Field measurement intercomparison

Field measurements of dissolved oxygen concentration

Mirja Leivuori, Teemu Näykki, Ivo Leito, Irja Helm, Lauri Jalukse, Lari Kaukonen, Panu Hänninen and Markku Ilmakunnas

Finnish Environment Institute

Field measurement intercomparison

Field measurements of dissolved oxygen concentration

Mirja Leivuori, Teemu Näykki, Ivo Leito, Irja Helm, Lauri Jalukse, Lari Kaukonen, Panu Hänninen and Markku Ilmakunnas

European Metrology Research Programme
▶ Programme of EURAMET
The EMRP is jointly funded by the EMRP participating countries
within EURAMET and the European Union

CONTENT

1 Introduction

In the framework of the European Metrology Research Programme (EMRP) project ENV05 OCEAN (Metrology for ocean salinity and acidity)[1], the dissolved oxygen concentration field (*in situ*) intercomparison (FieldOxy 2014) test was organized onboard R/V Aranda on April 23, 2014 in the Gulf of Finland (location called as "LL7": 59°50.79', 24°50.27'). The aim of the intercomparison was to enable the participants to assess their performance in measuring dissolved oxygen concentration in seawater under field conditions. The intercomparison measurement was organized jointly by the Finnish Environment Institute (Proftest SYKE, Envical SYKE) and University of Tartu (UT).

The proficiency test was carried out in accordance with the international guidelines ISO/IEC 17043 [1], ISO 13528 [2] and IUPAC Technical report [3]. The Proftest SYKE is accredited by the Finnish Accreditation Service as a proficiency testing provider (PT01, ISO/IEC 17043, www.finas.fi). This intercomparison test has not been carried out under the accreditation scope of the Proftest SYKE.

2 Organizing the proficiency test

2.1 Responsibilities

The responsibles for organizing the field intercomparison were Teemu Näykki and Lari Kaukonen (Envical SYKE), Mirja Leivuori (Proftest SYKE) and Ivo Leito (UT). Technical expertise was provided by Irja Helm, Lauri Jalukse (UT) and Keijo Tervonen (Proftest SYKE). Report layout was made by Markku Ilmakunnas (Proftest SYKE).

The contact of the organizers:
Proftest SYKE, Finnish Environment Institute (SYKE), Laboratory Centre,
Hakuninmaantie 6, 00430 Helsinki, Finland
Phone: +358 295 251 000, Fax. +358 9 448 320
http://www.syke.fi/proftest/en

University of Tartu, Institute of Chemistry
Ravila 14a, Tartu 50411, Estonia
http://www.chem.ut.ee/

---

[1] For more information, please see the ENV05 website: http://www.ptb.de/emrp/env05.html

2.2 Participants

Total of 21 participants from 10 institutes in Finland, Estonia, France, Germany and Sweden participated in the intercomparison (Table 1, Figure 1). Totally, 13-18 oxygen sensors were tested depending of the test depth. 15 optical and 3 electrochemical oxygen sensors were used. Additionally, six Winkler titrimetric setups participated in the intercomparison. The sensors used by Proftest SYKE are shown as labcode 4, 18 and 21 in the result tables. The metrologically traceable Winkler titration result by UT (the assigned value) is shown as labcode 12.

Table 1. The participants in the test.

Estonia	Estonian Marine Institute, University of Tartu
	University of Tartu
Finland	EHP-Tekniikka
	HSY Käyttölaboratorio Pitkäkoski
	Hyxo Oy
	SYKE Laboratory of Hakuninmaa, Helsinki
	SYKE Marine Research Centre, Helsinki
	SYKE Freshwater Centre, Oulu
France	IFREMER France
Germany	Federal Maritime and Hydrographic Agency Germany
Sweden	Stockholm University, Department of Ecology, Environment and Plant Sciences
	Umeå Marine Sciences Centre

Figure 1. Participants in the oxygen field intercomparison onboard Aranda 2014.

2.3 Testing site and times

The testing site situated two hours away in the offshore area in front of Helsinki (59°50.79', 24°50.27'). The station is called LL7 and it has been applied for a long time as monitoring site for water quality in the Baltic Sea. Based on the vertical water CTD profile measured onboard Aranda, the stable testing depths were chosen (Figure 2). The chosen testing depths were 5 m (D1), 23 m (D2) and 40 m (D3). Before the test the time of the participating sensors were synchronized with the provider's time. The testing times were: for 5 m 11:54, for 23 m 13:45 and for 40 m 16:25. In the intercomparison, the water samples for Winkler titrations were collected using the water samplers mounted into Rosette onboard Aranda. Also the participant's oxygen sensors were mounted on the Rosette (Figure 3) and they were transported to the testing depths using depth control and winch (Figure 4).

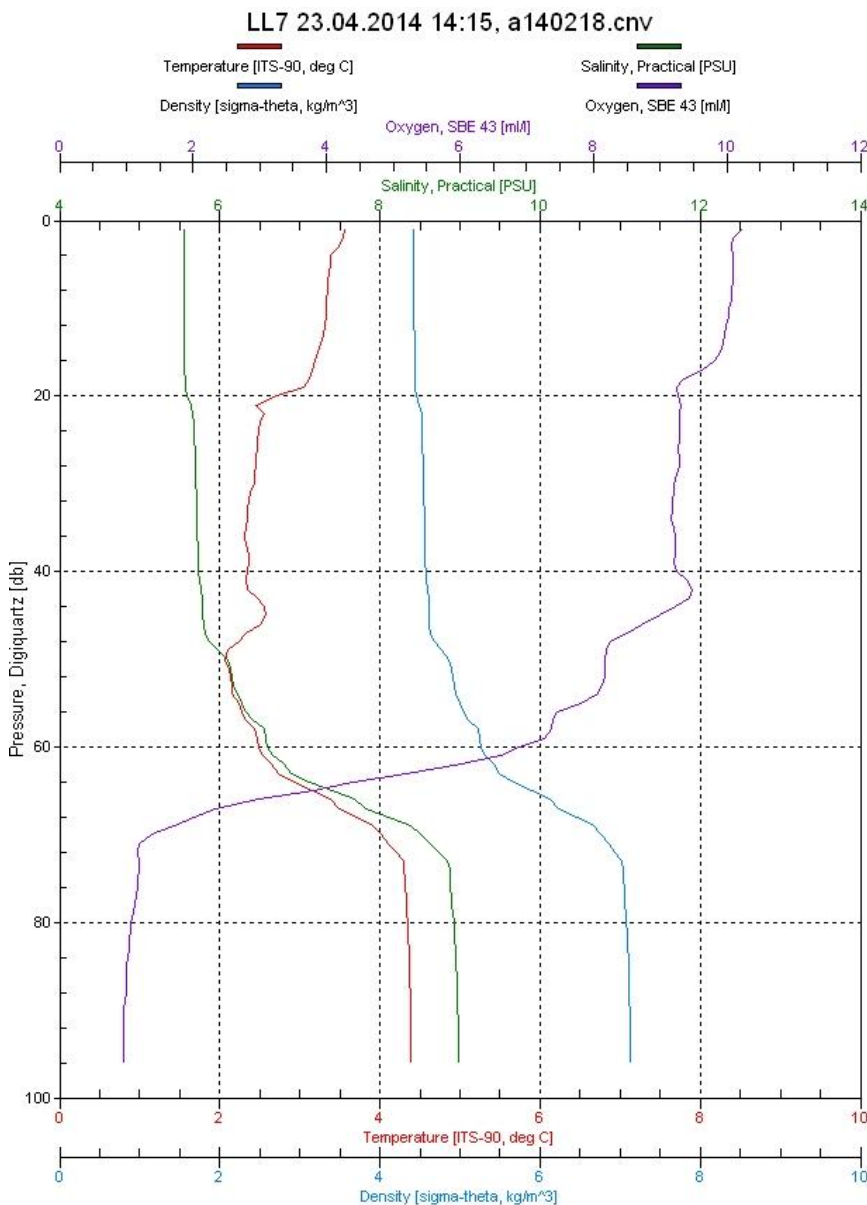Figure 2. The vertical CTD profile measured onboard Aranda at the testing site LL7. Please note, that in this figure the unit of DO (oxygen) concentration is ml/l deviating from the rest of report.
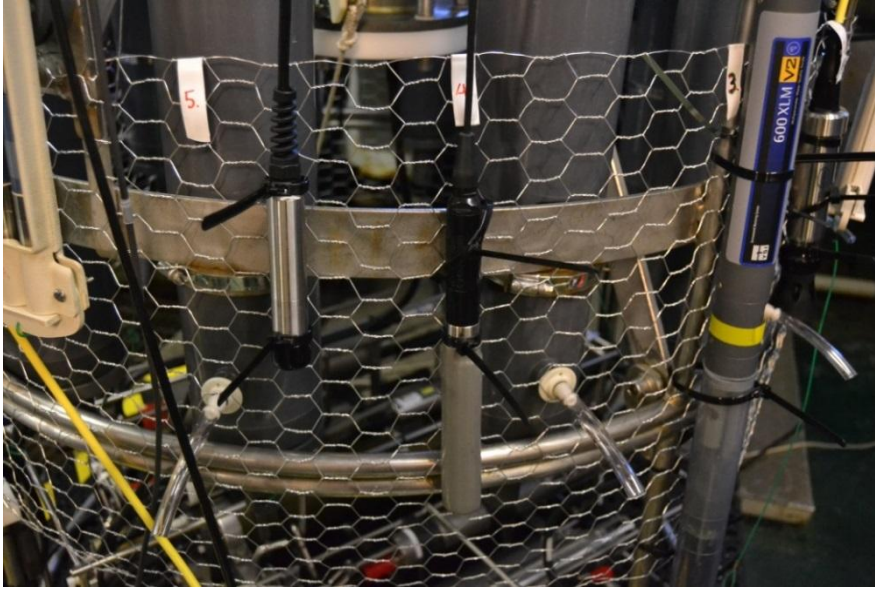
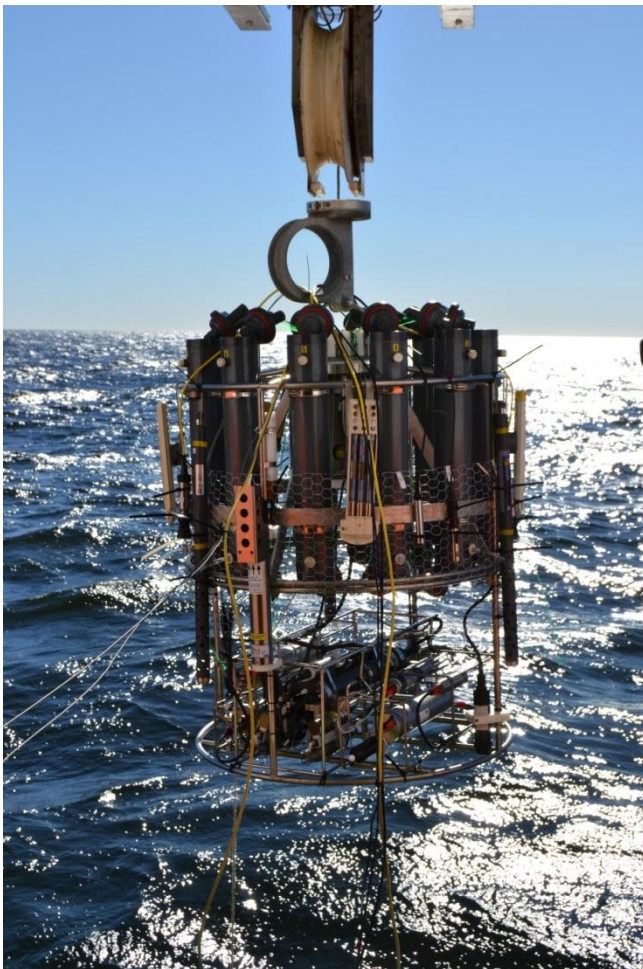Figure 3. The participant's oxygen sensors mounted on the Rosette for transporting to the testing depths.

Figure 4. The tested oxygen sensors, CTDs and water samplers going-on down to the testing depths.

2.4 Homogeneity and stability studies

The homogeneity and temporal stability of the sea water at the intercomparison site was studied by measuring dissolved oxygen with three oxygen sensors YSI 600 XLM V2. These sensors were mounted in the three different positions on the Rosette for obtaining information on spatial heterogeneity at each testing depths (Figure 5). During the intercomparison homogeneity test measurements were carried out at each depth from 5 minutes before to 5 minutes after the testing time.

In addition, the contribution of inhomogeneity of DO concentration was numerically evaluated by Winkler titration performed with the UT setup from rosette sampling vessels 1, 4, 7 and 10 (Figure 5).The contribution of DO concentration inhomogeneity was taken into account in the uncertainty estimate of the assigned (reference) value.
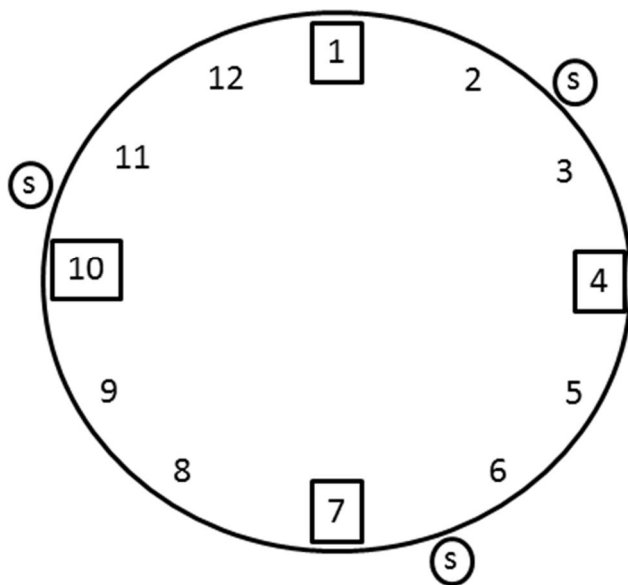

Figure 5. A top view of the Rosette used water sample collection and mounting the sensors. Numbers 1-12 describe the 12 vessels applied for water sampling. Vessels marked with square were used for sampling the water for high-accuracy Winkler procedure (reference method). Positioning of the sensors (YSI 600 XLM V2) of PT provider for homogeneity testing are marked with letter "s".

As the intercomparison was dealing with *in situ* measurements from naturally flowing water, the sources of uncertainties were not as easily controlled as in laboratory environment. By conducting the homogeneity testing simultaneously with the actual intercomparison experiments it was tried to ensure that the measured water was sufficiently homogeneous for all the participants.

For homogeneity test, 10 results were recorded within 10 minute's timeframe by the sensor used by Proftest SYKE. Sensors recorded values measured from water body every fifth seconds. Two consecutive measurements were regarded as replicate measurements for one homogeneity test sample.

To assess whether the homogeneity is sufficient for the intercomparison, both spatial and between measurement time variability standard deviations during the test were obtained using ANOVA variance component analysis in accordance with the Eurachem/CITAC guide [4]. The total variation components were divided into three parts: time of the measurement (temporal heterogeneity), spatial heterogeneity of the test area and analytical precision (Appendix 1). The homogeneity of testing area was estimated based on the recommendation of the IUPAC Technical report [3], with the exception of using variation of temporal heterogeneity instead of variation due to heterogeneity of the test area. This was used as the sample column is not stable and time of measurement described better the flowing water. Because PT provider used three sensors for homogeneity testing (see Figure 5), the variation of heterogeneity of the test area included also analytical variability.

2.5 Feedback from the proficiency test

The feedback received from the PT participants was very positive. The arrangements were successful and facilities were fit for purpose. It was appreciated that they could participate for free of charge including meals and accommodation in the ship due to the financial support of the project. The participants wished that similar intercomparisons would be organized also in the future.

2.6 Processing the data

2.6.1 Pretesting the data

The normality of the data was tested by the Kolmogorov-Smirnov test. The outliers were rejected according to the Grubbs or Hampel test before calculating the mean. More information about the statistical handling of the data is available from the Guide for participant [5].

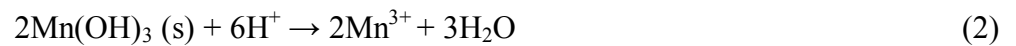### 2.6.2 Assigned values based on Winkler method

Winkler method (WM) was first published in 1888 by Hungarian analytical chemist Ludwig Wilhelm Winkler [6]. Although an old method, WM is still used for getting the reliable and traceable dissolved oxygen (DO) concentration values, because all sensors, in spite of being fast and convenient have a disadvantage: they all need to be calibrated with DO, the analyte.

The Winkler method is based on quantitative oxidation of Mn^{2+} to Mn^{3+} by oxygen in alkaline medium and on the subsequent quantitative oxidation of iodide to iodine by Mn^{3+} in acidic medium [7]. The formed iodine is titrated with thiosulfate.

First, two solutions (Winkler reagents) are added to the oxygen-containing sample: one containing I^- and OH^- and the other containing Mn^{2+}. Oxygen reacts under alkaline conditions with Mn^{2+} ions forming manganese(III)hydroxide [7]:

$$4Mn^{2+} + O_2 + 8OH^- + 2H_2O \rightarrow 4Mn(OH)_3 \downarrow \qquad (1)$$

The solution is then acidified. Under acidic conditions Mn^{3+} ions oxidize iodide to iodine, which eventually forms I_3^- ions with the excess of I^- [7]:

$$2Mn(OH)_3\ (s) + 6H^+ \rightarrow 2Mn^{3+} + 3H_2O \qquad (2)$$

$$2Mn^{3+} + 2I^- \rightarrow 2Mn^{2+} + I_2 \qquad (3)$$

$$I_2 + I^- \rightarrow I_3^- \qquad (4)$$

The concentration of the formed I_3^- ions is usually determined by titration with sodium thiosulfate solution:

$$I_3^- + 2S_2O_3^{2-} \rightarrow 3I^- + S_4O_6^{2-} \qquad (5)$$

Thiosulfate solution is standardized using potassium iodate (KIO_3) standard solution. So, the DO concentration in the sample is traceable to the KIO_3 mass. Under acidic conditions iodine is formed quantitatively according to the following reaction:

$$IO_3^- + 5I^- + 6H^+ \rightarrow 3I_2 + 3H_2O \qquad (6)$$

The procedure used for assigned value determination in FieldOxy 2014 is mainly based on the high-accuracy Winkler procedure [8]. This procedure has a number of modifications compared to the classical Winkler method (available e.g. as standard EN 25813 [9]) in order to achieve higher accuracy:

1) Oxygen content in the Winkler reagents is determined and accounted for (instead of using approximate values from literature [10]);

2) Iodine loss by volatilization is determined and accounted for and is additionally minimized by pre-titration;

3) Possible sample contamination with air is determined and accounted for as an uncertainty source;

4) Titration end-point is detected amperometrically using two Pt-electrodes.

The procedure uses gravimetric measurement of titrant and the KIO_3 solution is prepared and its amount measured gravimetrically [8]. It is impossible to weigh accurately on a ship. Therefore volumetric titration instead of gravimetric was used. The titrant standardization was also performed volumetrically. Standard solution was prepared in the laboratory by weighing certain amount of KIO_3 and dissolving it in 1 dm^3 (calibrated) flask. 5 ml of prepared standard solution was transferred into the titration vessel using calibrated glass pipette. Reagents were added and iodine was titrated. The Brand Liquid Handling Station LHS 600 was used for dosage of the titrant in case of titrant standardization as well as sample titrations.

The samples were not pure water (as in [8]), so the possible presence of oxidizing or reducing substances in the sea-water was determined after the testing cruise in the laboratory. It was done by using the same procedure as in the case of titrant standardization, but alternately (to eliminate all other influences) deionized water and sample water from the ship (5 ml each time in both cases) were added to the titrated solution. The relative differences between titrant concentrations in the case of these two titrations (at three depths, 5 replicates at every depth) and used as the estimates of uncertainty caused by possible interferences. They are converted to absolute (i.e. expressed in mg/dm^3) uncertainty estimates by multiplying the relative value with oxygen concentration determined at the same depth, multiplying also with the average sample volume and dividing by 5 ml. The corresponding standard uncertainty estimates are obtained by dividing with square root of three. The summary data of uncertainty due to possible oxidizing or reducing substances in the sea-water are presented in Table 2.

Table 2. Determination of uncertainty due to possible oxidizing or reducing substances in the sea-water.

Depth [m]	Relative differences between titrant concentrations [%]	$u_{\text{Interferences}}$ [mg/dm^3]
5	0.118	0.023
23	0.148	0.027
40	0.062	0.011

In order to be sure that the uncertainty due to oxidizing or reducing substances is not underestimated, for all testing depths the estimate $u_{\text{Interferences}} = 0.027$ mg/dm^3 was used.

In the intercomparison the DO concentration was determined separately by the above described high-accuracy Winkler procedure from the Rosette sampling vessels 1, 4, 7 and 10 at all testing depths (Figure 5). 17 sampling flasks (with calibrated volumes in the range of 11.1 to 11.5 ml) were used, so that 4 subsamples were taken from three sampling vessels and 5 from one. Some of the subsamples were discarded because of experimental failures (e.g. air bubbles in the flask, precipitate in the flask neck that was displaced by the reagent solution etc.). The numbers of used subsamples in the four sampling vessels at all depths are presented in Table 3. The subsamples were taken by thoroughly rinsing the sample bottles by ca 10-fold volume of the sampled water.

The results of determining the assigned values are presented in Table 3. At all depths the DO concentration assigned values ($C_{\text{O2\_Wink}}$ [mg/l]) are average values of the sampled Rosette vessels.

The measurement uncertainties were calculated mainly according to the same principles as in [8], except that volumetric solution measurement was used instead of weighing and the uncertainty source taking into account possible interferences (described above) was added. The uncertainty sources with their contributions in the case of one Rosette sampling vessel at one depth are presented in Figure 6 as an example (the descriptions of all input quantities not described here can be found in [8]).

The combined standard uncertainties ($u_c(C_{\text{O2}})$ [mg/l], see Table 3) take into account the averaged uncertainty of the Winkler titration procedure ($u_c(C_{\text{O2\_Wink\_averaged}})$ [mg/l], calculated as the pooled standard uncertainties of all the subsamples, as well as the differences between the Rosette sampling vessels (u(between vessels) [mg/l]). The latter uncertainty (which is the dominating uncertainty contribution at 5 m and 40 m depths) is expected to account for the inhomogeneity of DO concentration around the Rosette. The expanded uncertainties $U(C_{\text{O2}})$ were found with 95% coverage probability, taking into account the effective number of degrees of freedom. Because the inhomogeneity is taken into account, the assigned values are expected to be valid assigned values for all DO measurement devices attached to the Rosette, as well as the Winkler titration results of samples taken from other sampling vessels.

M$_{KIO3}$
0.00%

V$_{flask}$
1.10%

m$_{KIO3\_s}$
0.74%

$\Gamma^{V_{KIO_3}}_{V_{Na_2S_2O_3\_KIO_3}}$
1.5%

F$_{V\_KIO3}$
1.7%

F$_{V\_KIO3\_endp}$
0.2%

P$_{KIO3}$
2.71%

C$_{KIO3\_III}$
4.6%

F$_{I2\_vol}$
8.3%

F$_{V\_Na2S2O3\_KIO3}$
2%

C$_{Na2S2O3}$
18.1%

F$_{V\_Na2S2O3\_s}$
1.7%

$\Gamma^{V_{Na_2S_2O_3\_s}}_{V_s}$
40.9%

Int$_{O2}$
0.3%

M$_{O2}$
0.0%

Interferences
31%

F$_{V\_s}$
3.6%

CF$_{O2}$
0.8%

F$_{V\_s\_endp}$
0.5%

F$_{I2\_vol\_s}$
2.9%

C$_{O2\_s}$= 13.748
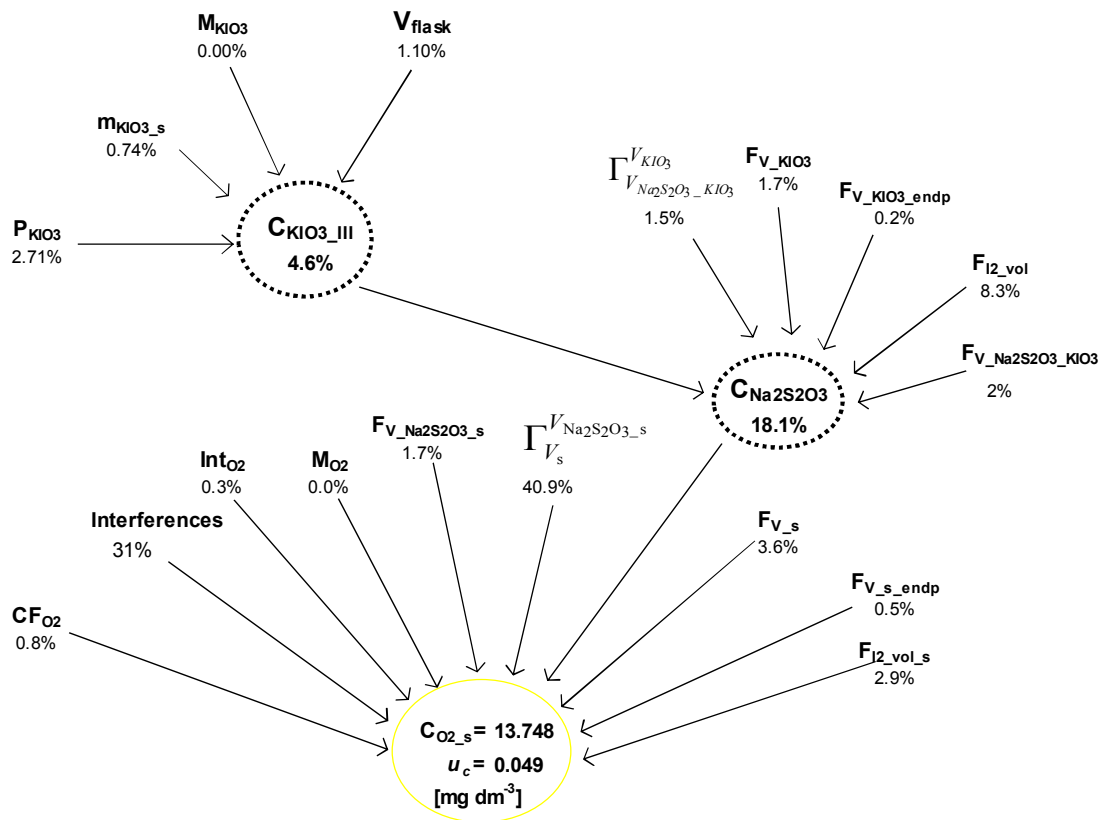u_c = 0.049
[mg dm^{-3}]

Figure 6. Cause-effect diagram at depth 23 m (Rosette vessel no 4).

Table 3. Results of Winkler titration for determination of assigned values for the Fieldoxy 2014 intercomparison (Gulf of Finland, 23.04.2014).

Depth (m)	Rosette no	C_{O2_Wink} [mg/l]	u_c (C_{O2_Wink}) [mg/l]	Number of values obtained	u_c ($C_{O2\_Wink\_averaged}$) [mg/l]	u (between vessels) [mg/l]	$u_c(C_{O2})$ [mg/l]	Effective degrees of freedom	Assigned value, C_{O2} [mg/l]	k (95% coverage probability)	$U(C_{O2})$ [mg/l]
5	all averaged	14.932		17	0.051	0.101	0.11	4.7	14.93	2.78	0.31
	1	14.875	0.047	4							
	4	14.966	0.040	4							
	7	14.843	0.075	4							
	10	15.067	0.039	5							
23	all averaged	13.794		16	0.045	0.033	0.06	16.9	13.79	2.12	0.12
	1	13.825	0.038	4							
	4	13.748	0.049	4							
	7	13.793	0.039	4							
	10	13.811	0.054	4							
40	all averaged	13.631		14	0.054	0.063	0.08	8.4	13.63	2.31	0.19
	1	13.533	0.089	3							
	4	13.679	0.038	4							
	7	13.648	0.044	3							
	10	13.640	0.043	4							

2.6.3 Standard deviation for proficiency assessment and z score

In this intercomparison, the assessment of the standard deviation for proficiency assessment (s_p) was based on perception and experience of the PT provider, taking into account the type of the sample, the concentration of the tested parameter, the results of homogeneity testing, the uncertainties of the assigned values and the long-term variation in previous proficiency tests for chemical laboratories. The target value for the standard deviation for proficiency assessment ($2*s_p$) was set 8 % for all testing depths. This is in accordance with the previous field intercomparison organized in 2013 [11, 12].

Additionally, the reliability was tested according to the criterion $u / s_p \leq 0.3$, where u is the standard uncertainty of the assigned value (the expanded uncertainty of the assigned value (U) divided by the coverage factor, which is show in Table 3) and s_p is the standard deviation for proficiency assessment [3]. The results show, that the criterion was fulfilled and the assigned values were considered reliable (Appendix 2).

The reliability of the target value of the standard deviation and the corresponding z score was estimated by comparing the deviation for proficiency assessment (s_p) with the robust standard deviation of the reported results (s_{rob}) [3]. The criterion $s_{rob} / s_p < 1.2$ was fulfilled.

3   Results and conclusions

3.1 Results

The terms used in the result tables are explained in Appendix 3.The results and the performance of each laboratory are presented in Appendix 4 and the summary of the results in Table 4. Results of replicate DO concentration for the participants using Winkler titration is shown in Table 5. The reported results with their expanded uncertainties ($k = 2$) are presented in Appendix 5. The summary of the z scores is shown in Appendix 6, participants' z scores in the ascending order in Appendix 7 and comparison of z and zeta scores in Appendix 8. The robust standard deviations of the results varied from 2.4 to 4.2 % (Table 4). The variability of participants' results was lower than expected for measurement in the field condition.

Table 4. The summary of the results in the field intercomparison.

Analyte	Sample	Unit	Assigned value	Mean	Rob. mean	Median	SD rob	SD rob %	$2*s_p$ %	n	Acc z %
O$_2$	D1_05	mg/l	14.93	14.90	14.89	14.90	0.36	2.4	8	24	92
	D2_23	mg/l	13.79	13.79	13.66	13.70	0.41	3.0	8	22	86
	D3_40	mg/l	13.63	13.70	13.56	13.63	0.57	4.2	8	19	84

Rob. mean: the robust mean, SD rob: the robust standard deviation, SD rob %: the robust standard deviation as percent, $2*s_p$ %: the total standard deviation for proficiency assessment at the 95 % confidence interval, Acc z %: the results (%), where $|z| \leq 2$, n: the number of the participants

Table 5. Results of the dissolved oxygen concentration (mg/l) for the participants using Winkler titration, who reported replicate results for their measurements. See results of laboratory 12 from Table 3.

LabCode	D1_05	D2_23	D3_40
11	15.18 15.12	13.92 13.89	13.88 13.92
14	15.02 14.96	14.05 13.92	13.83 13.75
20	14.79 14.74	13.76 13.79 13.73	13.68 13.66
24	14.637 14.663	13.557 13.570	13.423 13.546

3.2 Used oxygen measurement instruments

The participants were allowed to use different analytical methods for the oxygen measurements in the intercomparison. The used analytical methods and results of the participants grouped by methods are shown in more detail in Appendix 9. The statistical comparison of the analytical methods was not carried out for the data due to the low number of the results was ($n \leq 5$). Thus the comparison was carried out based on the graphical presentation.

Optical and electrochemical sensors

The participants used various oxygen sensors of which 15 were based on optical oxygen measurement and three were based on electrochemical oxygen measurement (Table 6). Based on the mean of the optical and electrochemical sensor and respectively recoveries, it is evident that the results of electrochemical oxygen sensors were lower than the results of optical oxygen sensors (Table 7).

Two possible reasons can be envisaged for the behavior of the electrochemical sensors. Firstly, the electrochemical sensors need water movement and if this is not sufficient then lowered readings are observed. Secondly, it is possible that the sensors' parameters have drifted during the time period from last calibration to the intercomparison. Such drift almost always leads to lowering of the values (not increasing), which is also observed here.

Table 6. The used oxygen sensors and Winkler titrimetric instrumentations in the test.

LabCode	Method	Oxygen sensor
1	Optical	SBE37-SMP-ODOMicroCat
2	Optical	Ponsel OPTOD
3	Electrochemical	OxyGuard Ocean Probe, attached to SAIV SD204 CTD
4	Optical	YSI ROX oxygen sensor YSI 600 XLM V2
5	Winkler titration	
6	Optical	YSI 6150 rox attached YSI 6600 V2
7	Optical	Hach HQ30d with sensor LDO10130
8	Optical	Hach Lange LDO101-30
9	Electrochemical	RBR duo T.DO
10	Optical	Ysi ProODO
11	Winkler titration	
12	Winkler titration (assigned value)	
13	Optical	Ysi ProODO
14	Winkler titration	
15	Optical	SS DO Sensor, Sea and Sun
16	Optical	Sea&Sun Fast Optical Oxygen Sensor (Optical DOSST)
17	Optical	Aanderaa Oxygen Optode 3835 + NKE Dortalogger
18	Optical	YSI ROX oxygen sensor YSI 600 XLM V2
19	Electrochemical	Dissolved Oxygen sensor SBE 13 attached Seabird SBE
20	Winkler titration	
21	Optical	YSI ROX oxygen sensor YSI 600 XLM V2
22	Optical	JFE Advantech, Rinko I aro-usb
23	Optical	Alec Rinko III
24	Winkler titration	

Table 7. The mean and recovery of the results of the optical and electrochemical oxygen sensors.

O$_2$ / sample	Optical		Electrochemical	
	Mean (mg/l)	Recovery (%)[1]	Mean (mg/l)	Recovery (%)[1]
D1_05	14.99	100.4	13.89	93.0
D1_23	13.83	100.3	12.30	89.2
D1_40	13.80	101.2	12.10	88.8

[1]Recovery (%) = 100*mean/assigned value

Winkler titrations

Four participants used iodometric method based on standard EN 25813 [9] in Winkler titrimetric titrations, one participant used the withdrawn SFS 3040 titrimetric method and UT used a high-accuracy titrimetric method [8]. The latter one was applied for establishing the metrologically traceable assigned values of DO concentration in the intercomparison. As seen in the Figures in Appendix 9 all the results based on Winkler titrimetric method were in the same range and all of them were acceptable for the intercomparison.

3.3 Measurement uncertainties of the results

Very few participants reported measurement uncertainties. The reported ones were mainly for Winkler titrimetric procedures and for some optical oxygen sensors operated by SYKE and UT. The comparison of z and zeta scores is shown in Appendix 8 for those participants who reported their measurement uncertainties.

Participants were encouraged to improve their analytical results by providing information about uncertainty of the measurement result. According to ISO 11352 [13] and Nordtest Handbook for measurement uncertainty estimation [14], uncertainty is broken down into two main components: (1) within-laboratory reproducibility and (2) method and laboratory bias. The first one covers the random effects of analytical results i.e. standard deviation of the measurement results. Sensor operator may for example record at least 10 replicate measurement results of same sample water in repeatability conditions and repeat this during five different days with the instrument calibrated just before the measurements. After that the (pooled) standard deviation (u_{Rw}) of the measurement results can be estimated.

The bias be calculated using the results of this PT. More detailed information for calculation of bias using PT results is described in Nordtest TR 537 [14]. Nordtest TR handbook suggests having at least six different PT results. However in this PT there were only three individual results available. When more PT data are available, the participants should revise their bias estimates. As the use of PT results for bias estimate is inferior to use of certified reference material for same purpose, the participants should also consider setting up facility for production of "in-house" reference material, water saturated with air, for dissolved oxygen determination. This is described in detail in reference [8].

4 Evaluation of the results

The evaluation of the participants was based on z scores, which were calculated using the assigned values and the estimated target values for the total standard deviation (Appendix 3). The z scores were interpreted as follows:

Criteria	Performance
$\|z\| \leq 2$	Satisfactory
$2 < \|z\| < 3$	Questionable
$\|z\| \geq 3$	Unsatisfactory

In total, 88 % of the results were satisfactory when total deviation of 8 % from the assigned values were accepted. More detailed summary of the type of oxygen sensor used or Winkler titrimetric results are shown in Table 8. Only three results were questionable and five results were unsatisfactory (Table 8, Appendix 7). The unsatisfactory results were found only for electrochemical oxygen sensors, which are based on the Clark cell type [16] measurement principle. Clark cell sensors measure DO indirectly through an electrochemical reaction. They are known to need careful and skilled maintenance, more frequent calibration and skilled

operation in order to perform well. This finding was similar than noticed in the previous intercomparisons [11, 15].

All electrochemical DO sensors have some flow dependency because they consume oxygen at the membrane surface. Therefore, water should be moving to obtain good measurement results, and in slow-moving water, mechanical stirring is necessary for most models. Under the sea conditions the Rosette is constantly moving due to movement of the ship in the sea. Also the water currents are moving around the Rosette and the sensors. In connection to this PT, additional tests were carried out by collaborator of EMPR project ENV05 (IOW Liebniz-institute for Baltic Sea research Warnemünde, Rostock, Germany) for flow dependency of DO measurement results on water flow rate in the surface of the sensor (SBE43; Seabird). It was noticed that flow velocities ca. 6-14 cm/s yielded DO results within 1.5% at 8.9 mg/l concentration level. If the flow rate of water was 0 cm/s, then the DO results were decreased dramatically resulting DO concentration ca 65% lower than compared to flow speed of 14 cm/s. However, the movement of the water during the PT may have been insufficient for the electrochemical sensors, and based on this dissolved oxygen concentration field measurement intercomparison, it cannot be reliably concluded that the electrochemical measurement principle is inferior to the optical one. For some oxygen sensors, the results were affected by the measurement depth and the measurement results were noticed to be systematically higher or lower. In these cases, the calibration and depth compensation of the oxygen sensor should be checked.

Table 8. Summary of the used oxygen sensor's type and performance (z score) in the field intercomparison.

LabCode	Sample_ depth (m)	z score	Oxygen sensor	Measurement principle
1	D1_05	-0,37	SBE37-SMP-ODOMicroCat	Optical
1	D2_23	-0,62	SBE37-SMP-ODOMicroCat	Optical
1	D3_40	-0,51	SBE37-SMP-ODOMicroCat	Optical
2	D1_05	-0.22	Ponsel OPTOD	Optical
2	D2_23	-0.16	Ponsel OPTOD	Optical
3	D1_05	-0.97	OxyGuard Ocean Probe,attached to SAIV SD204 CTD	Electrochemical
3	D2_23	-3.57	OxyGuard Ocean Probe,attached to SAIV SD204 CTD	Electrochemical
3	D3_40	-3.36	OxyGuard Ocean Probe, attached to SAIV SD204 CTD	Electrochemical
4	D1_05	-0.34	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
4	D2_23	-0.71	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
4	D3_40	-0.77	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
5	D1_05	-0.20	Winkler	Winkler
5	D2_23	0.14	Winkler	Winkler
5	D3_40	0.25	Winkler	Winkler
6	D1_05	-0.79	YSI 6150 rox attached YSI 6600 V2	Optical
6	D2_23	-0.91	YSI 6150 rox attached YSI 6600 V2	Optical
6	D3_40	-0.77	YSI 6150 rox attached YSI 6600 V2	Optical
7	D1_05	0.12	Hach HQ30d with sensor LDO10130	Optical
7	D2_23	-0.54	Hach HQ30d with sensor LDO10130	Optical
8	D1_05	0.45	Hach Lange LDO101-30	Optical
8	D2_23	-0.16	Hach Lange LDO101-30	Optical
9	D1_05	-3.23	RBR duo T.DO	Electrochemical
9	D2_23	-3.50	RBR duo T.DO	Electrochemical
9	D3_40	-4.18	RBR duo T.DO	Electrochemical
10	D1_05	0.40	Ysi ProODO	Optical
11	D1_05	0.37	Winkler	Winkler
11	D2_23	0.21	Winkler	Winkler
11	D3_40	0.50	Winkler	Winkler
12	D1_05	NA	Winkler (assigned value)	Winkler
12	D2_23	NA	Winkler (assigned value)	Winkler
12	D3_40	NA	Winkler (assigned value)	Winkler
13	D1_05	0.28	Ysi ProODO	Optical
14	D1_05	0.10	Winkler	Winkler
14	D2_23	0.35	Winkler	Winkler
14	D3_40	0.29	Winkler	Winkler

Table 8 continued.

LabCode	Sample_ depth (m)	z score	Oxygen sensor	Measurement principle
15	D1_05	1.16	SS DO Sensor,Sea and Sun	Optical
15	D2_23	1.65	SS DO Sensor, Sea and Sun	Optical
15	D3_40	1.82	SS DO Sensor, Sea and Sun	Optical
16	D1_05	0.52	Sea&Sun Fast Optical Oxygen Sensor (Optical DOSST)	Optical
16	D2_23	1.00	Sea&Sun Fast Optical Oxygen Sensor (Optical DOSST)	Optical
16	D3_40	1.32	Sea&Sun Fast Optical Oxygen Sensor (Optical DOSST)	Optical
17	D1_05	-0.37	Aanderaa Oxygen Optode 3835 + NKE Dortalogger	Optical
17	D2_23	-0.53	Aanderaa Oxygen Optode 3835 + NKE Dortalogger	Optical
17	D3_40	-0.02	Aanderaa Oxygen Optode 3835 + NKE Dortalogger	Optical
18	D1_05	-0.07	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
18	D2_23	-0.42	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
18	D3_40	-0.72	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
19	D1_05	-1.01	Dissolved Oxygen sensor SBE 13 attached Seabird SBE 43	Electrochemical
19	D2_23	-1.01	Dissolved Oxygen sensor SBE 13 attached Seabird SBE 43	Electrochemical
19	D3_40	-0.90	Dissolved Oxygen sensor SBE 13 attached Seabird SBE 43	Electrochemical
20	D1_05	-0.28	Winkler	Winkler
20	D2_23	-0.03	Winkler	Winkler
20	D3_40	0.07	Winkler	Winkler
21	D1_05	-0.03	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
21	D2_23	-0.45	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
21	D3_40	-0.75	YSI ROX oxygen sensor attached YSI 600 XLM V2	Optical
22	D1_05	2.33	JFE Advantech, Rinko I aro-usb	Optical
22	D2_23	2.28	JFE Advantech, Rinko I aro-usb	Optical
22	D3_40	2.15	JFE Advantech, Rinko I aro-usb	Optical
23	D1_05	0.45	Alec Rinko III	Optical
23	D2_23	0.27	Alec Rinko III	Optical
23	D3_40	0.35	Alec Rinko III	Optical
24	D1_05	-0.47	Winkler	Winkler
24	D2_23	-0.41	Winkler	Winkler
24	D3_40	-0.27	Winkler	Winkler

5   Summary

In the framework of the European Metrology Research Programme (EMRP) project ENV05 OCEAN (Metrology for ocean salinity and acidity), the dissolved oxygen field (*in situ*) intercomparison (FieldOxy 2014) test was organized onboard R/V Aranda on April 23, 2014 in the Gulf of Finland (location called as "LL7": 59°50.79', 24°50.27'). The aim of the intercomparison was to enable the participants to assess their performance in measuring dissolved oxygen concentration in seawater under field conditions. The intercomparison measurement was organized jointly by the Finnish Environment Institute (Proftest SYKE, Envical SYKE) and University of Tartu (UT).

Total of 21 participants from 10 institutes in Finland, Estonia, France, Germany and Sweden participated in the intercomparison. Totally, 13-18 oxygen sensors were tested depending of the test depth. Additionally, six Winkler titrimetric setups participated in the intercomparison. The metrologically traceable Winkler titration result (the assigned value) was measured by the Winkler setup of University of Tartu onboard R/V Aranda.

In total, 88 % of the results were satisfactory when total deviation of 8 % from the assigned values were accepted. Only three results were questionable and five results were unsatisfactory. A possible reason for several unsatisfactory results might be problems with calibration of electrochemical oxygen sensors. The electrochemical sensors need water movement and if this is not sufficient then lowered readings are observed. The movement of the water during the PT may have been insufficient for the electrochemical sensors, and based on this intercomparison, it cannot be reliably concluded that the electrochemical measurement principle is inferior to the optical one. For the most part the share of satisfactory results was very good.

6   Summary in Finnish

Euroopan metrologian tutkimusohjelman (EMRP) projektissa ENV05 OCEAN (Metrology for ocean salinity and acidity) järjestettiin meriveden liuenneen hapen kenttämittausten vertailukoe tutkimusalus Arandalla 23.4.2014. Vertailun tarkoituksena oli arvioida kentällä suoritettavien happimääritysten laatua ja keskinäistä vertailtavuutta. Vertailukokeen järjestivät Suomen Ympäristökeskus ja Tarton Yliopisto.

Vertailukokeeseen osallistui 21 osallistujaa kymmenestä eri laitoksesta Suomesta, Virosta, Ranskasta, Saksasta ja Ruotsista. Kaikkiaan kenttämittausvertailussa testattiin 13-18 happisensoria. Lisäksi testattiin kuusi Winklerin titrimetriseen määritykseen perustuvaa laitteistoa. Tarton Yliopiston määritti tutkimusaluksella metrologisesti jäljitettävän vertailuarvon perustuen Winkler titraukseen.

Kenttämittausvertailussa kaiken kaikkiaan 88 % tuloksista oli hyväksyttäviä, kun tulosten sallittiin vaihdella 8 % vertailuarvosta. Vain kolme tulosta oli kyseenalaisia ja viisi tulosta ei-hyväksyttäviä. Jälkimmäiseen tulokseen saattaa olla syynä elektrokemiallisten happisensorei-

den kalibrointiongelmat. Pääosin tulos oli hyvä, mutta elektrokemialliseen sensoritekniikkaan perustuvat kenttämittarit eivät menestyneet pätevyyskokeessa yhtä hyvin kuin optiseen mittaustekniikkaan perustuvat happisensorit. Edellisten käyttö edellyttää tarkkaa kalibrointia ja huolellista käyttöä kenttäolosuhteissa. Yksi syy poikkeaviin tuloksiin voi myös olla elektro-kemiallisten sensoreiden vaatima riittävä veden vaihtuminen mittauksen aikana. On mahdollis-ta, että tässä kenttämittausvertailussa veden liike ei ollut elektrokemiallisen mittauksen kannal-ta riittävää eikä siten luotettavasti voida todeta elektrokemialliseen mittaustapaan perustuvan tekniikan olevan huonompaa optiseen mittaustapaan perustuvaan tekniikkaan verrattuna.

REFERENCES

1. SFS-EN ISO 17043, (2010). Conformity assessment – General requirements for Proficiency Testing.

2. ISO 13528, (2005). Statistical methods for use in proficiency testing by interlaboratory comparisons.

3. Thompson, M., Ellison, S. L. R., Wood, R., 2006. The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry laboratories (IUPAC Technical report). Pure Appl. Chem. 78: 145-196. http://www.iupac.org.

4. Ramsey, M. H. and Ellison, S. L. R. (eds.) Eurachem/EUROLAB/CITAC/Nordtest/ACM Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches Eurachem, (2007). (http://eurachem.org/index.php/publications/guides/musamp).

5. Proftest SYKE Guide for laboratories: www.syke.fi/proftest/en → Running proficiency test (http://www.syke.fi/download/noname/%7B3FFB2F05-9363-4208-9265-1E2CE936D48C%7D/39886).

6. Winkler, L. W. (1888). Die Bestimmung des in Wasser Gelösten Sauerstoffes. Berichte der Deutschen Chemischen Gesellschaft, 21 (2): 2843–2855.

7. Carritt, D. E. and Carpenter, J. H., (1966). Comparison and evaluation of currently employed modifications of the Winkler method for determining dissolved oxygen in seawater. NASCO Report. J. mar. Res. 24: 286-318.

8. Helm, I., Jalukse L. and Leito I., (2012). A highly accurate method for determination of dissolved oxygen: gravimetric Winkler method. Anal. Chim. Acta. 74: 21-31.

9. EN 25813, (1992). Water quality - Determination of dissolved oxygen - Iodometric method. European Committee for Standardization.

10. Murray. C.N., Riley, J.P. and Wilson, T.R.S., (1968). The solubility of oxygen in Winkler reagents used for the determination of dissolved oxygen. Deep-Sea Res., 15: 237–238.

11. Leivuori, M., Björklöf, K., Näykki, T. and Väisänen, R.. Laboratorioiden välinen pätevyyskoe 5/2013. Kenttämittaukset - vesien happi, lämpötila, pH ja sähkönjohtavuus. Suomen ympäristökeskus. Suomen ympäristökeskuksen raportteja 30/2013. (http://hdl.handle.net/10138/41687).

12. Näykki, T., Leivuori, M., Björklöf, K., Väisänen, R., Laine. M., Väisänen, T. and Leito, I., (2014). Proficiency test of pH, conductivity and dissolved oxygen field measurements in river water. Accred. Qual. Assur., 19:259–268

13. ISO 11352, (2012). Water quality -- Estimation of measurement uncertainty based on validation and quality control data.

14. Magnusson, B. Näykki. T., Hovind, H. and Krysell, M., (2012). Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories. NT Technical Report 537. Nordtest. (http://www.nordtest.info/index.php/technical-reports.html)

15. Näykki, T., Jalukse, L., Helm, I. and Leito, I., (2013) Dissolved Oxygen Concentration Interlaboratory Comparison: What Can We Learn? Water. 5. 420-442.

16. Helm, I., Jalukse, L. and Leito, I., (2010) Measurement uncertainty estimation in amperometric sensors: a tutorial review. Sensors 10: 4430–4455.

APPENDIX 1: Homogeneity of the samples

Sample[a]		Mean	Unit	s_h (%)	s_p (%)	s_{obs} (%)	Component and it's standard deviation (%) during the measurement sequence			Criteria for homogeneity		
							s_a	s_t	s_s	$s_a/s_h<0.5$?	c	$s_t^2<c$?
O_2	D1_05	14.83	mg/l	1.0	4	0.70	0.02	0.11	0.69	yes	0.003	yes
	D2_23	13.51	mg/l	1.5	4	0.68	0.02	0.10	0.67	yes	0.005	yes
	D3_40	13.19	mg/l	1.1	4	0.62	0.02	0.09	0.62	yes	0.003	yes

[a] Time for measurement of sample D1: Ten points between 11:40 and 11:23.30. Time for measurement of sample D2: Ten points between 13:40 and 13: 50. Time for measurement of sample D3: Ten points between 16:20 and 16: 30.

s_h = Target standard deviation for homogeneity; s_p = Standard deviation for proficiency assessment;
s_{obs} = Observed total standard deviation = $\sqrt{s_a^2 + s_t^2 + s_s^2}$
s_a = Variation due to analytical precision; s_t = Variation due to temporal heterogeneity; s_s = Variation due to heterogeneity of the test area (sampling)
c = F1 x s_{all}^2 + F2 x s_a^2; s_{all}^2 = $(0.3$ x $s_p)^2$; F1 = 1.46 and F2 = 0.93

Conclusion: The analytical deviation fulfilled the criteria $s_a/s_p<0.5$ for each sample. Also in the each case the s_t^2 was smaller than the criteria c. The testing area could be regarded as homogenous.

APPENDIX 2: Evaluation of the assigned values and their uncertainties

Analyte	Sample	Unit	Assigned value	Standard uncertainty	Standard uncertainty, %	Evaluation method of assigned value	u/s_p
O_2	D1_05	mg/l	14.93	0.11	0.76	Metrologically traceable	0.18
	D2_23	mg/l	13.79	0.06	0.42	Metrologically traceable	0.11
	D3_40	mg/l	13.63	0.08	0.61	Metrologically traceable	0.15

Criterion for reliability of the assigned value $u/s_p \leq 0.3$, where:
s_p = target value of the standard deviation for proficiency assessment
u = standard uncertainty of the assigned value

Conclusion: The criterion was fulfilled and the assigned values were considered to be reliable.

APPENDIX 3: Terms in the results tables

Results of each participant

Analyte	The determined chemical species
Sample	The code of the sample
z score	Calculated as follows:

$z = (x_i - X)/s_p$, where

        x_i = the result of the individual laboratory

        X = the assigned value (*reference value*)

        s_p = the target value of the standard deviation for proficiency assessment

Assigned value	The reference value
2* s$_p$ %	The target value of total standard deviation for proficiency assessment (s_p) at the 95 % confidence level
Lab's result	The result reported by the participant (the mean value of the replicates)
Md	Median
Mean	Mean
SD	Standard deviation
SD%	Standard deviation, %
n (stat)	Number of results in statistical processing

Summary on the z scores

S – satisfactory ($-2 \leq z \leq 2$)

Q – questionable ($2 < z < 3$), positive error. the result deviates more than $2 \cdot s_p$ from the assigned value

q – questionable ($-3 < z < -2$), negative error. the result deviates more than $2 \cdot s_p$ from the assigned value

U – unsatisfactory ($z \geq 3$), positive error. the result deviates more than $3 \cdot s_p$ from the assigned value

u – unsatisfactory ($z \leq -3$), negative error. the result deviates more than $3 \cdot s_p$ from the assigned value

Robust analysis

The items of data are sorted into increasing order, $x_1, x_2, x_i. ....x_p$.

Initial values for x^* and s^* are calculated as:

x^* = median of x_i ($i = 1, 2. .....p$)

s^* = 1.483 · median of $|x_i – x^*|$ ($i = 1, 2. .....p$)

The mean x^* and s^* are updated as follows:

Calculate $\varphi = 1.5 \cdot s^*$. A new value is then calculated for each result x_i ($i = 1, 2 ...p$):

$$x_i^* = \begin{cases} x^* - \varphi, & \text{if } x_i < x^* - \varphi \\ x^* + \varphi, & \text{if } x_i > x^* + \varphi. \\ x_i, & \text{otherwise} \end{cases}$$
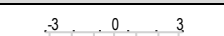
The new values of x^* and s^* are calculated from:
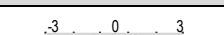
$$x^* = \sum x_i^* / p$$
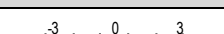
$$s^* = 1.134\sqrt{\sum (x_i^* - x^*)^2 /(p-1)}$$
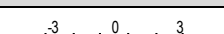
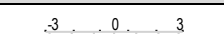The robust estimates x^* and s^* can be derived by an iterative calculation, i.e. by updating the values of x^* and s^* several times, until the process convergences [2].

APPENDIX 4: Results of each participant
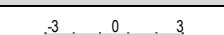
Laboratory 1												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.368	14.93	8	14.71	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.616	13.79	8	13.45	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.514	13.63	8	13.35	13.63	13.70	0.5	3.6	17
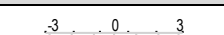
Laboratory 2												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.218	14.93	8	14.80	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.163	13.79	8	13.70	13.70	13.79	0.5	3.3	20
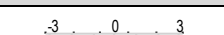
Laboratory 3												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.971	14.93	8	14.35	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-3.571	13.79	8	11.82	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-3.357	13.63	8	11.80	13.63	13.70	0.5	3.6	17

Laboratory 4												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.335	14.93	8	14.73	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.707	13.79	8	13.40	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.770	13.63	8	13.21	13.63	13.70	0.5	3.6	17
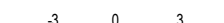
Laboratory 5												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.201	14.93	8	14.81	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		0.136	13.79	8	13.87	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		0.248	13.63	8	13.77	13.63	13.70	0.5	3.6	17
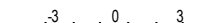
Laboratory 6												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.787	14.93	8	14.46	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.906	13.79	8	13.29	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.770	13.63	8	13.21	13.63	13.70	0.5	3.6	17
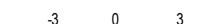
Laboratory 7												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		0.117	14.93	8	15.00	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.544	13.79	8	13.49	13.70	13.79	0.5	3.3	20
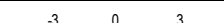
Laboratory 8												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		0.452	14.93	8	15.20	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.163	13.79	8	13.70	13.70	13.79	0.5	3.3	20
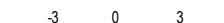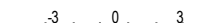
Laboratory 9												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		-3.265	14.93	8	12.98	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-3.499	13.79	8	11.86	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-4.182	13.63	8	11.35	13.63	13.70	0.5	3.6	17
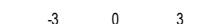
Laboratory 10												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		0.402	14.93	8	15.17	14.90	14.90	0.3	2.1	22
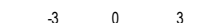
Laboratory 11												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		0.368	14.93	8	15.15	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		0.208	13.79	8	13.91	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		0.495	13.63	8	13.90	13.63	13.70	0.5	3.6	17
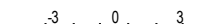
Laboratory 12												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		0.000	14.93	8	14.93	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		0.000	13.79	8	13.79	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		0.000	13.63	8	13.63	13.63	13.70	0.5	3.6	17
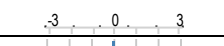
Laboratory 13												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		0.275	14.93	8	15.09	14.90	14.90	0.3	2.1	22

Laboratory 14												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		0.100	14.93	8	14.99	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		0.354	13.79	8	13.99	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		0.293	13.63	8	13.79	13.63	13.70	0.5	3.6	17
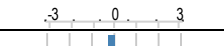
Laboratory 15												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		1.155	14.93	8	15.62	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		1.650	13.79	8	14.70	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		1.816	13.63	8	14.62	13.63	13.70	0.5	3.6	17

Laboratory 16												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		0.519	14.93	8	15.24	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		0.997	13.79	8	14.34	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		1.321	13.63	8	14.35	13.63	13.70	0.5	3.6	17
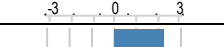
Laboratory 17												
Analyte	Unit	Sample	-3 0 3	z score	Assigned value	2*s$_p$, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O$_2$	mg/l	D1_05		-0.368	14.93	8	14.71	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.526	13.79	8	13.50	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.018	13.63	8	13.62	13.63	13.70	0.5	3.6	17
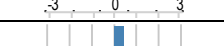
Laboratory 18												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.067	14.93	8	14.89	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.417	13.79	8	13.56	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.715	13.63	8	13.24	13.63	13.70	0.5	3.6	17

Laboratory 19												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-1.005	14.93	8	14.33	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-1.010	13.79	8	13.23	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.899	13.63	8	13.14	13.63	13.70	0.5	3.6	17

Laboratory 20												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.276	14.93	8	14.77	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.027	13.79	8	13.78	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		0.073	13.63	8	13.67	13.63	13.70	0.5	3.6	17

Laboratory 21												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.033	14.93	8	14.91	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.453	13.79	8	13.54	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.752	13.63	8	13.22	13.63	13.70	0.5	3.6	17

Laboratory 22												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		2.328	14.93	8	16.32	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		2.284	13.79	8	15.05	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		2.146	13.63	8	14.80	13.63	13.70	0.5	3.6	17

Laboratory 23												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		0.452	14.93	8	15.20	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		0.272	13.79	8	13.94	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		0.348	13.63	8	13.82	13.63	13.70	0.5	3.6	17

Laboratory 24												
Analyte	Unit	Sample	-3 . 0 . 3	z score	Assigned value	2*s_p, %	Lab's result	Md	Mean	SD	SD%	n (stat)
O_2	mg/l	D1_05		-0.469	14.93	8	14.65	14.90	14.90	0.3	2.1	22
	mg/l	D2_23		-0.411	13.79	8	13.56	13.70	13.79	0.5	3.3	20
	mg/l	D3_40		-0.267	13.63	8	13.48	13.63	13.70	0.5	3.6	17

APPENDIX 5: Results of participants and their uncertainties

In figures:

- The dashed lines describe the standard deviation for the proficiency assessment, the red solid line shows the assigned value, the shaded area describes the expanded measurement uncertainty of the assigned value, and the arrow describes the value outside the scale.

Analyte O_2 Sample D1_05

Analyte O_2 Sample D2_23

Analyte O$_2$ Sample D3_40
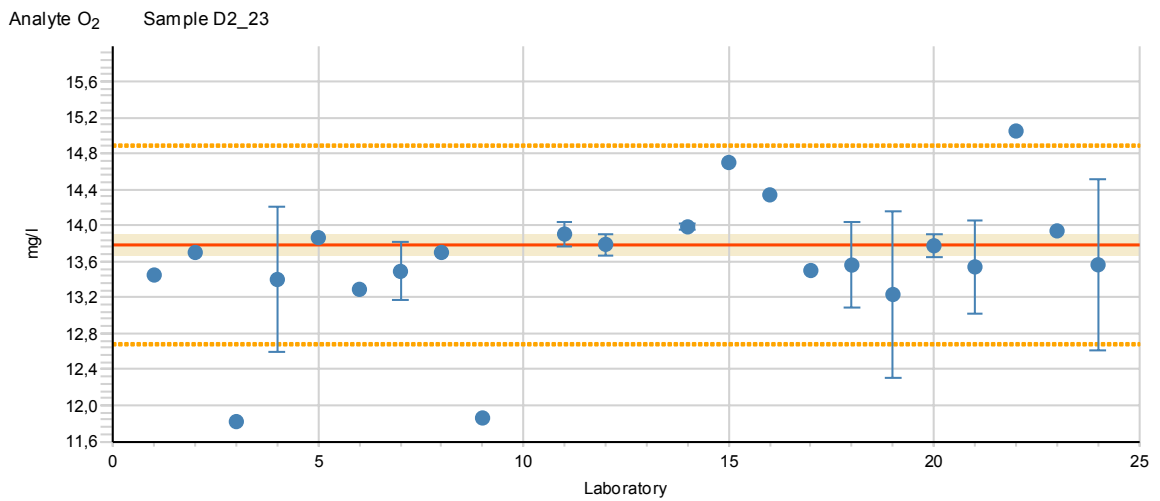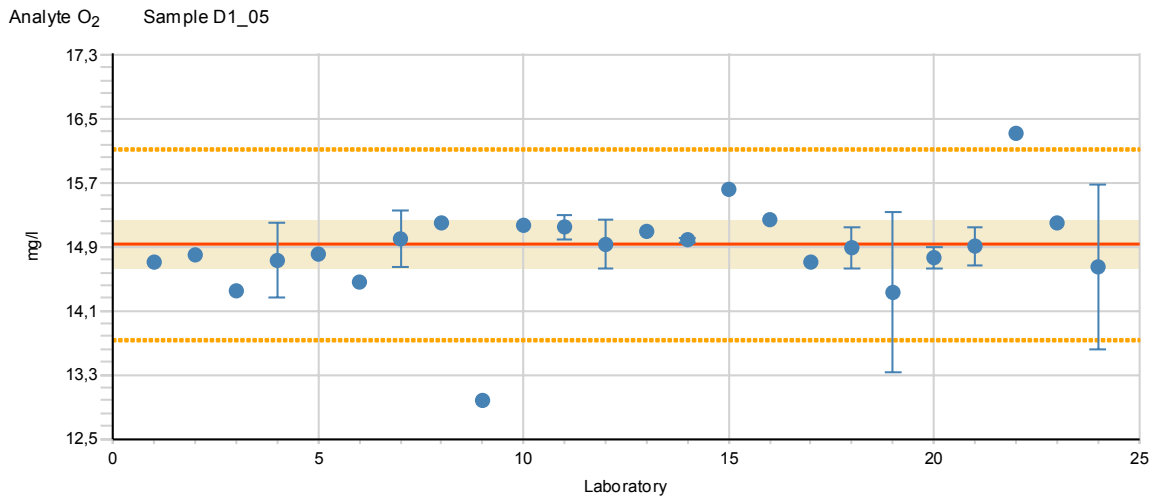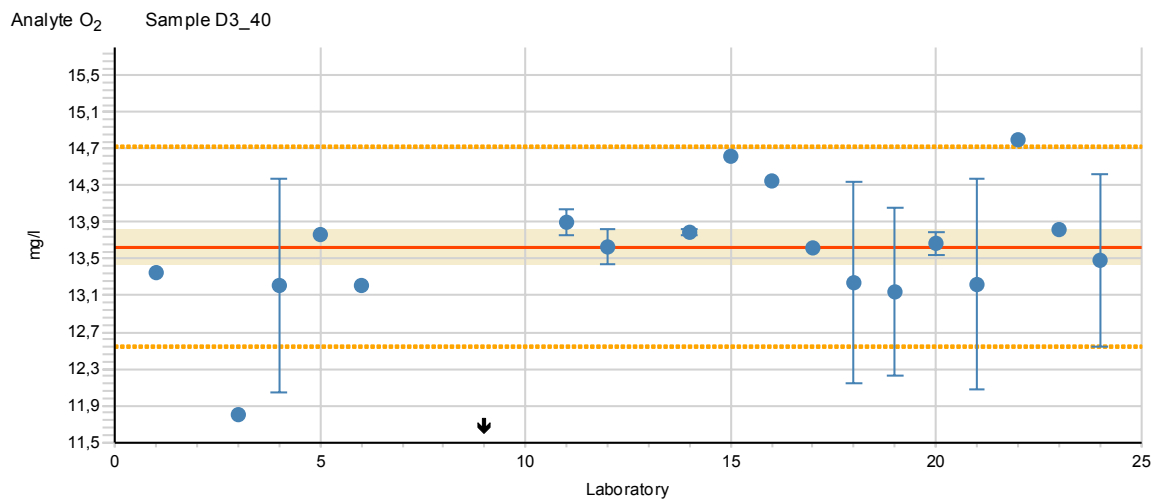
APPENDIX 6: Summary of the z scores

Analyte	Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	%
O$_2$	D1_05	S	S	S	S	**S**	**S**	**S**	**S**	u	*S*	**S**	**S**	*S*	**S**	**S**	**S**	**S**	**S**	**S**	*S*	*S*	Q	*S*	91.7
	D2_23	S	S	u	S	**S**	**S**	**S**	**S**	u	.	**S**	**S**	.	**S**	**S**	**S**	**S**	**S**	**S**	*S*	*S*	Q	*S*	86.4
	D3_40	S	.	u	S	**S**	**S**	.	.	u	.	**S**	**S**	.	**S**	**S**	**S**	**S**	**S**	**S**	*S*	*S*	Q	*S*	84.2
%		100	100	33	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100	100	100	0	100	
accredited						3	3	2	2	3		3	3		3	3	3	3	3	3					

Analyte	Sample	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	%
	D1_05	**S**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	91.7
	D2_23	**S**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	86.4
	D3_40	**S**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	84.2
%		100																							
accredited		3																							

S - satisfactory (-2 $\leq$ z $\leq$ 2), Q - questionable (2 < z < 3), q - questionable (-3 < z < -2),
U - unsatisfactory (z $\geq$ 3), u - unsatisfactory (z $\leq$ -3),
bold - accredited, italics - non-accredited, normal - other
% - percentage of satisfactory results

Totally satisfactory, % in all: 88   % in accredited: 93   % in non-accredited: 80

APPENDIX 7: z scores in ascending order

Analyte O$_2$ Sample D1_05
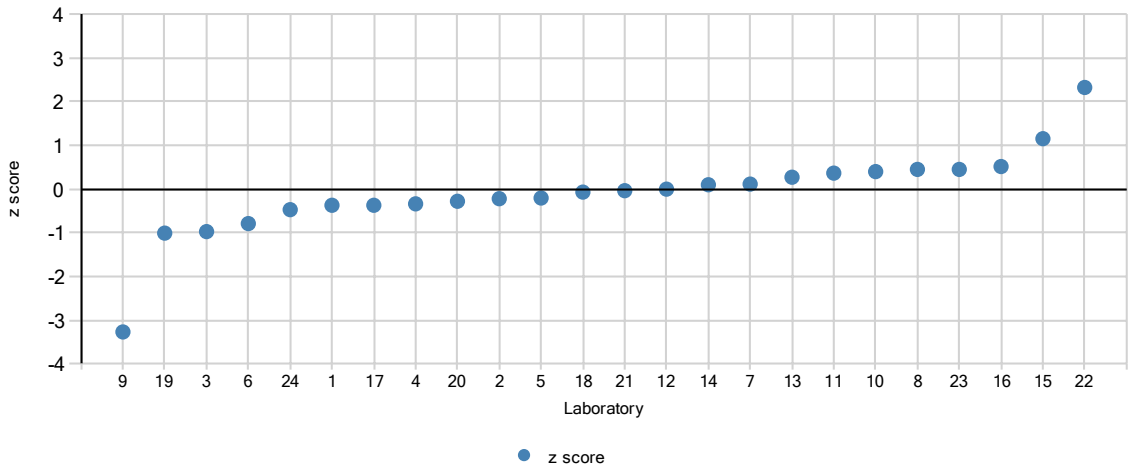
Analyte O$_2$ Sample D2_23
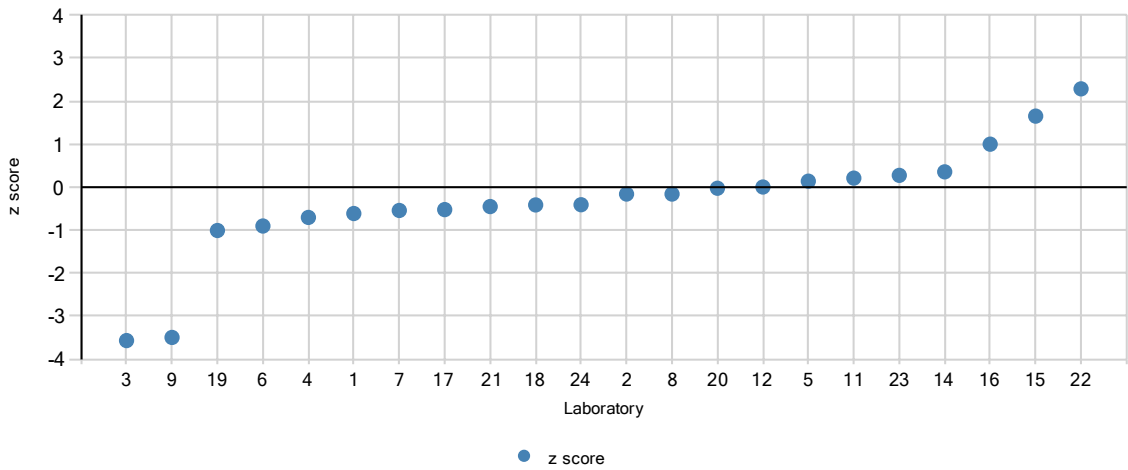
Analyte O$_2$ Sample D3_40
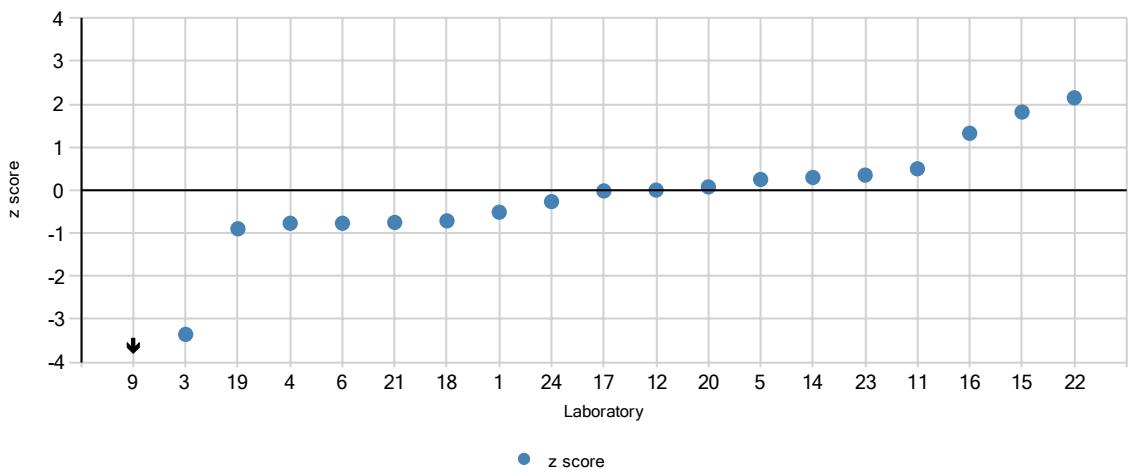
APPENDIX 8: Summary of the z and zeta scores

Zeta scores are not used for the evaluation of the performance of the laboratories. This information is however very useful when you re-evaluate the measurement uncertainties for your own laboratory (see below).

Explanations' for the z and zeta score sheet

Assigned value = the reference value
$k*u_C$ = the expanded uncertainty of the assigned value (%), the coverage factor k is shown in Table 3
2 s_p % = the target value for the total standard deviation at 95 % confidence interval

$$z = (x - X)/s_p, \text{ where}$$

x = the result of the individual participant
X = the assigned value
s_p = the standard deviation for proficiency assessment

$$zeta = (x - X)/\sqrt{u^2{}_{lab} + u^2{}_C}, \text{ where}$$

x = the result the assigned value of the individual participant
X = the assigned value
u_{lab} = the standard uncertainty of the participant's result
u_C = the standard uncertainty of the assigned value

How to interpret these results?

z score	zeta score	Action to take
Satisfactory	Satisfactory	No action; the result is good!
Satisfactory	Not satisfactory	The claimed uncertainty is too low, but it fills the requirement of the proficiency test.
Not satisfactory	Satisfactory	The result is within your claimed uncertainty, but not within the limits of proficiency test. The uncertainty might therefore be too high and should be checked against the uncertainty requirement of your client.
Not satisfactory	Not satisfactory	The result is too much biased and the reason should be clarified.

Analyte O_2 Sample D1_05
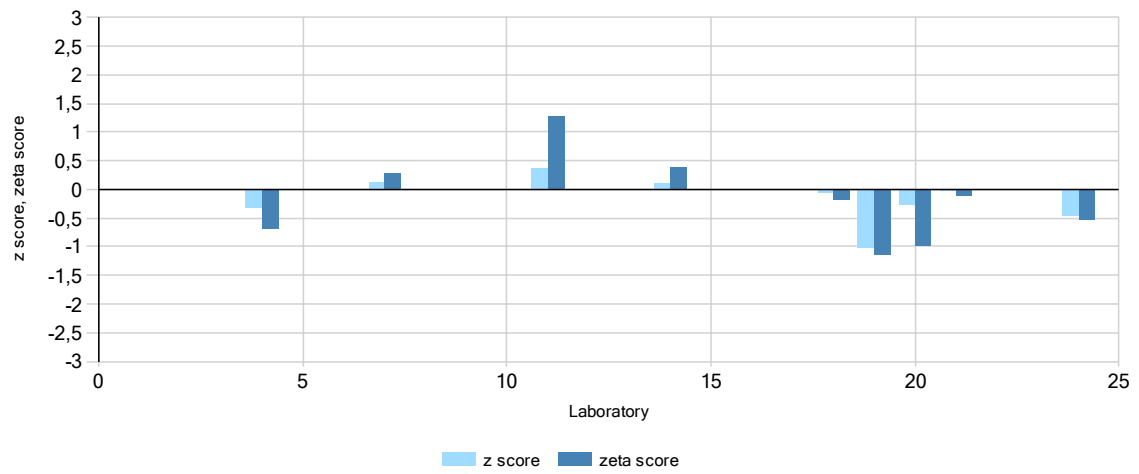
Assigned value	$k*u_C$	$2s_p$ %
14.93	2.1	8.0

Laboratory	Mean	U_{lab}, %	z	zeta
4	14.73	3.2	-0.33	-0.71
7	15.00	2.4	0.12	0.29
11	15.15	1.0	0.37	1.27
12	14.93	2.1	0.00	0.00
14	14.99	0.2	0.10	0.38
18	14.89	1.7	-0.07	-0.20
19	14.33	7.0	-1.00	-1.14
20	14.77	0.9	-0.28	-0.98
21	14.91	1.6	-0.03	-0.10
24	14.65	7.0	-0.47	-0.52

Analyte O$_2$ Sample D2_23
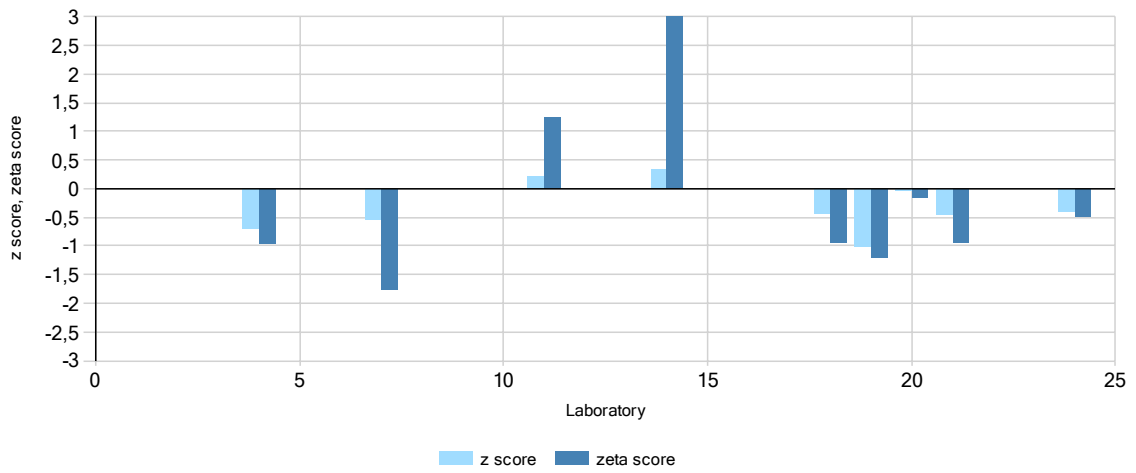
Assigned value k^*u_C	2s$_p$ %	
13.79	0.9	8.0

Laboratory	Mean	U$_{lab}$, %	z	zeta
4	13.40	6.0	-0.71	-0.96
7	13.49	2.4	-0.54	-1.76
11	13.91	1.0	0.21	1.25
12	13.79	0.9	0.00	0.00
14	13.99	0.2	0.35	3.17
18	13.56	3.5	-0.42	-0.94
19	13.23	7.0	-1.01	-1.19
20	13.78	0.9	-0.03	-0.17
21	13.54	3.8	-0.45	-0.95
24	13.56	7.0	-0.41	-0.47

Analyte O_2 Sample D3_40

Assigned value	k^*u_C	$2s_p$ %
13.63	1.4	8.0

Laboratory	Mean	U_{lab}, %	z	zeta
4	13.21	8.8	-0.77	-0.71
11	13.90	1.0	0.50	2.30
12	13.63	1.4	0.00	0.00
14	13.79	0.2	0.29	1.67
18	13.24	8.3	-0.72	-0.70
19	13.14	7.0	-0.90	-1.04
20	13.67	0.9	0.07	0.35
21	13.22	8.7	-0.75	-0.70
24	13.48	7.0	-0.27	-0.30

APPENDIX 9: Analytical methods

Results grouped according to the methods

Analyte O$_2$ Sample D1_05

✖ EN 25813 Iodometric method (Winkler)	◆ Withdrawn SFS 3040 Titrimetric method (Winkler)
■ High accuracy Winkler	★ Electrochemical oxygen sensor
▲ Optical oxygen sensor	

Analyte O$_2$ Sample D2_23

✖ EN 25813 Iodometric method (Winkler)	◆ Withdrawn SFS 3040 Titrimetric method (Winkler)
■ High accuracy Winkler	★ Electrochemical oxygen sensor
▲ Optical oxygen sensor	

Analyte O₂   Sample D3_40

Publisher	Finnish Environment Institute	Date	
		August 2014	
Author(s)	Mirja Leivuori, Teemu Näykki, Ivo Leito, Irja Helm, Lauri Jalukse, Lari Kaukonen, Panu Hänninen and Markku Ilmakunnas		
Title of publication	Field measurement intercomparison Field measurements of dissolved oxygen concentration		
Publication series and number	Reports of the Finnish Environment Institute 24/2014		
Theme of publication			
Parts of publication/ other project publications	The publication is available in the internet: www.syke.fi/publications | helda.helsinki.fi/syke		
Abstract	In the framework of the European Metrology Research Programme (EMRP) project ENV05 OCEAN (Metrology for ocean salinity and acidity), the dissolved oxygen field (*in situ*) intercomparison (FieldOxy 2014) test was organized onboard R/V Aranda on April 23, 2014 in the Gulf of Finland. The aim of the intercomparison was to enable the participants to assess their performance in measuring dissolved oxygen concentration in seawater under field conditions. The intercomparison measurement was organized jointly by the Finnish Environment Institute (Proftest SYKE. Envical SYKE) and University of Tartu. Total of 21 participants from 10 institutes in Finland, Estonia, France, Germany and Sweden participated in the intercomparison. Totally, 13-18 oxygen sensors were tested depending of the test depth. Additionally, six Winkler titrimetric setups participated in the intercomparison. The metrologically traceable Winkler titration result (the assigned value) was measured by the Winkler setup of University of Tartu onboard R/V Aranda. In total, 88 % of the results were satisfactory when total deviation of 8 % from the assigned values were accepted. Only three results were questionable and five results were unsatisfactory. A possible reason for several of the unsatisfactory results might be problems with calibration of electrochemical oxygen sensors. The movement of the water during the PT may have been insufficient for the electrochemical sensors, which may have resulted biased measurement results. Overall the share of satisfactory results was very good.		
Keywords	EMRP, Aranda, dissolved oxygen, field intercomparison, field measurement, Envical, Proftest, proficiency test, water analysis, water and environmental laboratories		
Financier/ commissioner			
	ISSN (pdf)	ISBN (online)	
	1796-1726 (pdf)	978-952-11-4356-4	
	No. of pages	Language	
	44	English	
	Restrictions		
	Public		
Distributor	Finnish Environment Institute (SYKE), neuvonta P.O. Box 140, FI-00251 Helsinki, Finland, Email: neuvonta.syke@ymparisto.fi		
Financier of publication	Finnish Environment Institute (SYKE), P.O. Box 140, FI-00251 Helsinki, Finland Phone +358 295 251 000		
Printing place and year	Helsinki 2014		