Results of the inter-laboratory comparison exercise for TC and EC measurements (ref.: OCEC-2017-1)

Summary

The European Centre for Aerosol Calibration (ECAC) under ACTRIS-2 completed in March 2017 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 Italy and Norway and an urban background site in Spain. A solution of phthalic acid prepared at JRC-ERLAP (the inter-laboratory comparison exercise coordinator) was also distributed.

Seventeen laboratories participated in this exercise running their usual thermal-optical protocol (16 applied EUSAAR_2 and one "Quartz") with their usual analytical instrument. Among those, twelve are responsible for the aerosol chemical speciation at the EMEP or ACTRIS stations located in their countries (i.e. Spain, Germany, Norway, Greece, Czech Republic, Poland, France, The Netherlands, Cyprus and Italy).

Measurement method performance: for TC determination, repeatability and reproducibility relative standard deviations ranged from 3% to 6% and from 5% to 8%, respectively. For the determination of the EC/TC ratio, repeatability and reproducibility relative standard deviations ranged from 3% to 7% and from 13% to 16%. Based on last four inter-laboratory comparisons, 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. <10 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, fourteen outliers –mainly from two laboratories– and two stragglers were identified; 87% of all entries were within 10% from the assigned TC concentration value.
Regarding EC/TC ratios, two outliers and eight stragglers were identified. 59% of all entries is within 10% of the assigned value and 93% 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 and the thermal protocol EUSAAR-2 with a transmittance optical correction for pyrolysis has been recently selected as the European standard thermal protocol (EN16909:2017).

The *European center for aerosol calibration* within the European project ACTRIS-2 has organized in January-March 2017 an inter-laboratory comparison exercise (ILCE) (ref. OCEC-2017-1) among seventeen applicants including, among others, laboratories in charge of OC and EC measurements at EMEP/ACTRIS station in Spain, Germany, Norway, Greece, Czech Republic, Poland, France, The Netherlands, Cyprus 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 three sites across Europe (Table 1). Upon receipt at ERLAP, 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
Barcelona	Spain	TER1 _	Urban background	Dec.2016	PM2.5	Pallflex
Ispra	Italy	IPR_	rural	Dec.2016/Jan. 2017	PM2.5	Pallflex
Birkenes	Norway	A2 _ _	rural	Apr./May 2000	PM10	Whatman QMA

Rectangular filter punches of ca. 3.6 cm x 1.8 cm (or of 1.6 cm dia.) were randomly distributed to participants to allow them to triplicate measurements.

The homogeneity of these test samples was investigated by ERLAP on one of the test samples for 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 ERLAP laboratory (< 3 and 5% for TC and EC, respectively). The homogeneity was better than 4 and 7% for TC and EC/TC, respectively (Table 2). If sampling at each location occurred

under repeatable conditions, it can be assumed that the remaining 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 (%)
IPR_ _ Ispra (I)	3.8	2.7
TER1 _ Barcelona(E)	2.4	1.9
A2_ _ Birkenes (N)	3.8	7.0

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 seventeen 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 17) on 26th Jan. 2017 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 March 2017. 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 (EN16909:2017).
In this exercise all laboratories but one (Lab 14) applied the EUSAAR_2 protocol (Table 4) with transmittance-based correction.

Table 3: List of participants in the inter-laboratory comparison 2017, and contact persons

Code	Participant	Acronym	Contact
1	Clemen, Sebastian	s-swu	sebastian.clemen@senstadtum.berlin.de
2	Virginia Andreoli	cnr-iia	v.andreoli@iia.cnr.it
3	Daniele Contini	cnr-isac	a.dinoi@le.isac.cnr.it
4	PONT Veronique	obs-mip	veronique.pont@aero.obs-mip.fr
5	Yubero, Eduardo	umh	eyubero@umh.es
6	Bieber, Elke	uba	elke.bieber@uba.de
7	Yttri, Karl Espen	nilu	Key@nilu.no
8	Eleftheriadis, Konstantinos	demokritos	ldiapouli@ipta.demokritos.gr
9	Milan Vana	chmi	Adela.holubova@chmi.cz
10	Holoubek, Ivan	cgri	Holubovasmejkalova.a@czechglobe.cz
11	Mathews Barbara	ipis	barbara.mathews@ipis.zabrze.pl
12	GROS Valérie	lsce	nicolas.bonnaire@lsce.ipsl.fr
13	Panteliadis, Pavlos	ggd	ppanteliadis@ggd.amsterdam.nl
14	Rosalía Fernández Patier	isciii	aaguiar@isciii.es
15	JAFFREZO Jean-Luc	univ-grenoble	jaffrezo@univ-grenoble-alpes.fr
16	SCIARE, Jean	cyi	k.oikonomou@cyi.ac.cy / i.hafez@cyi.ac.cy
17	Cavalli, Fabrizia	jrc	fabrizia.cavalli@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	s-swu	Sunset-lab analyser	EUSAAR_2	1.50
2	cnr-iia	Sunset-lab analyser	EUSAAR_2	1.00
3	cnr-isac	Sunset-lab analyser	EUSAAR_2	1.00
4	obs-mip	DRI	EUSAAR_2	0.55
5	umh	Sunset-lab analyser	EUSAAR_2	1.50
6	uba	Sunset-lab analyser	EUSAAR_2	1.50
7	nilu	Sunset-lab analyser	EUSAAR_2	1.50
8	demokritos	Sunset-field analyser	EUSAAR_2	2.14
9	chmi	Sunset-lab analyser	EUSAAR_2	1.50
10	czechglobe	Sunset-field analyser	EUSAAR_2	1.27
11	ipis	Sunset-lab analyser	EUSAAR_2	1.5
12	lsce	Sunset-lab analyser	EUSAAR_2	1.50
13	ggd	Sunset-field analyser	EUSAAR_2	2.05
14	isciii	Sunset-lab analyser	Quartz	1.50
15	univ-grenoble	Sunset-lab analyser	EUSAAR_2	1.5
16	cyi	Sunset-lab analyser	EUSAAR_2	1.50
17	jrc	Sunset-lab analyser	EUSAAR_2	1.00

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

	EUSAAR_2		QUARTZ	
Carrier gas	Time (s)	Temp. (°C)	Time (s)	Temp. (°C)
Helium	120	200	80	310
Helium	150	300	80	475
Helium	180	450	80	615
Helium	180	650	110	870
Helium			45	550
Oxygen in Helium (2%)	120	500	45	625
Oxygen in Helium	120	550	45	700
Oxygen in Helium	70	700	45	775
Oxygen in Helium	80/110	850	45	850
Oxygen in Helium			110	870

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 8 to 21 µg cm^{-2}, corresponding to atmospheric concentrations ranging from 1.5 to 5 µg m^{-3} collected for 24h at a face velocity of 54 cm s^{-1}. EC/TC ranged on average from 0.07 to 0.29. All submitted results for TC (in µg cm^{-2}) EC, OC (calculated as OC = TC-EC) and EC/TC ratio are presented in tables in Annex 1.

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 values from all participants (see Par 2.2).

Note that mean/robust averages and standard deviations were calculated from values from all participants, including also those from laboratory 14 applying the QUARTZ protocol.

TC values from various thermal protocols do not significantly differ but EC/TC ratios can with EC/TC ratio from the QUARTZ protocol being typically lower than those from the EUSAAR-2 protocol. In the present exercise EC/TC ratios from laboratory 14 were indeed generally lower

(on average, 1 σ) than the overall averages but similar deviation were observed for other laboratories applying the EUSAAR-2 protocol).

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 1.57 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).

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 provided three replicates (except lab 8 which provided two replicates) for every sample. Thus *Mandel's k* was calculated for an average case of three replicates and compared to the critical values 2.06 (outlier) and 1.70 (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 seventeen participants, the critical values for Mandel's *h* are 2.35 (outlier) and 1.87 (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).

Eleven outliers (lab/sample: 10/IPR1; 10/IPR2; 4/IPR3; 8/IPR3; 10/IPR3; 10/A210; 4/A223; 10/A223; 4/A224; 10/TER10; 4/TER11) and one straggler (lab/sample: 8/IPR1) were identified (Fig. 1). Cochran's test confirmed as outliers 10/IPR1; 10/IPR2; 4/IPR3; 8/IPR3; 10/IPR3; 4/A223; 10/A223; 4/A224; 10/TER10 and 10/A210 and 4/TER11 as stragglers.

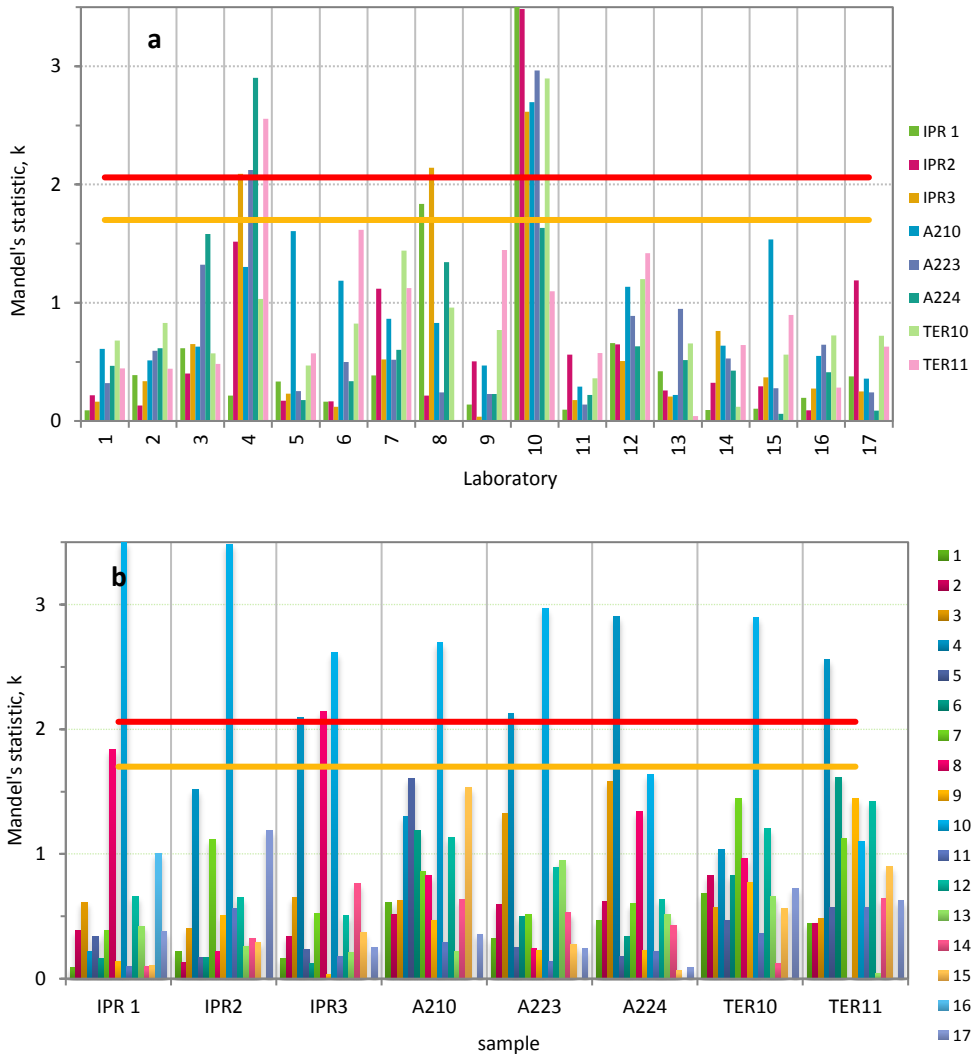


Figure 1. *Mandel's k* statistic values for within-laboratory consistency on TC data, grouped by laboratory (panel a) and by sample (panel b). For seventeen laboratories and three replicates, *k* values should be < 2.06 at the 1% significance level (red line) and < 1.7 at the 5% significance level (orange 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, eight outliers (lab/sample: 10/IPR1; 10/IPR2; 10/IPR3; 10/A210; 10/A223; 10/A224; 10/TER10 and 10/TER11) and two stragglers (lab/sample: 4/A223 and 8/TER11) were identified (Fig. 2). The Grubbs' tests confirmed as outliers 10/IPR1; 10/IPR2; 10/IPR3; 10/A210; 10/A223; 10/A224; 10/TER10; 8/TER11and 10/TER11 and identified no straggler.

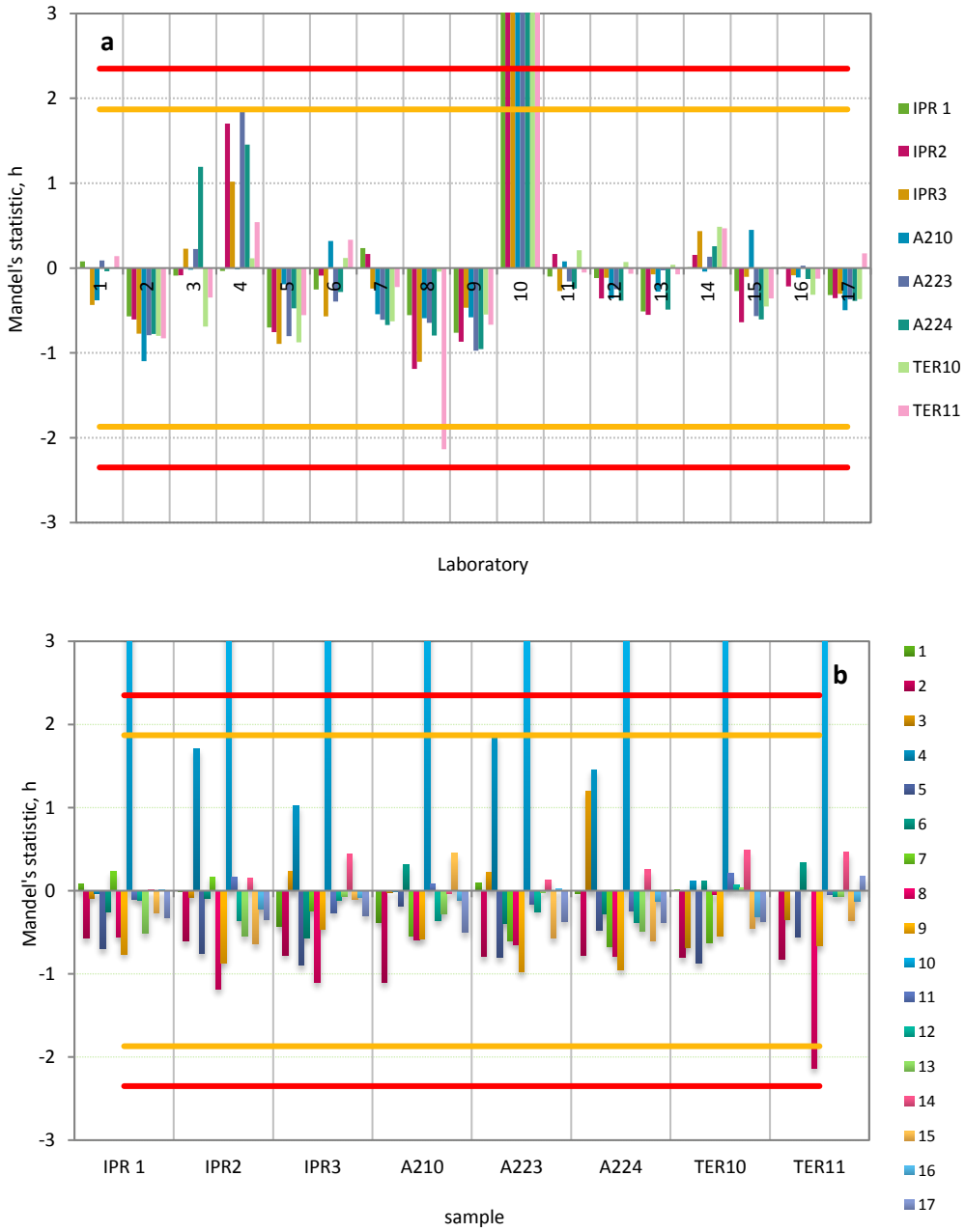


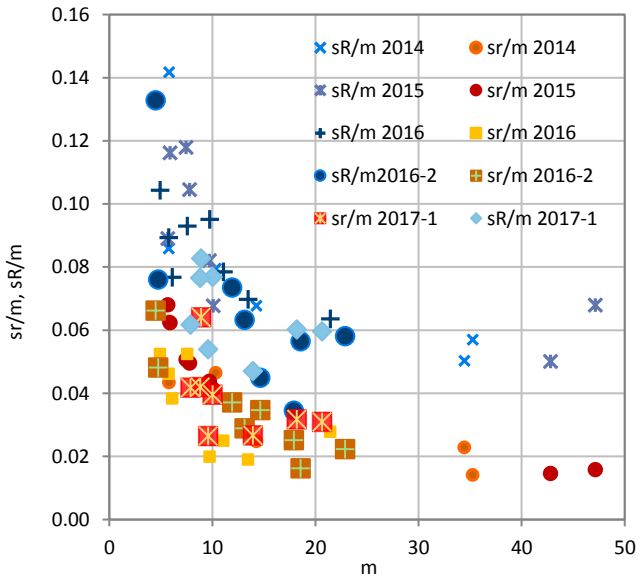
Figure 2. *Mandel's h* statistic values for between-laboratory consistency on TC data, grouped by laboratory (panel a) and by sample (panel b). For seventeen laboratories, *h* values should be < 2.35 at the 1% significance level (red line) and < 1.87 at 5% significance level (orange line).

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 4 and 10) 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 four 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. < 10 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	%
IPR 1	13.9	0.4	2.7	0.7	4.7
IPR2	10.0	0.4	4.0	0.8	7.7
IPR3	8.9	0.6	6.4	0.7	8.3
A210	7.9	0.3	4.2	0.5	6.2
A223	9.6	0.3	2.7	0.5	5.4
A224	8.8	0.4	4.2	0.7	7.7
TER10	18.2	0.6	3.2	1.1	6.0
TER11	20.7	0.6	3.1	1.2	6.0



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, twelve outliers (lab/sample: 8/IPR1; 4/IPR2; 12/IPR2,; 17/IPR2; 14/IPR3; 4/A210; 4/A223; 10/A224; 14/A224; 2/TER10; 3/TER10; 4/TER11) and four stragglers (lab/sample: 8/IPR3; 12/IPR3; 12/TER10;and 7/TER11) were identified (Fig. 4). Cochran's test identified the entries 8/IPR1; 14/IPR3; 4/A210; 4/A223; 10/A224; 14/A224; 4/TER11; and 7/TER11 as outliers and the entry 2/TER10 as straggler.

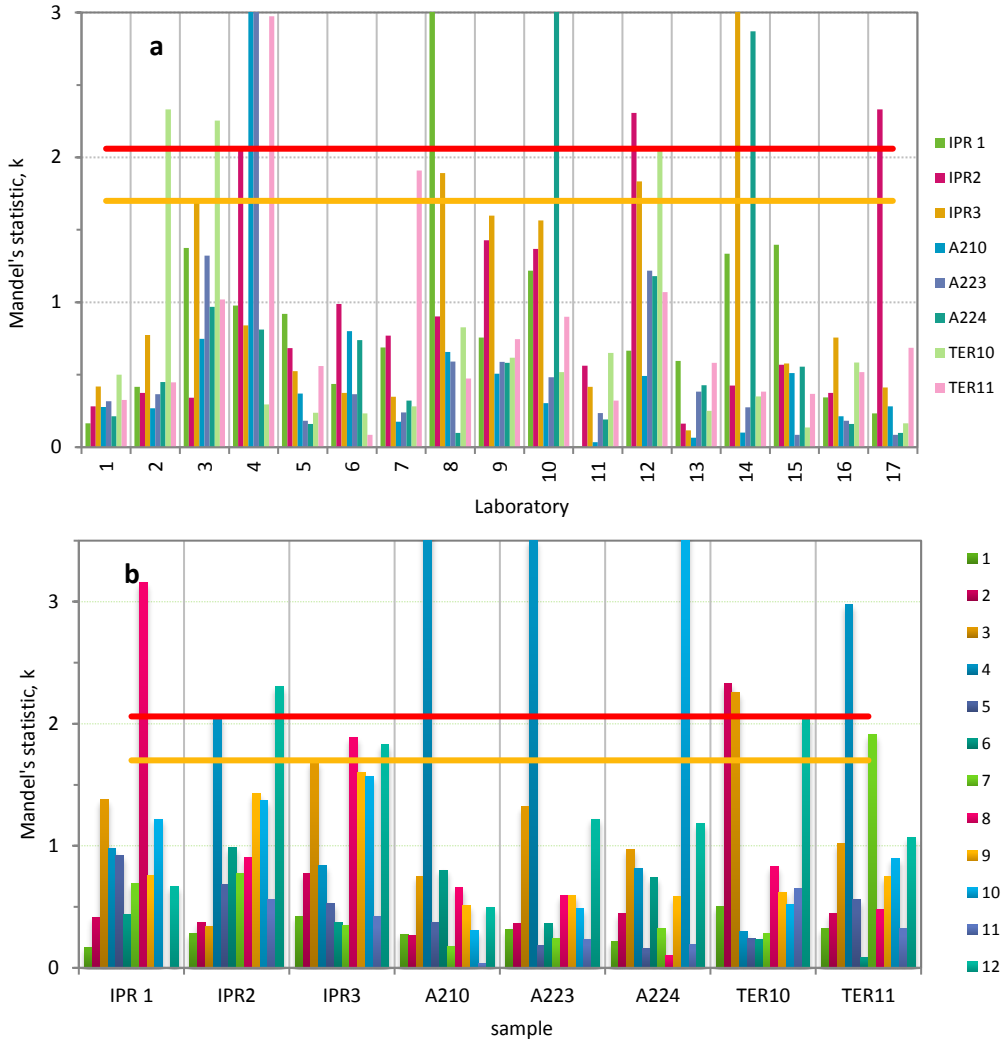


Figure 4. *Mandel's k* statistic values for within laboratory consistency on EC/TC ratio obtained from the entire database, grouped by laboratory (panel a) and by sample (panel b). For seventeen laboratories and three replicates k values should be < 2.06 at 1% significance level (red line) and < 1.7 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).

Two outliers (lab/sample: 4/A210 and 2/TER10) and eight stragglers (lab/sample: 7/IPR1; 7/IPR2; 4/IPR3; 10/A210; 4/A223; 2/A224; 10/A224 and 8/TER11) were identified. Grubbs' test identifies the entry 2/TER10 as outliers and no stragglers.

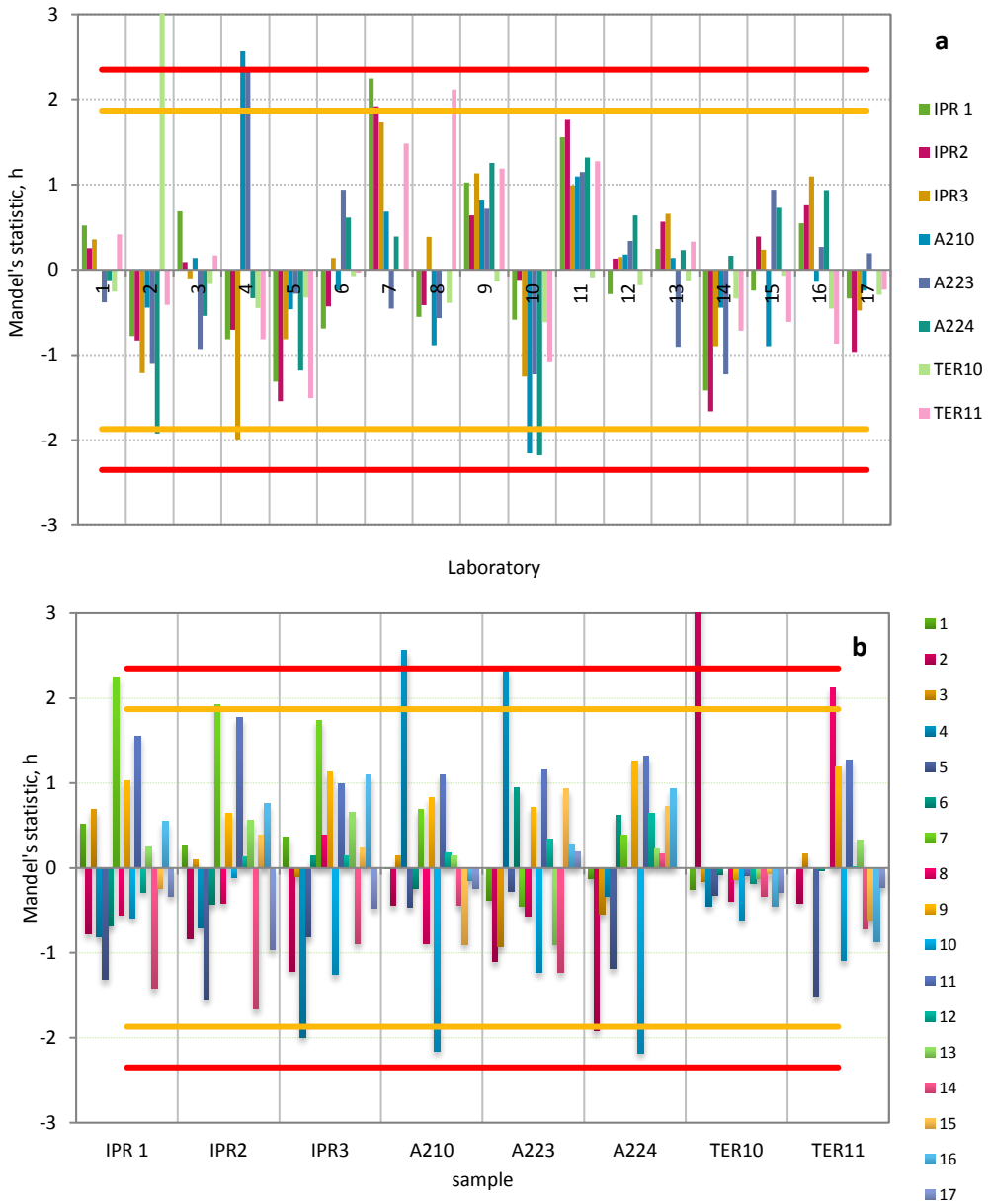


Figure 5. *Mandel's h* statistic values for between laboratory consistency on EC/TC ratio obtained from the entire database, grouped by laboratory (panel a) and by sample (panel b). For seventeen laboratories h values should be < 2.35 at 1% significance level (red line) and < 1.87 at 5% significance 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 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 four 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	
			%		%
IPR 1	0.18	0.01	2.8	0.03	14.7
IPR2	0.19	0.01	5.7	0.02	12.7
IPR3	0.15	0.01	5.8	0.02	12.2
A210	0.11	0.01	5.9	0.01	12.9
A223	0.07	0.00	6.5	0.01	16.4
A224	0.07	0.00	6.1	0.01	14.5
TER10	0.28	0.01	4.0	0.03	12.2
TER11	0.29	0.01	2.9	0.03	11.5

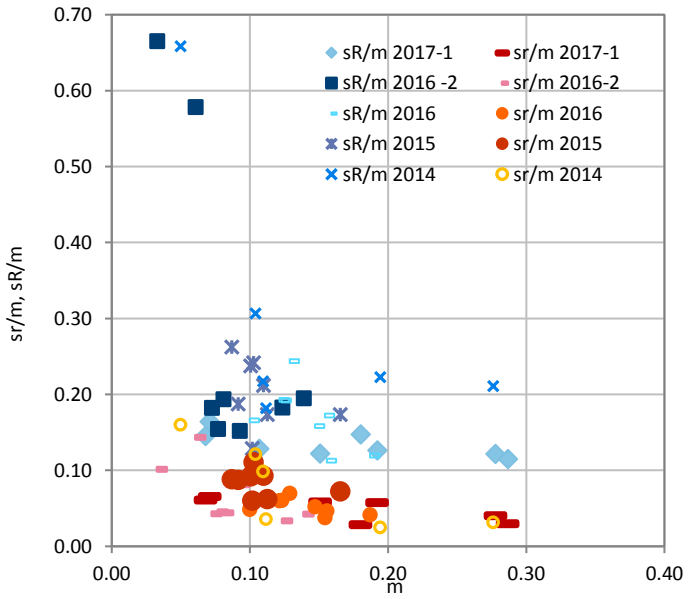


Figure 6. Repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol for EC/TC measurement obtained during the previous inter-laboratory comparisons and the present one.

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.

In Annex 1 tables 5, 6 and 7 are reported statistics (percentage bias and variability)

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.

		IPR 1	IPR2	IPR3	A210	A223	A224	TER10	TER11
assigned value	µg/cm2	14.0	10.0	8.9	7.9	9.7	8.8	18.3	20.7
standard	µg/cm2	0.6	0.6	0.5	0.4	0.7	0.5	1.2	1.3
deviation	%	4.5	6.0	5.1	5.5	6.7	5.9	6.4	6.5
2σ*	%	9	12	10	11	13	12	13	13
3σ*	%	14	18	15	16	20	18	19	19

Figure 7 shows z-scores calculated from σ*. Fourteen outliers, mainly from two laboratories (lab/sample: 10/IPR1; 4/IPR2; 10/IPR2; 4/IPR3; 10/IPR3; 10A210; 4/A223; 10/A223; 3/A224; 4/A224; 10/A224; 10/TER10; 8/TER11 and 10/TER11) and two stragglers (lab/sample: 8/IPR2; and 2/A210) were identified.

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

87% of all entries were within 10% from the assigned value.

A few participants showed the systematic tendency of overestimating –i.e. labs 4, 10, 11 and 14 - or underestimating –i.e. labs 2, and 9 - the assigned TC concentrations. A more accurate determination of the instrument's calibration constant (e.g. implementing CO_2 calibration where possible) would correct this tendency.

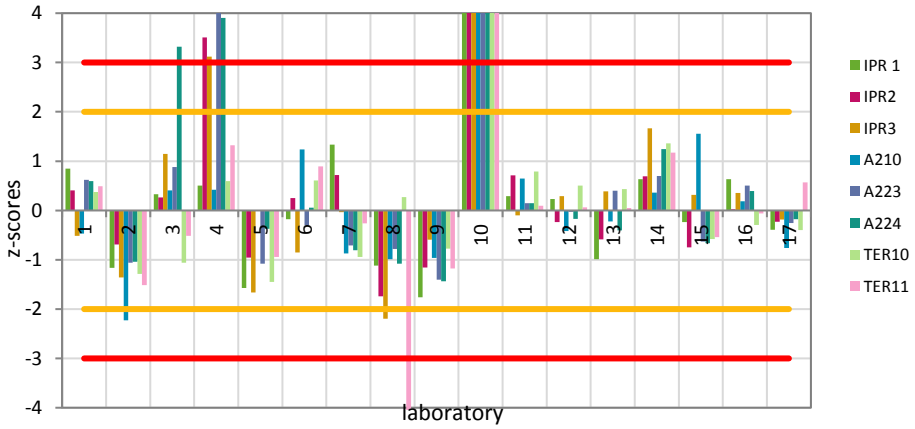


Figure 7. z-scores for TC calculated using σ* *from data obtained in a round of a proficiency testing scheme*.

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.

		IPR 1	IPR2	IPR3	A210	A223	A224	TER10	TER11
assigned value	ratio	0.18	0.19	0.15	0.11	0.07	0.07	0.28	0.29
standard	ratio	0.02	0.02	0.02	0.01	0.01	0.01	0.03	0.04
deviation	%	12.8	10.5	12.0	8.9	18.0	12.5	11.9	12.8
2σ*	%	26	21	24	18	36	25	24	26
3σ*	%	38	32	36	27	54	38	36	38

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.

Three outliers (lab/sample: 4/A210; 10/A210 and 2/TER10) and eight stragglers (lab/sample: 7/IPR1; 7/IPR2; 11/A210; 4/A223; 2/A224; 10/A224; 10/TER10; and 8/TER11) were identified. For each sample, ten to twelve out of seventeen laboratories showed deviations from the assigned values within +/- 1 σ* as listed in Table 9 (i.e. within 1 z-score).
59% of all entries were within 10% of the assigned value and 93% were within the 25% of the assigned value.
A contribution of filter heterogeneities to poor laboratory performances cannot be completely excluded. 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 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.

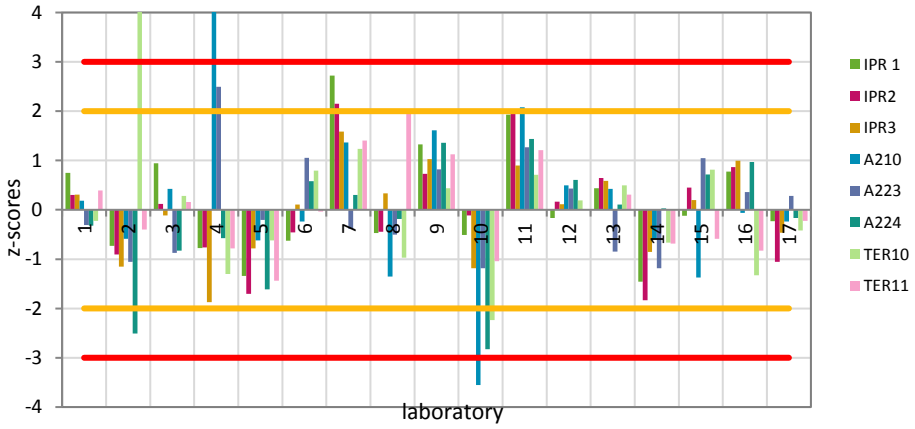


Figure 8. z-scores for EC/TC ratio calculated using σ* *from data obtained in a round of a proficiency testing scheme*.

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 (1.57 ± 0.02 gC l^{-1}, calculated from primary mass and water volume measurements) for each participant. Eleven laboratories out of seventeen 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 carbon 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.).

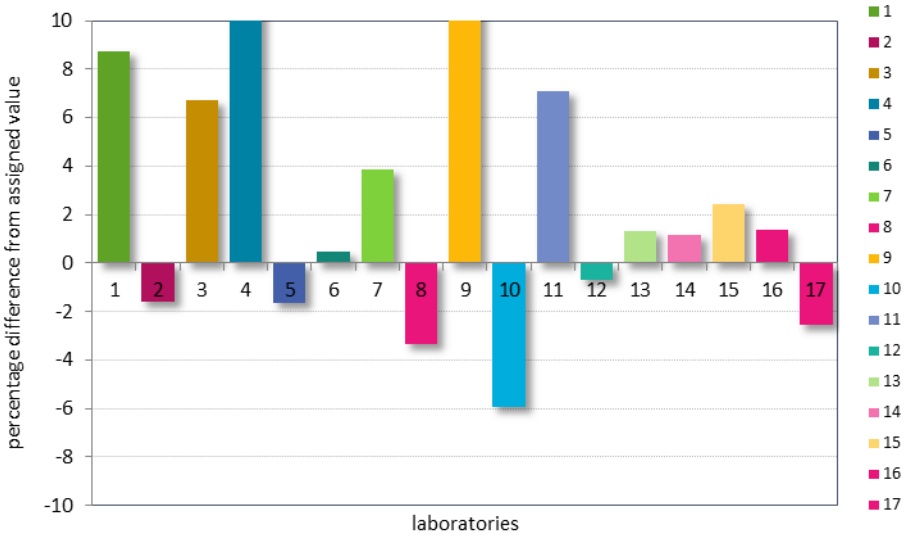


Figure 9. Phthalic acid solution –percentage differences from the assigned value, i.e. the C concentration of the test solution calculated from the mass of phthalic acid and the volume of ultra-pure water used to make the solution.

Conclusions

This inter-laboratory comparison involved seventeen participants all applying thermal-optical analyses with the EUSAAR_2 protocol but one applying the Quartz protocol.
The measurement method ***repeatability and reproducibility for TC*** ranged from **3% to 6%** and from **5% to 8%** (as one relative standard deviation), respectively. For the ***EC/TC ratio, repeatability and reproducibility*** ranged from **3% to 7%** and from **13% to 16%** (as one relative standard deviation), respectively. Combining the repeatability and reproducibility relative standard deviation for the EUSAAR-2 protocol obtained during the previous four ILCEs and the present one, we observed that the method precision (both sr and sR) becomes exponentially poorer toward lower TC contents i.e. <10 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 mainly produced 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, eleven outliers and one stragglers were identified; and 87% of all entries were within 10% from the assigned TC concentration value.
A few participants showed the systematic tendency of overestimating –i.e. labs 4, 10, 11 and 14 - or underestimating –i.e. labs 2, and 9 - the assigned TC concentrations. 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, three outliers and eight stragglers were identified. 59% of all entries were within 10% of the assigned value and 93% were within the 25% of the assigned value.
Participants 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)

Laboratory	IPR 1	IPR 2	IPR 3	A 210	A 223	A 224	TER 10	TER 11
1	14.619	10.382	8.722	7.956	10.013	9.438	18.206	21.648
	14.582	10.223	8.516	7.741	10.278	8.992	18.981	21.160
	14.479	10.118	8.738	7.509	10.117	8.930	19.116	21.233
2	13.051	9.528	8.479	7.135	8.824	8.638	16.942	18.375
	13.635	9.674	7.990	7.019	9.311	7.905	17.359	18.690
	13.170	9.545	8.359	6.769	8.990	8.271	16.179	18.896
3	14.378	10.437	9.831	8.380	9.809	9.517	17.550	20.008
	13.685	9.982	8.867	7.958	10.900	10.729	16.976	19.705
	14.627	10.047	9.538	8.012	10.210	11.368	16.749	20.278
4	14.180	12.350	9.580	7.850	12.130	9.440	18.300	23.960
	14.320	12.900	9.210	8.670	13.330	10.310	19.000	22.490
	14.520	11.090	12.130	7.840	11.600	12.770	19.790	20.930
5	12.800	9.320	8.310	8.580	8.910	8.720	16.250	19.690
	12.960	9.530	7.960	7.800	9.110	8.630	16.860	19.040
	13.320	9.410	8.150	7.430	9.070	8.510	16.810	19.530
6	13.945	10.184	8.605	7.982	9.362	8.613	19.652	20.883
	14.021	10.034	8.492	8.801	9.478	8.918	19.006	21.964
	13.767	10.229	8.424	8.642	9.766	8.991	18.464	22.794
7	14.600	11.200	8.450	7.330	9.300	8.010	16.900	20.900
	14.800	9.890	9.230	7.430	9.040	8.720	18.400	20.500
	15.200	10.200	8.950	7.920	9.470	8.450	16.400	19.600
8	12.285	8.854	9.050	7.726	9.294	7.686	19.138	15.028
	14.349	9.040	6.746	7.297	9.152	8.816	18.158	15.028
9	13.025	8.953	8.594	7.471	8.925	8.178	17.642	18.813
	12.897	9.547	8.627	7.715	8.748	7.916	17.851	20.069
	12.807	9.395	8.649	7.383	8.776	8.107	16.803	18.430
10	22.118	15.483	10.784	11.191	13.576	13.761	25.900	29.200
	24.396	14.222	13.523	13.106	12.675	11.875	29.913	29.009
	18.848	11.322	14.652	11.740	15.123	12.411	26.881	30.220
11	14.290	10.111	8.962	8.128	9.763	9.033	19.078	20.423
	14.138	10.379	8.698	8.196	9.865	8.777	19.557	20.995
	14.192	10.794	8.879	8.336	9.862	8.852	19.137	21.027
12	13.951	9.671	9.116	7.857	9.726	8.771	18.434	20.119
	14.768	10.010	8.907	7.840	9.731	8.600	18.922	20.734
	13.792	9.881	9.047	7.577	9.689	8.798	19.405	21.448
13	13.398	9.644	9.068	7.843	9.995	8.601	18.841	20.749
	13.615	9.544	9.048	7.834	9.636	8.500	17.902	20.736
	14.056	9.334	8.787	7.699	9.204	9.074	18.481	20.784
14	14.362	10.612	9.519	8.029	10.444	9.282	19.898	21.892
	14.506	10.218	10.279	8.353	10.079	9.345	19.854	22.235
	14.400	10.418	9.143	7.902	10.047	9.749	20.021	22.652
15	13.791	9.749	9.112	8.052	9.451	8.431	18.077	19.351
	13.869	9.407	9.269	8.612	9.229	8.504	17.627	20.197
	13.956	9.485	8.726	9.175	9.284	8.471	17.269	20.331
16	14.351	10.022	9.059	8.252	9.858	8.879	17.960	20.552
	14.601	10.059	8.843	7.896	10.367	9.299	17.483	20.794
	14.318	9.951	9.261	7.913	9.958	8.871	18.525	20.471
17	13.438	9.551	8.530	7.649	9.571	8.705	18.423	21.291
	13.863	10.580	8.799	7.464	9.429	8.632	17.688	21.819
	14.027	9.452	9.098	7.716	9.698	8.824	17.485	21.219

Table 2: Elemental carbon / total carbon (ratios)

Laboratory	IPR 1	IPR 2	IPR 3	A 210	A 223	A 224	TER 10	TER 11
1	0.192	0.195	0.157	0.113	0.071	0.067	0.271	0.308
	0.194	0.200	0.159	0.111	0.066	0.065	0.273	0.301
	0.193	0.199	0.152	0.105	0.067	0.064	0.283	0.300
2	0.160	0.170	0.122	0.106	0.061	0.043	1.047	0.274
	0.157	0.177	0.134	0.103	0.055	0.048	1.068	0.268
	0.162	0.174	0.133	0.098	0.059	0.049	1.009	0.280
3	0.200	0.192	0.162	0.106	0.062	0.065	0.267	0.310
	0.188	0.198	0.150	0.125	0.071	0.053	0.287	0.288
	0.204	0.193	0.133	0.105	0.049	0.065	0.324	0.285
4	0.154	0.194	0.114	0.115	0.121	0.059	0.237	0.222
	0.166	0.156	0.125	0.218	0.071	0.070	0.237	0.256
	0.156	0.179	0.111	0.124	0.122	0.060	0.244	0.301
5	0.140	0.155	0.132	0.107	0.071	0.053	0.265	0.240
	0.147	0.153	0.136	0.103	0.068	0.055	0.259	0.227
	0.151	0.165	0.141	0.096	0.069	0.055	0.263	0.240
6	0.159	0.172	0.150	0.117	0.083	0.071	0.313	0.288
	0.163	0.189	0.151	0.107	0.085	0.079	0.310	0.288
	0.164	0.187	0.156	0.093	0.089	0.069	0.307	0.286
7	0.236	0.227	0.179	0.123	0.069	0.072	0.326	0.358
	0.234	0.240	0.176	0.122	0.065	0.072	0.328	0.311
	0.242	0.239	0.182	0.118	0.067	0.068	0.321	0.352
8	0.179	0.177	0.145	0.102	0.062	0.067	0.258	0.367
	0.152	0.189	0.168	0.088	0.069	0.066	0.243	0.358
9	0.202	0.197	0.157	0.130	0.086	0.075	0.295	0.330
	0.205	0.222	0.166	0.115	0.085	0.083	0.292	0.340
	0.211	0.201	0.184	0.125	0.077	0.081	0.307	0.320
10	0.159	0.175	0.140	0.073	0.061	0.015	0.215	0.249
	0.162	0.195	0.133	0.070	0.056	0.054	0.206	0.263
	0.173	0.199	0.114	0.079	0.053	0.063	0.202	0.239
11	0.220	0.226	0.166	0.127	0.091	0.079	0.298	0.329
	0.220	0.235	0.170	0.128	0.088	0.082	0.313	0.338
	0.220	0.236	0.163	0.128	0.087	0.080	0.311	0.333
12	0.175	0.195	0.147	0.108	0.077	0.070	0.298	0.293
	0.168	0.191	0.153	0.113	0.074	0.076	0.288	0.286
	0.174	0.200	0.158	0.117	0.082	0.073	0.283	0.288
13	0.186	0.205	0.161	0.112	0.061	0.069	0.300	0.300
	0.181	0.202	0.159	0.113	0.066	0.063	0.295	0.287
	0.188	0.203	0.160	0.111	0.067	0.067	0.301	0.286
14	0.134	0.151	0.120	0.104	0.058	0.055	0.261	0.262
	0.148	0.159	0.168	0.101	0.058	0.058	0.265	0.259
	0.148	0.155	0.117	0.102	0.054	0.092	0.256	0.269
15	0.180	0.195	0.152	0.103	0.086	0.070	0.312	0.266
	0.177	0.204	0.160	0.092	0.086	0.078	0.309	0.263
	0.164	0.204	0.150	0.089	0.085	0.075	0.310	0.272
16	0.192	0.213	0.168	0.106	0.078	0.077	0.238	0.262
	0.193	0.206	0.175	0.111	0.075	0.075	0.231	0.262
	0.196	0.209	0.162	0.105	0.077	0.077	0.246	0.250
17	0.167	0.182	0.138	0.111	0.074	0.068	0.268	0.287
	0.169	0.151	0.143	0.105	0.075	0.069	0.265	0.274
	0.177	0.179	0.145	0.101	0.078	0.063	0.274	0.280

Table 3: Elemental carbon loadings (µg/cm^2)

Laboratory	IPR 1	IPR 2	IPR 3	A 210	A 223	A 224	TER 10	TER 11
1	2.804	2.028	1.371	0.897	0.710	0.631	4.934	6.657
	2.823	2.044	1.351	0.857	0.677	0.586	5.188	6.369
	2.799	2.015	1.331	0.788	0.673	0.570	5.418	6.360
2	2.085	1.617	1.032	0.755	0.542	0.370	4.592	5.040
	2.146	1.716	1.069	0.720	0.489	0.382	4.295	5.005
	2.131	1.659	1.112	0.661	0.535	0.408	4.144	5.293
3	2.869	2.003	1.592	0.888	0.606	0.617	4.688	6.204
	2.566	1.981	1.332	0.996	0.715	0.574	4.872	5.673
	2.988	1.938	1.267	0.844	0.498	0.736	5.435	5.781
4	2.190	2.400	1.090	0.900	1.470	0.560	4.340	5.320
	2.370	2.010	1.150	1.890	0.940	0.720	4.500	5.760
	2.270	1.990	1.350	0.970	1.410	0.770	4.820	6.310
5	1.800	1.450	1.100	0.920	0.640	0.460	4.310	4.720
	1.910	1.460	1.080	0.800	0.620	0.470	4.360	4.320
	2.010	1.550	1.150	0.720	0.630	0.460	4.410	4.690
6	2.214	1.755	1.287	0.937	0.778	0.608	6.142	6.022
	2.291	1.893	1.283	0.944	0.802	0.708	5.899	6.335
	2.253	1.916	1.318	0.801	0.872	0.623	5.670	6.529
7	3.440	2.550	1.510	0.900	0.640	0.580	5.500	7.470
	3.450	2.370	1.630	0.910	0.590	0.630	6.040	6.370
	3.690	2.440	1.630	0.930	0.640	0.580	5.270	6.890
8	2.202	1.569	1.316	0.787	0.581	0.518	4.931	5.516
	2.176	1.708	1.130	0.641	0.634	0.579	4.409	5.387
9	2.628	1.767	1.347	0.975	0.769	0.609	5.200	6.212
	2.639	2.122	1.432	0.888	0.743	0.655	5.219	6.828
	2.703	1.891	1.596	0.919	0.672	0.659	5.157	5.900
10	4.466	3.451	1.918	1.042	1.048	0.254	7.073	9.217
	5.012	3.520	2.285	1.167	0.908	0.811	7.838	9.683
	4.141	2.865	2.128	1.182	1.023	0.987	6.884	9.183
11	3.138	2.285	1.487	1.035	0.884	0.716	5.679	6.718
	3.104	2.434	1.483	1.047	0.870	0.719	6.120	7.087
	3.126	2.543	1.451	1.070	0.854	0.706	5.955	6.995
12	2.445	1.890	1.339	0.848	0.745	0.617	5.491	5.888
	2.478	1.909	1.364	0.884	0.724	0.653	5.458	5.934
	2.402	1.973	1.425	0.889	0.793	0.647	5.483	6.170
13	2.491	1.974	1.464	0.876	0.615	0.593	5.650	6.234
	2.469	1.927	1.439	0.886	0.632	0.539	5.285	5.952
	2.649	1.895	1.403	0.854	0.614	0.605	5.564	5.940
14	1.927	1.604	1.142	0.835	0.607	0.512	5.184	5.735
	2.152	1.622	1.726	0.845	0.585	0.545	5.267	5.758
	2.126	1.618	1.065	0.804	0.541	0.894	5.119	6.101
15	2.480	1.900	1.386	0.833	0.815	0.590	5.649	5.155
	2.450	1.919	1.479	0.793	0.792	0.662	5.447	5.302
	2.287	1.938	1.310	0.815	0.788	0.633	5.360	5.536
16	2.761	2.130	1.521	0.876	0.764	0.681	4.266	5.382
	2.815	2.070	1.545	0.877	0.780	0.695	4.043	5.443
	2.811	2.081	1.500	0.833	0.772	0.679	4.566	5.113
17	2.250	1.745	1.183	0.855	0.709	0.593	4.954	6.121
	2.343	1.604	1.263	0.789	0.709	0.595	4.701	5.980
	2.496	1.697	1.328	0.786	0.757	0.562	4.802	5.949

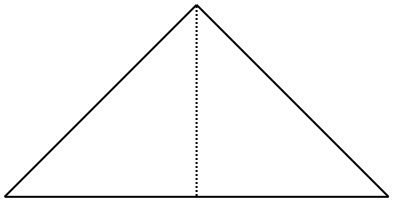
Table 4: Organic carbon [OC = TC-EC loadings] (µg/cm^2)

Laboratory	IPR 1	IPR 2	IPR 3	A 210	A 223	A 224	TER 10	TER 11
1	11.82	8.35	7.35	7.06	9.30	8.81	13.27	14.99
	11.76	8.18	7.17	6.88	9.60	8.41	13.79	14.79
	11.68	8.10	7.41	6.72	9.44	8.36	13.70	14.87
2	10.97	7.91	7.45	6.38	8.28	8.27	12.35	13.34
	11.49	7.96	6.92	6.30	8.82	7.52	13.06	13.69
	11.04	7.89	7.25	6.11	8.46	7.86	12.04	13.60
3	11.51	8.43	8.24	7.49	9.20	8.90	12.86	13.80
	11.12	8.00	7.54	6.96	10.19	10.16	12.10	14.03
	11.64	8.11	8.27	7.17	9.71	10.63	11.31	14.50
4	11.99	9.95	8.49	6.95	10.66	8.88	13.96	18.64
	11.95	10.89	8.06	6.78	12.39	9.59	14.50	16.73
	12.25	9.10	10.78	6.87	10.19	12.00	14.97	14.62
5	11.00	7.87	7.21	7.66	8.27	8.26	11.94	14.97
	11.05	8.07	6.88	7.00	8.49	8.16	12.50	14.72
	11.31	7.86	7.00	6.71	8.44	8.05	12.40	14.84
6	11.73	8.43	7.32	7.05	8.58	8.01	13.51	14.86
	11.73	8.14	7.21	7.86	8.68	8.21	13.11	15.63
	11.51	8.31	7.11	7.84	8.89	8.37	12.79	16.27
7	11.16	8.65	6.94	6.43	8.66	7.43	11.40	13.43
	11.35	7.52	7.60	6.52	8.45	8.09	12.36	14.13
	11.51	7.76	7.32	6.99	8.83	7.87	11.13	12.71
8	10.08	7.29	7.73	6.94	8.71	7.17	14.21	9.51
	12.17	7.33	5.62	6.66	8.52	8.24	13.75	9.64
9	10.40	7.19	7.25	6.50	8.16	7.57	12.44	12.60
	10.26	7.43	7.20	6.83	8.01	7.26	12.63	13.24
	10.10	7.50	7.05	6.46	8.10	7.45	11.65	12.53
10	17.65	12.03	8.87	10.15	12.53	13.51	18.83	19.98
	19.38	10.70	11.24	11.94	11.77	11.06	22.08	19.33
	14.71	8.46	12.52	10.56	14.10	11.42	20.00	21.04
11	11.15	7.83	7.48	7.09	8.88	8.32	13.40	13.71
	11.03	7.94	7.21	7.15	8.99	8.06	13.44	13.91
	11.07	8.25	7.43	7.27	9.01	8.15	13.18	14.03
12	11.51	7.78	7.78	7.01	8.98	8.15	12.94	14.23
	12.29	8.10	7.54	6.96	9.01	7.95	13.46	14.80
	11.39	7.91	7.62	6.69	8.90	8.15	13.92	15.28
13	10.91	7.67	7.60	6.97	9.38	8.01	13.19	14.52
	11.15	7.62	7.61	6.95	9.00	7.96	12.62	14.78
	11.41	7.44	7.38	6.85	8.59	8.47	12.92	14.84
14	12.44	9.01	8.38	7.19	9.84	8.77	14.71	16.16
	12.35	8.60	8.55	7.51	9.49	8.80	14.59	16.48
	12.27	8.80	8.08	7.10	9.51	8.86	14.90	16.55
15	11.31	7.85	7.73	7.22	8.64	7.84	12.43	14.20
	11.42	7.49	7.79	7.82	8.44	7.84	12.18	14.89
	11.67	7.55	7.42	8.36	8.50	7.84	11.91	14.79
16	11.59	7.89	7.54	7.38	9.09	8.20	13.69	15.17
	11.79	7.99	7.30	7.02	9.59	8.60	13.44	15.35
	11.51	7.87	7.76	7.08	9.19	8.19	13.96	15.36
17	11.19	7.81	7.35	6.79	8.86	8.11	13.47	15.17
	11.52	8.98	7.54	6.68	8.72	8.04	12.99	15.84
	11.53	7.76	7.77	6.93	8.94	8.26	12.68	15.27

Annex 2. QA measures

Calculation of QA variability = Random errors (2RSD)

It is assumed that laboratories taking part in inter-laboratory comparisons will obtain results near the expected ones when this bias is removed, and that the differences between expected and obtained results more often will be close to zero than not. Based upon this assumption, a triangular distribution can be used to quantify the random errors in the laboratory results (Eurachem, 2000; EMEP CCC report 6/2003).



The triangle distribution is symmetric with a baseline 2a. The height in the triangle will be 1/a when the triangle area equals 1. The standard uncertainty is given by

$$u(x) = \frac{a}{\sqrt{6}} \qquad (1)$$

The distance from –a to a (i.e. 2a) is called the range. When applied on the inter-laboratory comparison results, the range equals the distance between the largest and smallest of the differences between expected and found concentrations. L and T represent the laboratories' and the expected concentrations respectively, and D is the difference:

$$D_i = L_i - T_i \qquad (2)$$

The range (2a) is then the difference between the highest and minimum differences (D_{max} – D_{min}) and the uncertainty $u(D)$, for the differences becomes

$$u(D) = \frac{\left(D_{\max} - D_{\min}\right)}{\left(2 \cdot \sqrt{6}\right)}. \qquad (3)$$

and more than 95 % of the data will be within $\pm 2 \cdot u(D)$. The QA variability is defined as the relative standard deviation (RSD) given by the 95% confidence limit, thus:

$$\text{QA variability} \quad = \quad 2 \cdot RSD = \frac{2 \cdot u(D) \cdot 100}{\frac{\sum_{i=1}^{n} T_i}{n}}\% = \frac{n \cdot (D_{\max} - D_{\min})}{\sqrt{6} \cdot \sum_{i=1}^{n} T_i}\% \qquad (4)$$

Calculating the QA bias = systematic error (RB%)

An estimation of bias in single measurements requires a long data series, and only a few samples in a laboratory comparison will only give a very coarse estimate or indication of the bias. However looking at the bias in inter-laboratory comparison over years will give a good indication of the performance of the laboratory.

The absolute bias may be dependent upon the concentrations, though the relative bias are considered approximate constant for the concentrations range used in the comparisons. The differences D_i, as defined above are calculated as relative difference, and a median of these relative difference are defined as the QA bias. Median is chosen instead of average to avoid that one outlier get too high influence on the results.

$$\text{QA variability} \quad = \quad \text{RB} = \text{median} \left[\frac{D_i}{T_i} \% \right] \tag{5}$$

In Tables 1, 2, 3 are reported QA measures for TC, OC and EC from the present inter-laboratory comparison.

Table 1. QA bias and QA variability for TC

TC QA measure	QA_bias	QA_variability	
s-swu	3%	3%	
cnr-iia	-7%	5%	systematic
cnr-isac	2%	10%	
obs-mip	12%	8%	systematic
umh	-7%	6%	systematic
uba	1%	5%	
nilu	-3%	6%	
demokritos	-6%	20%	systematic
chmi	-7%	4%	systematic
gcri	45%	18%	systematic
ipis	1%	3%	systematic
lsce	0%	3%	
ggd	0%	4%	
isciii	6%	5%	systematic
univ-grenoble	-4%	5%	
cyi	1%	2%	
jrc	-2%	4%	systematic

Table 2. QA bias and QA variability for OC

OC QA measure	QA_bias	QA_variability	
s-swu	2%	3%	
cnr-iia	-5%	4%	systematic
cnr-isac	2%	11%	
obs-mip	17%	9%	systematic
umh	-2%	4%	
uba	1%	5%	
nilu	-5%	6%	systematic
demokritos	-5%	24%	systematic
chmi	-9%	6%	systematic
gcri	45%	19%	systematic
ipis	-1%	4%	
lsce	0%	2%	
ggd	-1%	4%	
isciii	9%	6%	systematic
univ-grenoble	-3%	7%	
cyi	2%	3%	
jrc	-1%	4%	

Table 3. QA bias and QA variability for EC

EC QA measure	QA_bias	QA_variability	
s-swu	0%	8%	
cnr-iia	-17%	12%	systematic
cnr-isac	2%	8%	
obs-mip	3%	20%	
umh	-21%	23%	systematic
uba	3%	17%	
nilu	12%	19%	
demokritos	-13%	8%	systematic
chmi	4%	6%	systematic
gcri	48%	56%	systematic
ipis	18%	14%	systematic
lsce	0%	7%	
ggd	1%	10%	
isciii	-5%	8%	
univ-grenoble	0%	17%	
cyi	8%	19%	
jrc	-7%	6%	