of participants in the inter-laboratory comparison 2015, and contact persons

Code	Participant	Acronym	Contact
1	FONSNY Eric	ISSeP	e.fonsny@issep.be
2	Szidat, Sönke	LARA_UniBe	szidat@dcb.unibe.ch
3	Lewandowska Anita	UGPoland	a.lewandowska@ug.edu.pl
4	Holubová Šmejkalová, Adéla	Kresin	Holubovasmejkalova.a@czechglobe.cz
5	BONNAIRE Nicolas	LSCE	nicolas.bonnaire@lsce.ipsl.fr
6	Maria Rita Perrone	UniSalento	perrone@le.infn.it
7	Quincey Paul	NPL	paul.quincey@npl.co.uk
8	Burger Judita	ARSO-KAL	judita.burger@gov.si
9	Panteliadis, Pavlos	GGD Amsterdam	ppanteliadis@ggd.amsterdam.nl
10	Martinsson, Johan	ULUND	johan.martinsson@nuclear.lu.se
11	Fischer Andrea and Hueglin Christoph	EMPA	andrea.fischer@empa.ch christoph.hueglin@empa.ch
13	Nøjgaard, Jacob, Klenø	ENVS-AU-Dk	jakn@envs.au.dk
15	Fabrizia Cavalli	EC-JRC	fabrizia .cavalli@jrc.ec.europa.eu

Table 4: List of the analytical protocol and punch size used by each participant

Code	Participant	Instrument	Protocol	Punch size (cm^2)
1	ISSeP	Sunset	EUSAAR_2	1.5
2	LARA_UniBe	Sunset	EUSAAR_2	1.5
3	UGPoland	Sunset	EUSAAR_2	1.5
4	Kresin	Sunset	EUSAAR_2	1.27
5	LSCE	Sunset	EUSAAR_2	1.5
6	UniSalento	Sunset	EUSAAR_2	1.176
7	NPL	Sunset	EUSAAR_2	1.5
8	ARSO-KAL	Sunset	EUSAAR_2	1.5
9	GGD Amsterdam	Sunset	EUSAAR_2	1.5
10	ULUND	DRI	EUSAAR_2	0.5
11	EMPA	Sunset	EUSAAR_2	1.5
13	ENVS-AU-Dk	Sunset	EUSAAR_2	1.5
15	EC-JRC	Sunset	EUSAAR_2	1

Table 5: Details of the analytical protocol implemented by all participants

	EUSAAR_2	
Carrier gas	Time (s)	Temp. (°C)
Helium	120	200
Helium	150	300
Helium	180	450
Helium	180	650
Helium		
Oxygen in Helium (2%)	120	500
Oxygen in Helium	120	550
Oxygen in Helium	70	700
Oxygen in Helium	80/110	850
Oxygen in Helium		
Oxygen in Helium		

2 Data evaluation

Ambient PM filter samples: In absence of certified reference material for atmospheric TC, OC and EC deposited on filters, the *measurement method performance* (par. 2.1) and *laboratory performances* (par. 2.2) were evaluated using atmospheric PM collected on filters as test samples.

In this report we focus on the *TC loadings* (in µg cm^{-2}) and *EC/TC ratios* reported by each participants for each test sample. TC represents the most robust (and protocol-independent) output of TOA analyses, while EC/TC ratios are free from biases in the total carbon determination, and reflect possible differences in the OC/EC split determination among participants.

On average, reported TC loadings ranged from 5 to 23 µg cm^{-2}, corresponding to atmospheric concentrations ranging from 1 to 6 µg m^{-3} collected for 24h at a face velocity of 54 cm s^{-1}. EC/TC ranged on average from 0.03 to 0.14. All submitted results for TC (in µg cm^{-2}) and EC/TC ratio are presented in Annex 1 in Tables 1 and 2, respectively.

Aqueous solution of phthalic acid: This solution was used to assess the uncertainty of the instrument calibration constant determination. Results were analysed in terms of percentage differences from the assigned value.

Assigned values:

As *ambient PM collected* on filters was used as test samples, the true values for *TC and EC/TC loadings* were not known. The assigned value and its standard uncertainty for TC and EC/TC ratio on each filter were calculated as the robust average among all participants (see Par 2.2).

For the *phthalic acid solution,* the assigned OC concentration value was calculated from the water volume used to make the solution, the mass of phthalic acid dissolved in this water volume, and the chemical formula of phthalic acid. The assigned value was 1.52 gC l^{-1} (traceable to primary measurements) with an expanded combined relative uncertainty (k = 2) of 1.0%.

2.1 TEST FILTER SAMPLES - Method performance

2.1.1 Data evaluation description

The assessment of the *method performance* aims at deriving, from the results of the present exercise, the precisions of the measurement method in terms of repeatability and reproducibility standard deviations. For this, the consistency of the dataset is evaluated, at first graphically, by means of Mandel's *h* and *k* statistics [ISO5725-2] for possible outliers (i.e. observations greater than the critical value at the 99% confidence level) or stragglers (i.e. observations greater than the critical value at the 95% confidence level but less or equal to the critical value at the 99% confidence level).

[Note: Participant 4 reported two replicates for four test samples and only one replicate for the remaining four ones. The statistical methods applied to estimate the precision of the measurement method strictly require the same number (or as much as possible) of replicates per test sample per participants. Therefore, values produced by participant 4 could not be included in this evaluation.]

Mandel's *k* parameter estimates the within-laboratory consistency (repeatability). The critical values for *Mandel's k* indicators (i.e. outlier and straggler) vary upon the number of replicate measurements. In this comparison exercise, all laboratories (lab 4 excluded) provided three replicates for every sample (except lab 1 for sample T6691). Thus *Mandel's k* was calculated for an average case of three replicates and compared to the critical values 2.02 (outlier) and 1.69 (straggler).

Mandel's *h* parameter describes the between-laboratory consistency (reproducibility) and has been calculated for every laboratory and every sample. For an inter-laboratory comparison among twelve participants, the critical values for Mandel's *h* are 2.25 (outlier) and 1.83 (straggler).

To confirm the identified outliers and stragglers, statistical Cochran's and G_1 and G_2-Grubbs' test are applied for testing the within-laboratory and the between-laboratory variances, respectively [ISO5725-2].

Based on the outcomes of above statistical treatments (Grubbs' and Cochran's tests), outliers are discarded for the calculation of the mean value, the method repeatability and reproducibility standard deviations. Subsequently, the dependence of precision (i.e. repeatability and reproducibility) upon the mean values is investigated [ISO5725-2].

2.1.2 Results: Method performance for TC

Within-laboratory consistency. In Figure 1, the *Mandel's k* statistic values for TC are presented grouped by laboratory (Panel a) and, separately, by sample (Panel b).

Eight outliers (lab/sample: 6/4833; 6/4834; 3/DEM4; 11/DEM9; 5/T6690; 2/T6691; 10/A219; 5/A222) and two stragglers (lab/sample: 3/DEM9; 7/T6691) were identified (Fig. 1). Cochran's test confirmed as outliers (lab/sample) 6/4833, 6/4834, 11/4834 and 10/A219 and 3/DEM9 and 7/T6691 as stragglers.

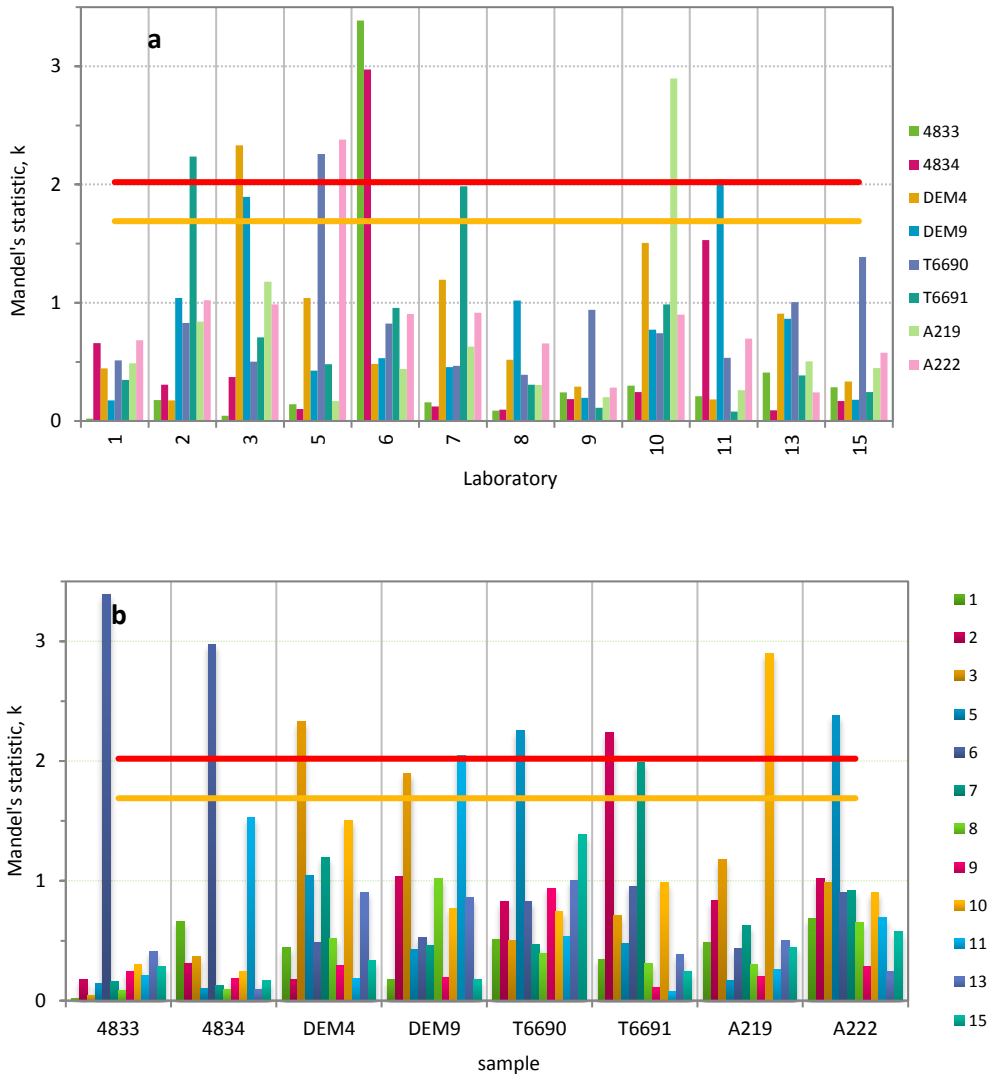


Figure 1. *Mandel's k* statistic values for within-laboratory consistency on TC data, grouped by laboratory (panel a) and by sample (panel b). For twelve laboratories and three replicates, k values should be < 2.02 at the 1% significance level (red line) and < 1.69 at the 5% significance level (orange line).

Between-laboratory consistency. In Figure 2, the *Mandel's h* statistic values are presented grouped for each laboratory (Panel a) and, separately, for each sample (Panel b).
In the TC dataset, six outliers (lab/sample: 6/4833; 6/4834; 3/DEM4; 3/DEM9; 3/T6691 and 6/A222) and three stragglers (lab/sample: 6/T6690; 10/T6690 and 6/A219) were identified (Fig. 2). The Grubbs' tests confirmed as outliers 6/4833, 6/4834, 3/DEM4, 6/DEM4, and 3/DEM9 and identified four stragglers 3/T6691, 6/T6691, 3/A222 and 6/A222.

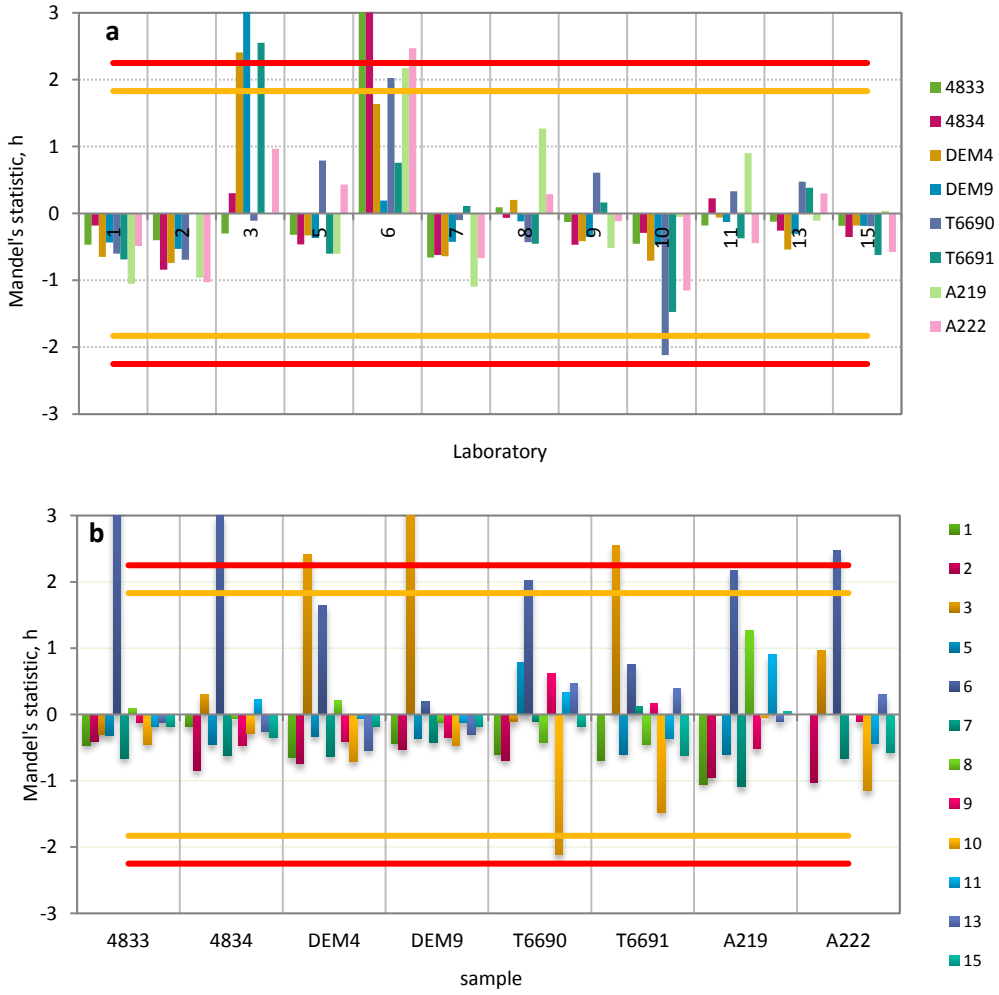


Figure 2. *Mandel's h* statistic values for between-laboratory consistency on TC data, grouped by laboratory (panel a) and by sample (panel b). For twelve laboratories, *h* values should be < 2.25 at the 1% significance level (red line) and < 1.83 at 5% significance level (orange line).

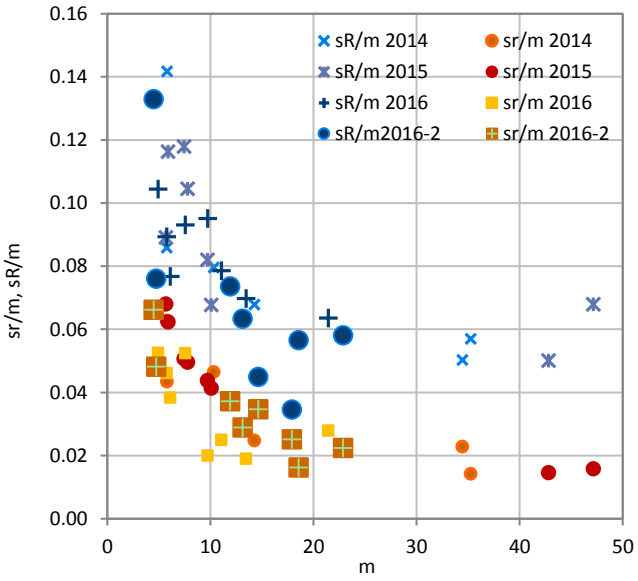
Localized sample heterogeneities or contaminations cannot rigorously be excluded, but the occurrence of several stragglers and/or outliers from a single laboratory (case of labs 3 and 6) most probably suggests unsatisfactory laboratory precision for the determination of the TC loadings as compared to the other laboratories.

From the retained values and for each sample separately, the mean value, the method repeatability (s_r) and reproducibility (s_R) standard deviations were calculated. The general means and values of s_r and s_R for the eight test filter samples are listed in Table 6. Both

repeatability and reproducibility relative standard deviations show an inverse dependence on TC. Combining the repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol obtained during the previous ILCEs and the present one, we observe that the method precision (both sR and sr) for TC measurement becomes exponentially poorer toward lower TC contents i.e. < 13 µgC / cm² (Fig. 3).

Table 6: General mean**,** repeatability (s_r) and reproducibility (s_R) standard and relative standard deviations for TC.

test sample	general mean	sr		sR	
	µgC / cm²	µgC / cm²	%	µgC / cm²	%
4833	14.7	0.5	3.5	0.7	4.5
4834	11.9	0.4	3.7	0.9	7.4
DEM4	17.9	0.5	2.5	0.6	3.5
DEM9	22.9	0.5	2.2	1.3	5.8
T6690	4.8	0.2	4.8	0.4	7.6
T6691	4.5	0.3	6.6	0.6	13.3
A219	18.6	0.3	1.6	1.0	5.7
A222	13.1	0.4	2.9	0.8	6.3



F**igure 3**. Repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol for TC measurement obtained during the previous inter-laboratory comparisons and the present one.

2.1.3 Results: Method performance for EC/TC

Within-laboratory consistency. In Figure 4 the *Mandel's k* statistic values are presented grouped for each laboratory (Panel a) and, separately, for each sample (Panel b). In the EC/TC dataset, six outliers (lab/sample: 6/4833, 6/T6694; 10/DEM9; 2/T6690; 6/T6691; 10/A219) and seven stragglers (lab/sample: 2/4833; 11/4834; 8/DEM4; 6/DEM9; 6/T6690; 2/A219 and 8/A222) were identified (Fig. 4). Cochran's test identified the entries 6/4833, 6/4834, 10/DEM9, 2/T6690, 6/T6691 and 10/A219 to be outliers and the entries 2/4833, 11/4834, 8/DEM4, 6/DEM9, 6/T6690, 2/A219, 6/T6690 and 8/A222 as stragglers.

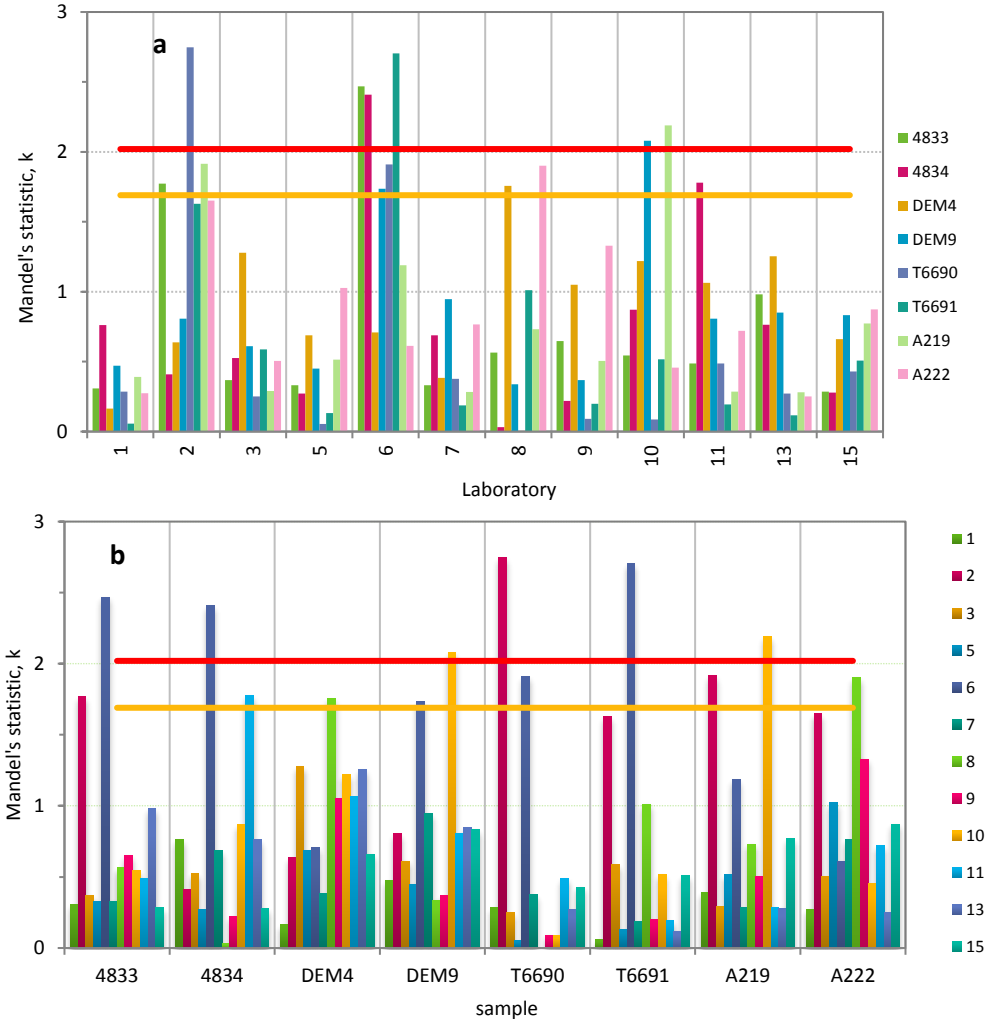


Figure 4. *Mandel's k* statistic values for within laboratory consistency on EC/TC ratio obtained from the entire database, grouped by laboratory (panel a) and by sample (panel b). For twelve laboratories and three replicates k values should be < 2.02 at 1% significance level (red line) and < 1.69 at 5% significance level (orange line).

Between-laboratory consistency. Figure 5 shows the *Mandel's h* statistic values for EC/TC ratio calculated on the entire database for each laboratory (Panel a) and, separately, for each sample (Panel b).

Seven outliers (lab/sample: 6/4833; 6/4834; 6/DEM4; 6/DEM9; 3/T6691; 6/A219; 6/A222) and no stragglers were identified. Grubbs' test identifies as outliers the entries 6/4833; 10/4833; 6/4834; 10/4834; 6/DEM4; 6/DEM9 and 6/A219 and as straggler the entry 6/A222.

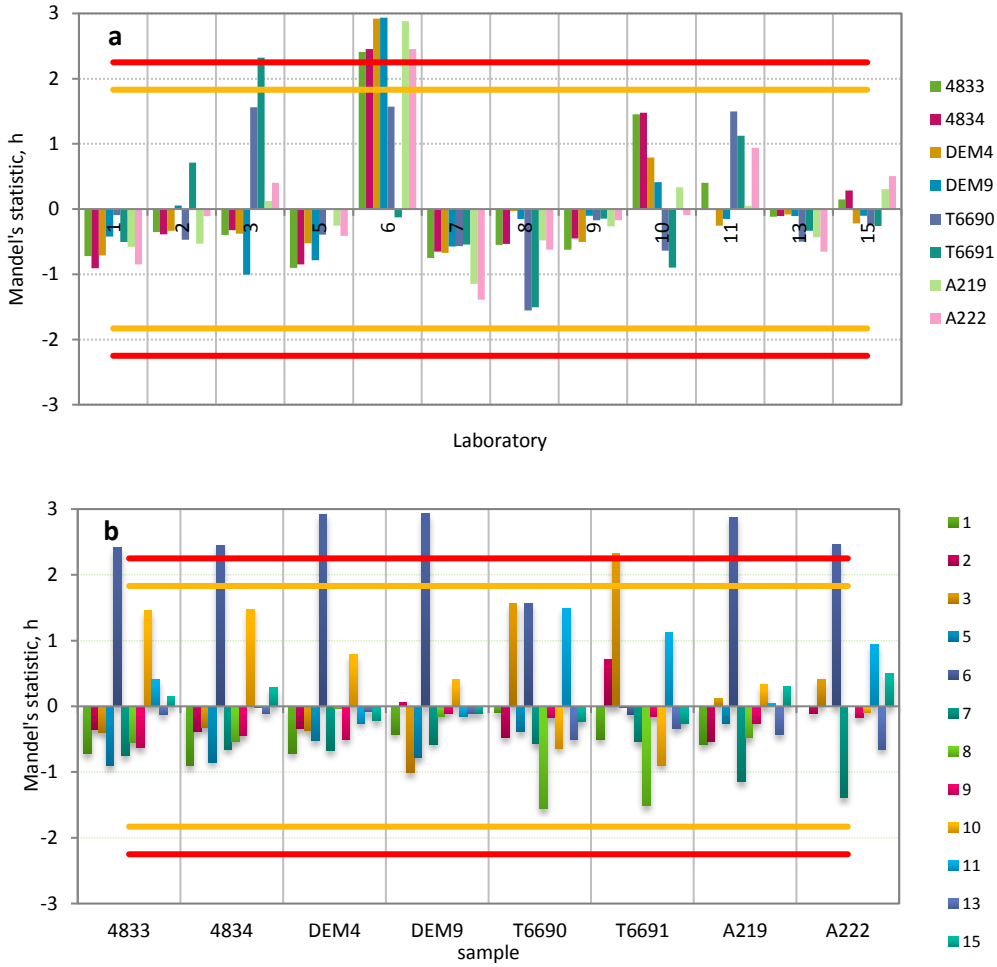


Figure 5. *Mandel's h* statistic values for between laboratory consistency on EC/TC ratio obtained from the entire database, grouped by laboratory (panel a) and by sample (panel b). For twelve laboratories h values should be < 2.25 at 1% significance level (red line) and < 1.83 at 5% significance level (orange line).

Localized sample heterogeneities or contaminations cannot rigorously be excluded, but the occurrence of several stragglers and/or outliers from a single laboratory (case of labs 6 and 10) most probably suggests unsatisfactory laboratory precision for the determination of the EC/TC ratio as compared to the other laboratories.

The entries identified as outliers by the statistical tests were discarded from the dataset, and the mean value, the repeatability (s_r) and the reproducibility (s_R) standard deviations for EC/TC were calculated for each sample from the retained values (Table 7). Combining the repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol obtained during the previous ILCEs and the present one, we observe that the method precision (both sR and sr) for

EC/TC ratio measurement becomes exponentially poorer toward lower EC/TC ratio i.e. < 0.07 (Fig. 6).

Table 7: General mean, repeatability (s_r) and reproducibility (s_R) standard and relative standard deviations for EC/TC.

test sample	general mean	sr		sR	
			%		%
4833	0.07	0.00	4.3	0.01	18.2
4834	0.08	0.00	4.5	0.01	15.4
DEM4	0.14	0.01	4.3	0.03	19.5
DEM9	0.12	0.00	3.4	0.02	18.3
T6690	0.03	0.00	10.2	0.02	66.5
T6691	0.06	0.01	14.3	0.04	57.8
A219	0.09	0.01	8.1	0.01	15.2
A222	0.08	0.00	4.4	0.02	19.4

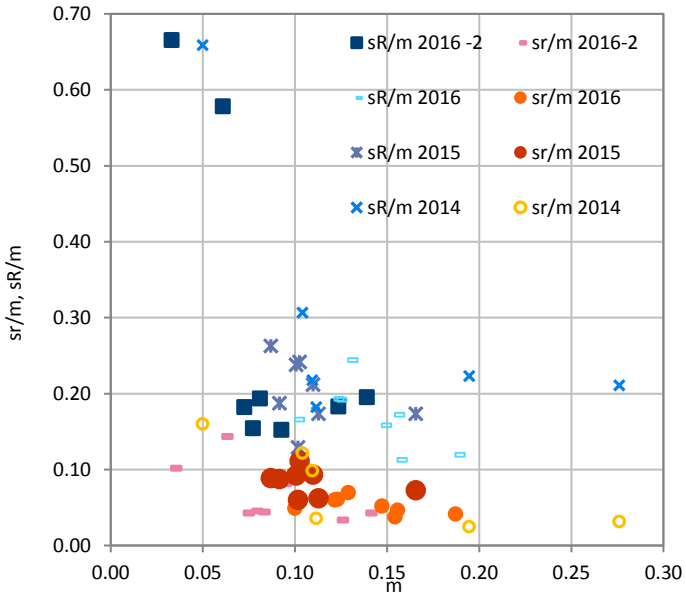


Figure 6. Repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol for EC/TC measurement obtained during the previous inter-laboratory comparisons and the present one.

2.2 FILTER TEST SAMPLES - Laboratory performance

2.2.1 Data evaluation description

The assessment of the *laboratory performance* aims at describing the laboratory bias compared to the assigned value associated with its standard deviation. Each participant's performance is determined in terms of *z-scores,* a measure of the deviation from the assigned value. To calculate *z-scores*, an assigned value and its standard deviation have to be determined for each test sample.

- *Determining the assigned value*: Among the available methods for determining the assigned value, the approach of the *consensus value from participants to a round of a proficiency testing scheme* was chosen, in absence of a reference or certified reference material. With this approach, the assigned value X for each test sample used in the ILCE is the robust average calculated, with a recursive algorithm, from the results reported by all participant (See ISO 13528:2005(E), Annex C).

- *Determining the standard deviation for proficiency assessment*: Among the available methods for determining the standard deviation for proficiency assessment (σ*), the approach of calculating σ* *from data obtained in a round of a proficiency testing scheme* was chosen. With this approach, σ* is the robust standard deviation calculated, with a recursive algorithm, from the results reported by all participant testing (See ISO 13528:2005(E), Annex C.

These approaches might become statically ineffective [ISO 13528:2015 (E)], for example, if the number of participant is lower than twenty. To verify their reliability the robust mean and its standard deviation were also calculated applying the Q/Hampel method (ISO 13528:2015 (E)). The obtained values do not significantly differ from those obtained by the *consensus value from participant results*, in Table 8, which are then used for the following elaboration.

For each laboratory and test sample, the *z-score* was calculated as:

$$z = (x_i - X)/ \sigma^*$$

where x_i is the result from the participant i; X is the assigned value for the sample; and σ^* is the standard deviation for proficiency assessment.

When a participant reports an entry that produces a bias greater than +3 z or less than -3 z (i.e. deviating from the assigned value for more than 3 standard deviations), this entry is considered to give an "action signal". Likewise, a laboratory bias above +2 z or below -2 z (i.e. deviating from the assigned value for more than 2 but less than 3 standard deviations) is considered to give a "warning signal". A laboratory bias between -2 z and +2 z indicates a satisfactory laboratory performance with respect to the standard deviation for proficiency assessment.

2.2.2 Results: Laboratory performance for TC

The assigned values X and the related standard deviations for proficiency assessment σ* calculated from the entire database for each sample, are reported in Table 8. Following ISO13528, σ* were calculated *from data obtained in a round of a proficiency testing scheme*.

Table 8: Assigned values and standard deviations for proficiency assessment σ* *from data obtained in a round of a proficiency testing scheme* for TC.

		4833	4834	DEM4	DEM9	T6690	T6691	A219	A222
assigned value	µg/cm2	14.7	12.1	18.1	22.9	4.8	4.4	18.5	13.0
standard	µg/cm2	0.6	0.9	0.7	1.2	0.2	0.4	1.0	0.6
deviation	%	3.8	7.3	3.7	5.3	5.0	8.3	5.2	4.8
2σ*	%	8	15	7	11	10	17	10	10
3σ*	%	12	22	11	16	15	25	16	15

Figure 7 shows z-scores calculated from σ*. Fifteen outliers (lab/sample: 4/4833; 6/4833; 4/4834; 6/4834; 3/DEM4; 4/DEM4; 6/DEM4; 3/DEM9; 4/DEM9; 4/T6690; 3/T6691; 4/T6691; 4/A219; 4/A222 and 6/A222) and four stragglers (lab/sample: 6/DEM9; 6/T6690; 10/T6690; 6/A219) can be identified.

For each sample, seven to nine out of thirteen participants showed deviations from the assigned values within +/- 1 σ* as listed in Table 8 (i.e. within 1 z-score).

78% of all entries were within 10% from the assigned value.

A few participants showed the systematic tendency of overestimating or underestimating the assigned TC concentrations – i.e. participants 4 and 6 for all test samples and participants 1, 2, 7, and 10 for all samples except one. A more accurate determination of the instrument's calibration constant (e.g. implementing CO_2 calibration where possible) would correct this tendency.

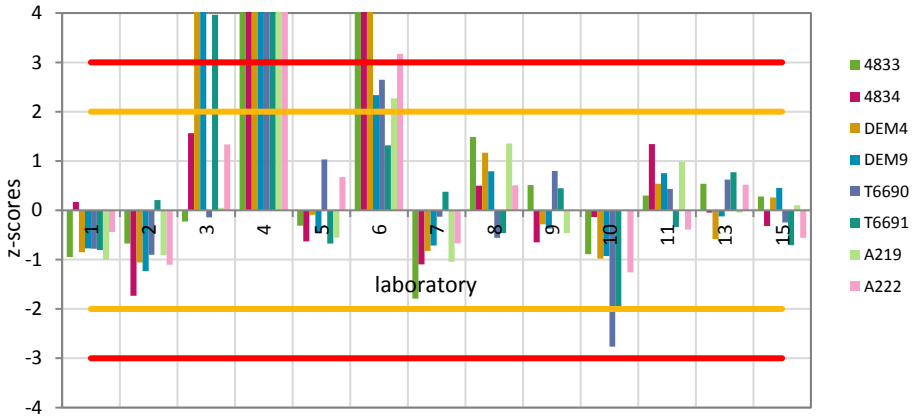


Figure 7. z-scores for TC calculated using σ* *from data obtained in a round of a proficiency testing scheme*.

2.2.3 Results: Laboratory performance for EC/TC

The assigned values, X, and the related standard deviations for proficiency assessment, σ*, are reported in Table 9. Following ISO13528, σ* were calculated *from data obtained in a round of a proficiency testing scheme* and corresponding z-scores are shown in Figure 8.

Table 9: Assigned values and standard deviations for proficiency assessment σ* *from data obtained in a round of a proficiency testing scheme* for EC/TC.

		4833	4834	DEM4	DEM9	T6690	T6691	A219	A222
assigned value	ratio	0.08	0.08	0.14	0.13	0.03	0.06	0.09	0.08
standard	ratio	0.02	0.02	0.02	0.02	0.01	0.02	0.01	0.01
deviation	%	21.32	19.24	16.15	17.70	31.47	31.57	14.75	15.85
2σ*	%	43	38	32	35	63	63	29	32
3σ*	%	64	58	48	53	94	95	44	48

Fifteen outliers (lab/sample: 6/4833; 10/4833; 4/4834; 6/4834; 10/4834; 6/DEM4; 10/DEM4; 6/DEM9; 3/T6690; 6/T6690; 8/T6690; 11/T6690; 3/T6691; 6/A219 and 6/A222) and six stragglers (lab/sample: 4/4833; 4/4834; 3/DEM9; 4/T6691; 8/T6691 and 11/T6691) were identified. For each sample, seven to nine out of thirteen laboratories showed deviations from the assigned values within +/- 1 σ* as listed in Table 9 (i.e. within 1 z-score).

Only 36% of all entries were within 10% of the assigned value and 72% were within the 25% of the assigned value.

A contribution of filter heterogeneities to poor laboratory performances cannot be completely excluded. However, the majority of outliers and stragglers were produced by participants 3, 4, 6, and 10. The recurrence (more than two) of stragglers and/or outliers for single laboratories as observed in this exercise most probably suggest biases in EC/TC determination compared to the other laboratories. Participants 3, 4, 6, and 10 showing large biases (|z-scores|> 2) shall carefully examine their procedures and identify appropriate corrective actions that are likely to prevent the recurrence of such results in the future.

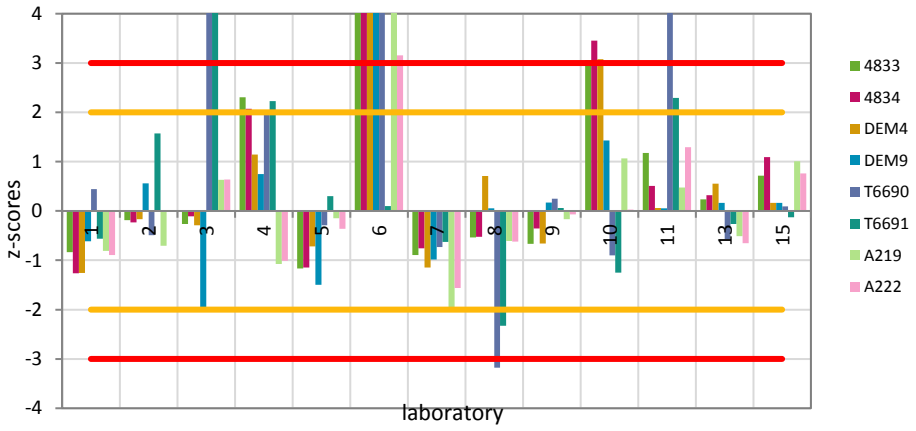


Figure 8. z-scores for EC/TC ratio calculated using σ* *from data obtained in a round of a proficiency testing scheme*.

2.3 PHTHALIC ACID SOLUTION – Percentage differences

Participants were asked to report the OC content of 10 µl of phthalic acid solution. This included the analysis of samples prepared by spiking a pre-cleaned filter punch with 10 µl solution. This is the procedure normally used by laboratories to determine and verify the FID calibration constant.

Figure 9 shows the percentage differences from the assigned value (1.52 ± 0.02 gC l^{-1}, calculated from primary mass and water volume measurements) for each participant. The observed percentage differences range -5% to +10%, while nine out of twelve laboratories reported OC deviating from the assigned value by less than ±5%. Since each phthalic acid solution flask was not checked individually, deviations from the assigned value of the standard solutions cannot be completely excluded.

This exercise did not aim at identifying systematic tendency of a laboratory to underestimate or overestimate the C content of analysed samples but rather to highlight the potential uncertainty (and variability) that can affect TC determination, when the spiking procedure is applied to determine the FID calibration constant.

It is recommended to implement the calibration with CO_2 injections where possible, or to carefully revise the accuracy of all steps involved in the external solution spiking procedure (calibration of the pipette volume, complete deposition of the volume onto a punch filter, drying etc.).

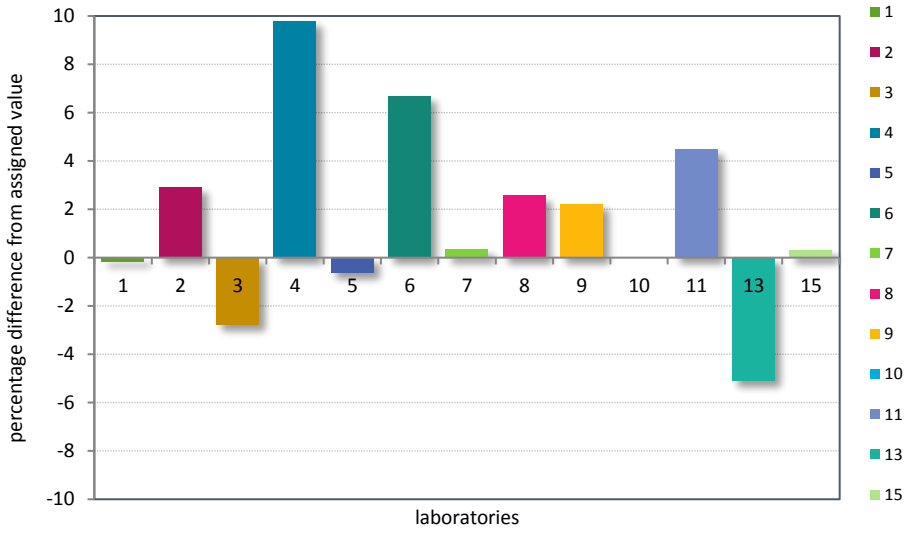


Figure 9. Phthalic acid solution –percentage differences from the assigned value, i.e. the C concentration of the test solution calculated from the mass of phthalic acid and the volume of ultra-pure water used to make the solution. (No data from participant 10).

Conclusions

This inter-laboratory comparison involved thirteen participants applying all thermal-optical analyses with the EUSAAR_2 protocol.
The measurement method ***repeatability and reproducibility for TC*** ranged from 2% to 4% and from 3.5% to 13% (as one relative standard deviation), respectively. For the ***EC/TC ratio, repeatability and reproducibility*** ranged from 3% to 14 and from 15% to 66% (as one relative standard deviation), respectively. Combining the repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol obtained during the previous ILCEs and the present one, we observed that the method precision (both sr and sR) becomes exponentially poorer toward lower TC contents i.e. <13 µgC / cm² and EC/TC ratio. i.e. <0.07.
Stragglers and outliers in terms repeatability and reproducibility of the determination of TC loadings and EC/TC ratios were produced mainly by single participants. Although the contribution of localized sample heterogeneities and /or contaminations to biased data cannot be totally excluded, the random scheme adopted to distribute sub-samples was such that the recurrence of stragglers or outliers for single laboratories most probably indicates an unsatisfactory laboratory precision as compared to the other participants.

Still in absence of a suitable certified reference material for atmospheric OC and EC, assigned values for TC loadings and EC/TC ratios in the filter test samples were calculated as robust averages among all participants.
Laboratory performances were assessed for both TC loadings and EC/TC ratios determinations based on z-scores, applying as assigned values and *standard deviation for proficiency assessment* the ones calculated *from data obtained in a round of a proficiency testing scheme.*
For TC loadings, fifteen outliers and four stragglers were identified; and 78% of all entries were within 10% from the assigned TC concentration value.
A few participants showed the systematic tendency of overestimating or underestimating the assigned TC concentrations – i.e. participants 4 and 6 for all test samples whereas participants 1, 2, 7, and 10 for all samples except one. A more accurate determination of the calibration constant (e.g. implementing where possible CO_2 calibration) would probably correct this tendency.
Regarding EC/TC ratios, the majority of the fourteen outliers and six stragglers were produced by four participants (i.e. 3, 4, 6 and 10). Only 36% of all entries were within 10% of the assigned value and 72% were within the 25% of the assigned EC/TC ratio.
Participants (i.e. 3, 4, 6 and 10) showing large biases (|z-scores|> 2) shall carefully examine their procedures and identify appropriate corrective actions that are likely to prevent the recurrence of such results in the future. A more solid and stable in time instrument set-up in terms of i) laser stability; ii) FID response in He and He/O2 phases; iii) temperature calibration and iv) transit time would correct this behavior and reduce the observed variability in EC/TC ratio determination.

References

EU Directive 2008/50/EC on ambient air and cleaner air for Europe.

FprCEN/TR 16243. Ambient air quality - Guide for the measurement of elemental carbon (EC) and organic carbon (OC) deposited on filters. CEN, Brussels, 2011.

FprEN 16909. Ambient air – Measurement of elemental carbon (EC) and organic carbon (OC) collected on filters. CEN, Brussels, 2016.

Cavalli F., Putaud J., Viana M., Yttri K., Genberg J., Toward a Standardised Thermal-Optical Protocol for Measuring Atmospheric Organic and Elemental Carbon: The EUSAAR Protocol. Atmospheric Measurement Techniques 3 (1); p. 79-89, 2010.

ISO 13528. Statistical methods for use in proficiency testing by inter-laboratory comparisons. ISO, Geneva, 2005.

ISO 13528. Statistical methods for use in proficiency testing by inter-laboratory comparisons. ISO, Geneva, 2015.

ISO 5725-2. Accuracy (trueness and precision) of measurement methods and results -- Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. ISO, Geneva, 1994.

Annex 1. Numerical results reported by participants

Table 1: Total carbon loadings (µg/cm^2)

TC								
Laboratory	4833	4834	DEM4	DEM9	T6690	T6691	A219	A222
1	14.153	13.419	17.767	22.091	4.449	4.217	17.280	12.541
	14.249	11.444	17.542	21.890	4.673	4.073	17.520	12.716
	14.205	12.025	17.264	21.954	4.623		17.794	13.052
2	14.321	10.703	17.275	21.730	4.500	5.216	17.318	12.414
	14.783	10.093	17.468	20.712	4.394	4.437	18.128	12.707
	13.965	11.023	17.405	21.802	4.764	3.909	17.418	11.937
3	14.506	12.988	20.948	44.651	4.865	5.760	19.100	13.607
	14.612	14.126	22.718	42.921	4.648	5.828	18.657	13.756
	14.712	13.481	23.532	42.577	4.689	6.150	17.875	14.318
4	19.574	21.913	32.608	29.265	6.899	13.003	31.612	22.240
	25.863	–	30.970	–	8.183	9.647	–	–
5	14.469	11.599	17.713	22.385	5.609	4.109	18.068	12.831
	14.931	11.419	17.675	22.635	4.664	4.357	17.901	13.088
	14.288	11.731	18.714	22.137	4.766	4.116	17.934	14.514
6	16.840	17.060	21.410	25.920	5.530	5.070	20.790	15.450
	31.900	18.640	20.870	25.420	5.480	4.610	20.810	14.880
	20.400	25.650	21.220	25.990	5.180	5.120	20.400	14.830
7	13.351	11.111	17.796	21.847	4.860	4.325	17.120	12.933
	14.080	11.380	18.051	22.357	4.678	4.173	17.759	12.703
	13.746	11.017	16.772	21.960	4.671	5.251	17.594	12.247
8	15.786	12.456	18.555	23.319	4.534	4.378	19.726	13.121
	15.582	12.746	19.116	23.834	4.663	4.236	19.972	13.364
	15.375	12.569	18.985	24.511	4.706	4.210	19.672	13.620
9	15.590	11.452	17.768	22.354	5.204	4.648	17.960	13.167
	14.466	11.358	17.852	22.582	4.804	4.594	18.025	12.953
	15.040	11.894	18.086	22.474	4.863	4.589	18.167	13.050
10	14.980	11.590	16.820	21.280	4.060	3.400	17.380	12.020
	14.130	12.280	17.080	22.140	4.300	3.970	20.240	12.100
	13.600	12.200	18.410	21.960	3.970	3.780	17.890	12.650
11	14.912	12.191	18.488	25.198	4.942	4.305	19.460	12.610
	14.418	16.046	18.544	22.915	4.941	4.346	19.278	12.697
	15.396	11.782	18.345	23.428	4.729	4.307	19.547	13.106
13	14.534	12.222	17.430	22.822	5.062	4.836	18.391	13.288
	16.142	12.144	17.381	22.247	4.648	4.745	18.734	13.372
	14.459	11.952	18.295	23.256	5.034	4.611	18.213	13.472
15	14.488	11.784	18.464	23.374	4.463	4.145	18.414	12.908
	14.544	11.653	18.263	23.486	4.596	4.135	18.504	12.717
	15.661	12.153	18.086	23.584	5.069	4.264	18.860	12.470

Table 2: Elemental carbon / total carbon (ratios)

EC/TC								
Laboratory	4833	4834	DEM4	DEM9	T6690	T6691	A219	A222
1	0.065	0.056	0.111	0.117	0.030	0.048	0.085	0.067
	0.063	0.065	0.111	0.115	0.036	0.046	0.085	0.069
	0.061	0.066	0.109	0.113	0.035		0.080	0.067
2	0.062	0.082	0.131	0.146	-0.008	0.104	0.073	0.077
	0.077	0.076	0.138	0.140	0.026	0.102	0.101	0.086
	0.082	0.079	0.135	0.139	0.057	0.050	0.081	0.075
3	0.074	0.079	0.137	0.079	0.074	0.149	0.102	0.086
	0.070	0.078	0.136	0.083	0.071	0.128	0.103	0.089
	0.073	0.085	0.123	0.085	0.068	0.132	0.106	0.087
4	0.135	0.115	0.159	0.146	0.043	0.175	0.080	0.067
	0.093	–	0.169	–	0.051	0.020	–	–
5	0.055	0.062	0.119	0.092	0.026	0.062	0.096	0.075
	0.059	0.066	0.127	0.095	0.027	0.065	0.088	0.078
	0.059	0.065	0.121	0.096	0.026	0.061	0.094	0.071
6	0.156	0.175	0.349	0.310	0.052	0.000	0.173	0.117
	0.139	0.149	0.343	0.303	0.097	0.087	0.186	0.118
	0.169	0.180	0.342	0.294	0.066	0.090	0.190	0.121
7	0.062	0.065	0.115	0.109	0.019	0.042	0.065	0.058
	0.064	0.075	0.112	0.109	0.022	0.049	0.070	0.063
	0.060	0.072	0.111	0.101	0.027	0.046	0.067	0.059
8	0.068	0.074	0.143	0.129	0.000	0.007	0.093	0.079
	0.071	0.074	0.157	0.130	0.000	0.002	0.082	0.070
	0.065	0.074	0.162	0.132	0.000	0.037	0.085	0.066
9	0.063	0.075	0.123	0.134	0.031	0.054	0.091	0.079
	0.070	0.077	0.130	0.131	0.032	0.059	0.090	0.083
	0.063	0.078	0.118	0.133	0.033	0.062	0.097	0.074
10	0.123	0.134	0.202	0.151	0.022	0.026	0.091	0.078
	0.128	0.144	0.205	0.170	0.021	0.045	0.123	0.079
	0.129	0.134	0.215	0.163	0.020	0.032	0.116	0.081
11	0.096	0.094	0.140	0.126	0.065	0.094	0.099	0.093
	0.099	0.077	0.134	0.133	0.076	0.101	0.102	0.096
	0.093	0.101	0.146	0.131	0.069	0.099	0.104	0.098
13	0.079	0.090	0.150	0.128	0.020	0.050	0.090	0.070
	0.076	0.090	0.158	0.134	0.026	0.053	0.086	0.072
	0.087	0.081	0.144	0.135	0.026	0.054	0.088	0.071
15	0.088	0.098	0.142	0.130	0.024	0.044	0.106	0.085
	0.090	0.100	0.146	0.131	0.034	0.057	0.105	0.090
	0.087	0.102	0.138	0.137	0.032	0.063	0.116	0.091