

Dissolved Oxygen Instrument Commissioning

The two new Dissolved Oxygen instruments purchased from SCRIPPS required testing to ensure they generate the same measurements as our existing instrument. Instruments were tested on RV Investigator voyage in2020_e01.

The new instruments are much newer than the existing Hobart and RV Investigator instruments. The dosimats are the newest versions - 876 series – which require a different software version. LVO2 software to work with the new dosimats has the dosimats model trailing, i.e. LVO2_876.

In this report the instruments are referred to as New A, New B and Old.

Table of Contents

Dissolved Oxygen Instrument Commissioning	1
Table of Contents	1
1. One Page Summary	3
2. Methods	4
2.1. Laboratory Setup	4
2.2. Instrument Configurations	4
2.3. Experimental Methods	5
2.3.1. Independent Iodate Standards	5
2.3.2. Repeated measurements of deep sample replicates (#1)	5
2.3.3. Repeated Measurements of Atmospheric Sample Replicates (All Instruments)	5
2.3.4. Repeated Measurements of Atmospheric Sample Replicates (Instrument New B)	6
2.3.5. Water Profile Comparison (Instrument New A & Old)	6
2.3.6. Repeated measurements of deep sample replicates (#2)	7
3. Results	8
Experiments Summary	8
3.1. Independent Iodate Standards	9
3.1.1. Iodate Standards across Instruments Boxplot	9
3.1.2. Iodate Standards Descriptive Statistics	9
3.2. Repeated Deep Sample Measurement: 1	10

3.2.1.	Samples from One Niskin per Instrument	10
3.2.2.	Samples from Two Niskins for all Instruments	12
3.3.	Atmospheric Saturated Sample: All Instruments.....	13
3.3.1.	Atmospheric Saturated Sample Boxplot (auto-scale).....	13
3.3.2.	Atmospheric Saturated Sample Boxplot (QC Control Lines)	14
3.3.3.	Descriptive Statistics	14
3.4.	Atmospheric Saturated Sample: One Instrument.....	14
3.4.1.	Atmospheric Saturated Sample: Instrument New B (auto-scale).....	15
3.4.2.	Atmospheric Saturated Sample: Instrument New B (auto-scale).....	15
3.4.3.	Atmospheric Saturated Sample: Instrument New B Boxplot (QC Control Limits).....	16
3.4.4.	Descriptive Statistics	16
3.5.	Water Profile Comparison.....	17
3.5.1.	Water Profile Plot.....	17
3.5.2.	T-Test Comparison of Means at specific depths.....	18
3.6.	Repeated Deep Sample Measurement: 2	18
3.6.1.	Deployment 2 Replicates Boxplot.....	19
3.6.2.	Descriptive Statistics	19
3.6.3.	T-Test Comparison of Means	19
3.7.	Experiment-Wide Meta-Analysis.....	20
3.7.1.	Calculation of Mean Measurement Standard Deviation	20
3.7.2.	F-Test Comparison of Variances	20
4.	Discussion.....	21
5.	Appendix	24
5.1.	CTD deployment locations	24

1. One Page Summary/Conclusion

The new dissolved oxygen instruments perform just as well as the old instrument, these experiments even found that the new instruments have potentially better measurement precision than the old. Instrument New B (marked B on right side inside box) has been installed and setup in the Hydrochemistry laboratory.

More testing is required to tune in the titration parameters as each sample can take roughly 20 - 30 seconds longer than on the old instrument. This is not an issue in general usage – but would become frustrating (and cumulative - very time consuming) on large oceanographic voyages. Important addendum, the new instruments and thermistors have the plug gender reversed compared to the old instrument. SITS could change the plug if required.

In these series of experiments the focus was on ensuring the new dissolved oxygen instruments could perform as well as the existing instrument and be suitable as replacements aboard RV Investigator. For the purposes of this investigation, a bigger focus was put upon the precision of measurements, rather than the exact accuracy.

To address any significant differences between instruments, most variables were attempted to be accounted for.

- Same temperature and air pressure environment
- Thermistors calibrated to same reference
- Burettes calibrated and volumes input to software
- Instruments standardised using same Potassium Iodate and Dosimat

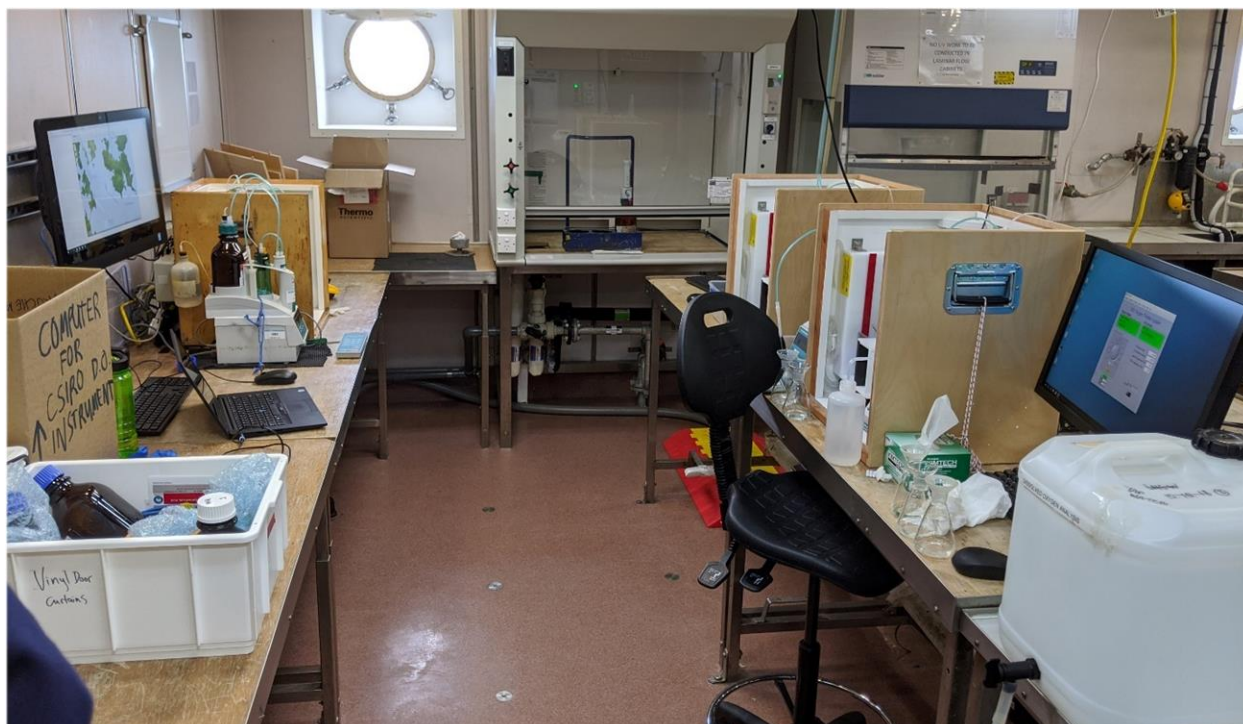
This meant the accuracy between instruments was often statistically indifferent, though there were 2 cases where instrument New A was very slightly offset from the two other instruments. This offset equated to approximately 0.3% ($<0.5\mu\text{M}$) in both instances. One case of this difference was likely due to an outlier samples from Niskin bottles, when these were removed there was no statistically significant difference between New A and Old.

In terms of precision, instrument New A had the tightest groupings of data when the cross experiment meta-analysis was completed. It was followed by instrument new B, with the old instrument having the worst precision across all experiments. However, these differences were statistically insignificant ($p=0.80$), perhaps likely due to the small sample size of standard deviations ($n=5$).

2. Methods

2.1. Laboratory Setup

The dissolved oxygen instruments were setup in the wet/clean laboratory on RV Investigator main deck. As the benchtops were sacrificial plywood, 40mm screw eyes were fixed into the bench. This allowed the instruments to be secured with ratchet straps or rope going through the instrument box handles and hooked onto the screw eyes. A similar method was used for securing the desktop computers, the monitors were simply screwed to the bench with screws through the base into the benchtop. There was two Potassium Iodate standard Dosimats, these had their own plastic containers which were fixed to the benchtop, which the Dosimats would then sit in.



2.2. Instrument Configurations

For the voyage significant variables were attempted to be controlled between all instruments. This list included:

- Same temperature and air pressure environment
- Thermistors calibrated to same reference
- Burettes calibrated and volumes input to software
- Instruments standardised using same Potassium Iodate and Dosimat
- Same titration parameters (where possible)

While this list of variables between instruments is not completely exhaustive, it was assumed it would be more than sufficient to produce comparable results between all 3 instruments.

The new instruments use the updated 876 Dosimat, as opposed to the existing instrument which has the older 676 model. While the titrator is mechanically very similar - can even use the old burettes - it has been significantly updated on the software side. This means the control system for titrating had to be updated for the 876 model, resulting in a fork of the SCRIPPS dissolved oxygen software LVO2. The older instrument uses the base LVO2 software, while the new instruments must use the software suffixed with 876, i.e. LVO2_876.

For the voyage software version 2.35 of the base software LVO2 was used with the older instrument. A different version was used for the 876 Dosimat instruments, 2.36g.

2.3. Experimental Methods

2.3.1. Independent Iodate Standards

This experiment was completed by dispensing a total of 12 Potassium Iodate standards from an independent Dosimat. The Dosimat was setup in the laboratory alongside the instruments and had its temperature monitored by one of the instruments temporarily. The Potassium Iodate was dispensed into flasks and then prepared in a similar fashion to standards used for standardisation. This involved topping the flasks 90% of the way up with Milli-Q water then adding reagents in a 3,2,1 direction.

These were then treated in the same and analysed in the same way as samples. For the analysis, flask 200 was entered for all of these, as in the test for comparability the volume should stay the same. In hindsight (and for a slightly simpler calculation) this bottle volume should just be set to 100mL for all.

2.3.2. Repeated measurements of deep sample replicates (#1)

The samples collected for this experiment were from deployment 1 of the voyage, where 5 Niskins were all fired at the cast bottom depth of 1000 meters. For each Niskin fired at the bottom depth, 6 dissolved oxygen samples were collected.

Part 1 of this experiment involved taking all the samples collected from 1 of the Niskins and measuring them on one instrument. So that would be the 6 samples collected at RP 3 all analysed on instrument New A etc.

In part 2 of this experiment, the remaining 2 Niskins worth of samples (12), was even split up between the instruments, meaning each instrument got 4 samples collected from 2 Niskins. b

2.3.3. Repeated Measurements of Atmospheric Sample Replicates (All Instruments)

Saturated dissolved oxygen samples at atmospheric pressure were collected using the new rig. The new rig is constructed from a round 20L Nalgene carboy fixed atop a stirring plate with an aquarium air stone inside. The carboy is filled with Milli-Q water and the aquarium air stone is

secured just under the water level, the water is then left bubbling and stirring for 24 hours before collection of samples.

Collection of atmospheric samples mimics the collection of dissolved oxygen samples from a Niskin. The Nalgene carboy has a tap and silicone tube attached, allowing for the exact same technique to be used. For the voyage the rig was setup next to the dissolved oxygen sink in the Hydrochemistry lab, this meant when collecting sample, excess water was draining into the sink (and not onto the floor).

During the collection of samples, an additional sample bottle was used to get an accurate estimate of the water temperature. The temperature probe was sitting in the bottle and used to measure the temperature a few times during collection. This was extremely consistent over the duration of sample collection.

Additionally, for calculation of the dissolved oxygen saturation, air pressure is required. The air pressure sensor in the Hydrochemistry laboratory was used to get this value, this was acquired from Grafana by looking at timestamps.

Twelve samples were collected from the rig, this took approximately 20 minutes to complete. During this time the water temperature was checked 4 times, all the same at 21.5°C. The air pressure did not vary significantly so an average over the sampling time was taken, 1010.7 atm.

The twelve samples were split up evenly amongst the instruments in a sequential pattern i.e. sample 1 for Old, sample 2 for New A, sample 3 for New B, sample 4 for Old, sample 5 for New A and so on.

2.3.4. Repeated Measurements of Atmospheric Sample Replicates (Instrument New B)

Samples were collected using the same technique as in section 2.3.3.

In this section 12 samples were collected from the rig, sampling temperature was 21.5°C and the averaged air pressure was 1010.7 atm. All samples were measured on instrument New B.

2.3.5. Water Profile Comparison (Instrument New A & Old)

For the water profile comparison samples were collected throughout the water column on deployment 1. Each Niskin had duplicate samples collected from it, the duplicates were then split so that one went to each of the tested instruments, either New A or Old.

DEPTH	RP FIRED
5	22, 23, 24
40	19, 20, 21
100	17, 18

200	15, 16
400	13, 14
600	11, 12
800	8, 9, 10
1000	3, 4

Table 2.3.5.1: Depths at which the different Niskins were fired at. There are some Niskins missing from the Table, those are due to the samples being used for other experiments. Namely 5,6,7 for deep replicates and 1,2 for making a BQC.

A T-Test for comparison of group means was used to ascertain if the measurements made were significant different from one another throughout the water column.

2.3.6. Repeated measurements of deep sample replicates (#2)

This experiment involved firing all Niskins at the bottom depth of voyage CTD deployment 2. The bottom depth for the deployment was 1000 meters where Niskins were fired quickly in sequential order.

A dissolved oxygen sample was collected from each Niskin, samples were then distributed evenly to each instrument being tested.

INSTRUMENT	ROSETTE POSITIONS
New A	2, 5, 8, 11, 14, 17, 20, 23
New B	3, 6, 9, 12, 15, 18, 21, 24
Old	1, 4, 7, 10, 13, 16, 19, 22

Table 2.3.6.1: Summary of the rosette positions where samples were collected and their respective instrument that was used to measure the sample.

3.Results

Experiments Summary

REF	EXPERIMENT OVERVIEW	INSTRUMENTS TESTED
3.1	Measurement of an independently dispensed Iodate standard as a sample	New A, New B, Old
3.2	Repeated measurements of deep sample replicates: 1	New A, New B, Old
3.3	Repeated measurements of atmospheric sample replicates	New A, New B, Old
3.4	Repeated measurements of atmospheric sample replicates	New B
3.5	Water Profile Comparison	New A, Old
3.6	Repeated measurements of deep sample replicates: 2	New A, New B, Old

3.1. Independent Iodate Standards

This section 3.1 includes results from the measurement of independently dispensed Potassium Iodate standards. Standards were dispensed from a separate 10mL Dosimat that was calibrated prior to the voyage. The Dosimat also used a different batch of Potassium Iodate to independently verify the Thiosulfate normality.

3.1.1. Iodate Standards across Instruments Boxplot

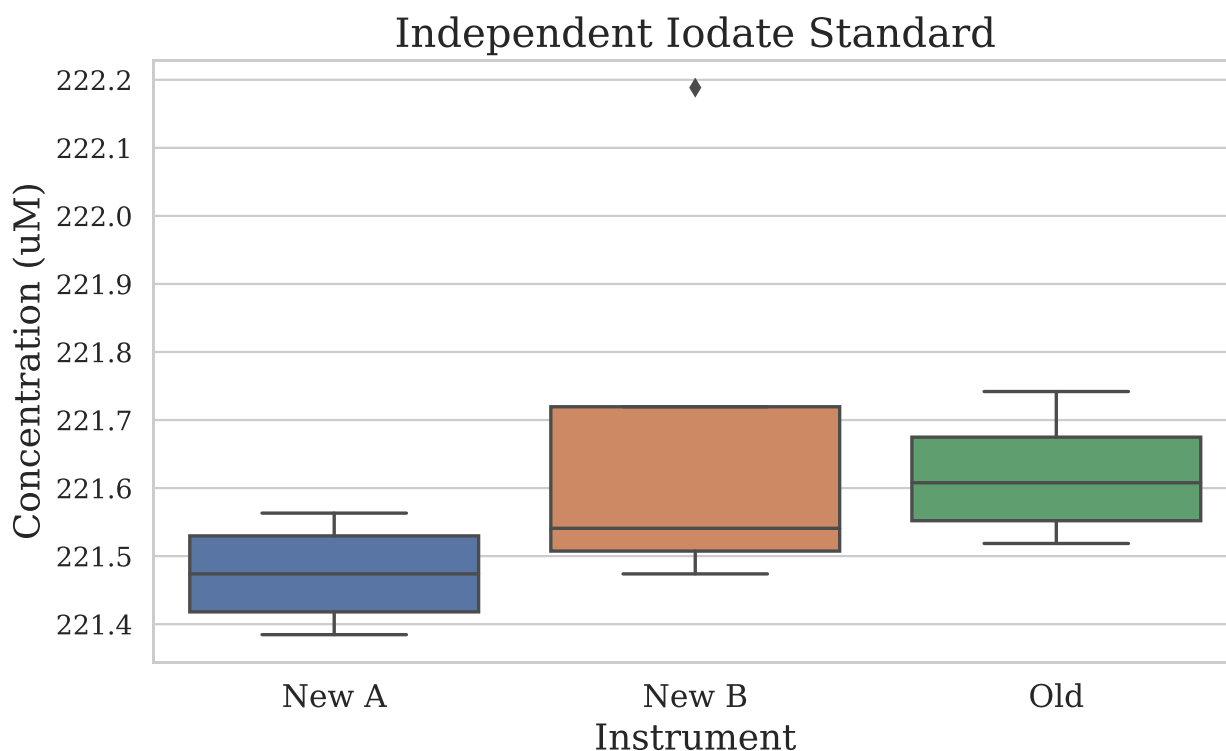


Figure 1.1.1: Boxplot style chart depicting the distribution of the independently dispensed Potassium Iodate standards measured by each instrument. Important to note that there was only 4 measurements made on each instruments.

3.1.2. Iodate Standards Descriptive Statistics

STAT	NEW A	NEW B	OLD
Mean	221.4739	221.6861	221.6191
Median	221.4739	221.5409	221.6079
Standard Deviation	0.0815	0.3369	0.0990
% RSD	0.037%	0.152%	0.045%
n	4	4	4

Table 1.1.2: Basic descriptive statistics of the independently dispensed Potassium Iodate standards.

3.2.Repeated Deep Sample Measurement: 1

Results section 3.2 pertains to the measurement of samples collected on deployment 1 from a depth of 1000 metres. Six sample replicates were taken from Niskins RP 3 to 7. For the first sub-section of this experiment, the 6 replicates from each Niskin were assigned to one instrument. For the second sub-section of this experiment, two replicates from each Niskin was measured by each instrument.

3.2.1. Samples from One Niskin per Instrument

For each instrument tested, the samples were collected from a single Niskin. See table below for a tabulated view.

INSTRUMENT	DEPLOYMENT	NISKIN (RP)	REPLICATES
New A	1	4	6
New B	1	7	6
Old	1	3	6

Table 1.2.1: Each instrument measured 6 replicate samples from a single niskin, the rosette position of the niskin is shown. Each bottle was fired sequentially, one after the other as quickly as possible at a depth of 1000 meters.

3.2.1.1. Samples from One Niskin Boxplot

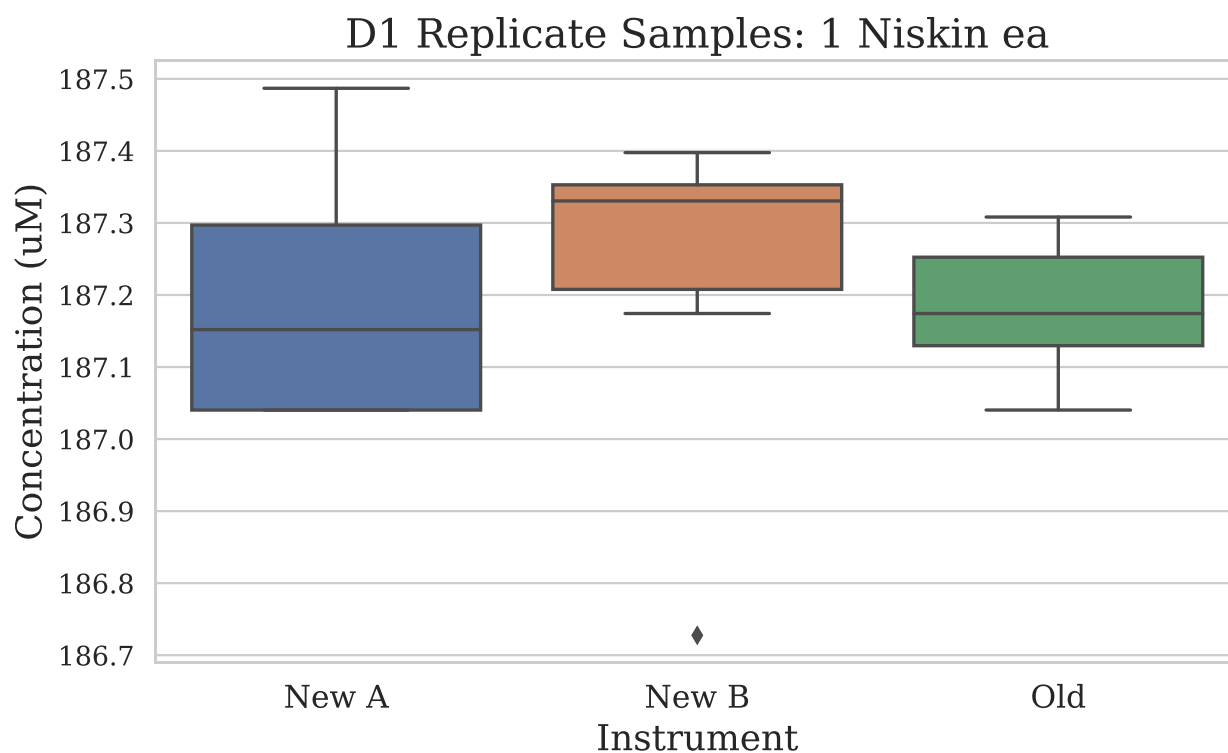


Figure 1.2.1.1: Boxplot style chart displaying the distribution of measurements for the sample replicates from each instrument.

3.2.1.2. Descriptive Statistics

STAT	NEW A	NEW B	OLD
Mean (uM)	187.1966	187.2189	187.1817
Median (uM)	187.1520	187.3306	187.1743
Standard Deviation	0.1868	0.2526	0.0995
% RSD	0.100%	0.135%	0.053%
n	6	6	6

Table 1.2.1.2: The basic descriptive statistics for the deep deployment sample replicates.

3.2.1.3. T-Test Comparison of Means

COMPARISON	P-VALUE	SIGNIFICANT DIFFERENCE
------------	---------	------------------------

New A to Old	0.866	No
New B to Old	0.744	No

3.2.2. Samples from Two Niskins for all Instruments

The samples to test all instruments were collected from two Niskins, this resulted in each instrument having 2 samples from each niskin – as 6 samples were collected from either Niskin.

3.2.2.1. Sample from Shared Niskins Boxplot

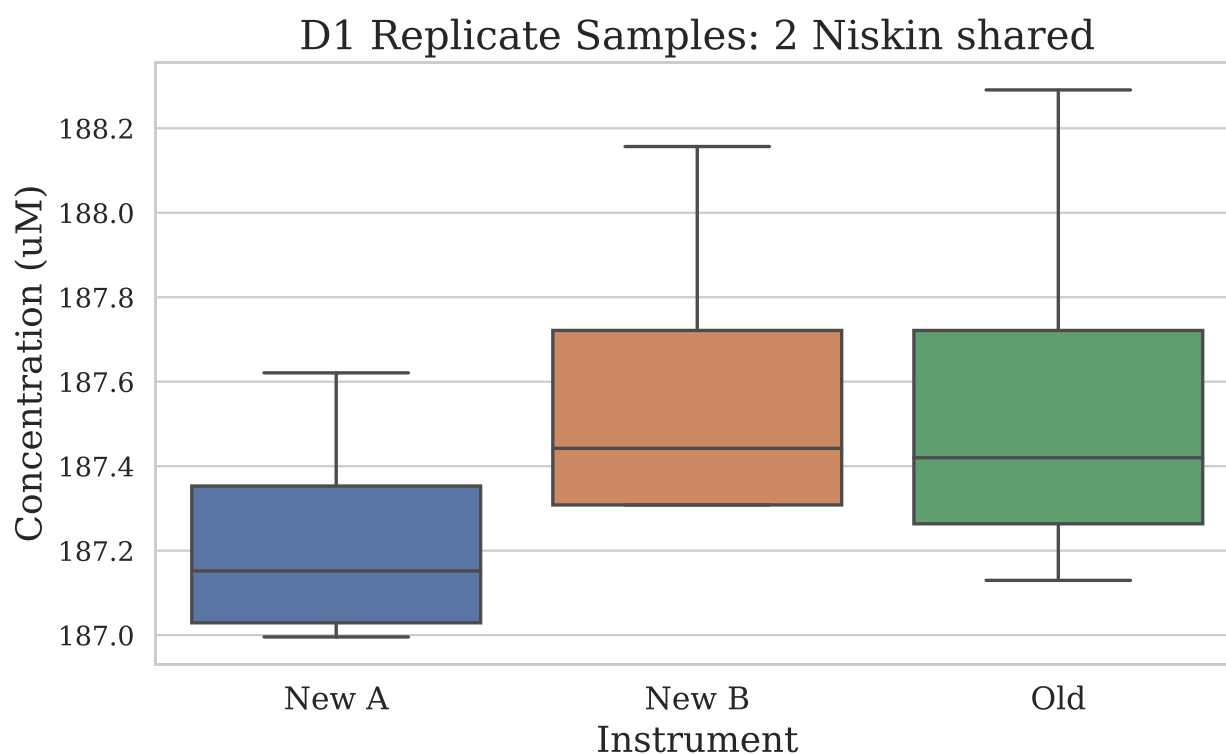


Figure 3.2.2.1: Depicted is a boxplot style chart where the samples measured are from 2 niskins and the replicates split between the 3 instruments. Each instrument made 4 measurements from the replicates pool.

3.2.2.2. Descriptive Statistics

STAT	NEW A	NEW B	OLD
Mean (uM)	187.2301	187.5874	187.5650
Median (uM)	187.152	187.4422	187.4199
Standard Deviation	0.2856	0.4001	0.5110
% RSD	0.153%	0.213%	0.272%
n	4	4	4

Table 3.2.2.2: Basic descriptive statistics of the shared deep water sample replicates.

3.3. Atmospheric Saturated Sample: All Instruments

The results shown in this section, 3.3, were generated from the measurement of samples collected off the saturated oxygen rig. The rig was kept in the Hydrochemistry laboratory, where it equilibrated close to 21.5°C. Previous installation of the pressure monitor in the laboratory was used to get the air pressure at the time of sample collection. Both instances of sample collection, 12 samples were taken.

3.3.1. Atmospheric Saturated Sample Boxplot (auto-scale)

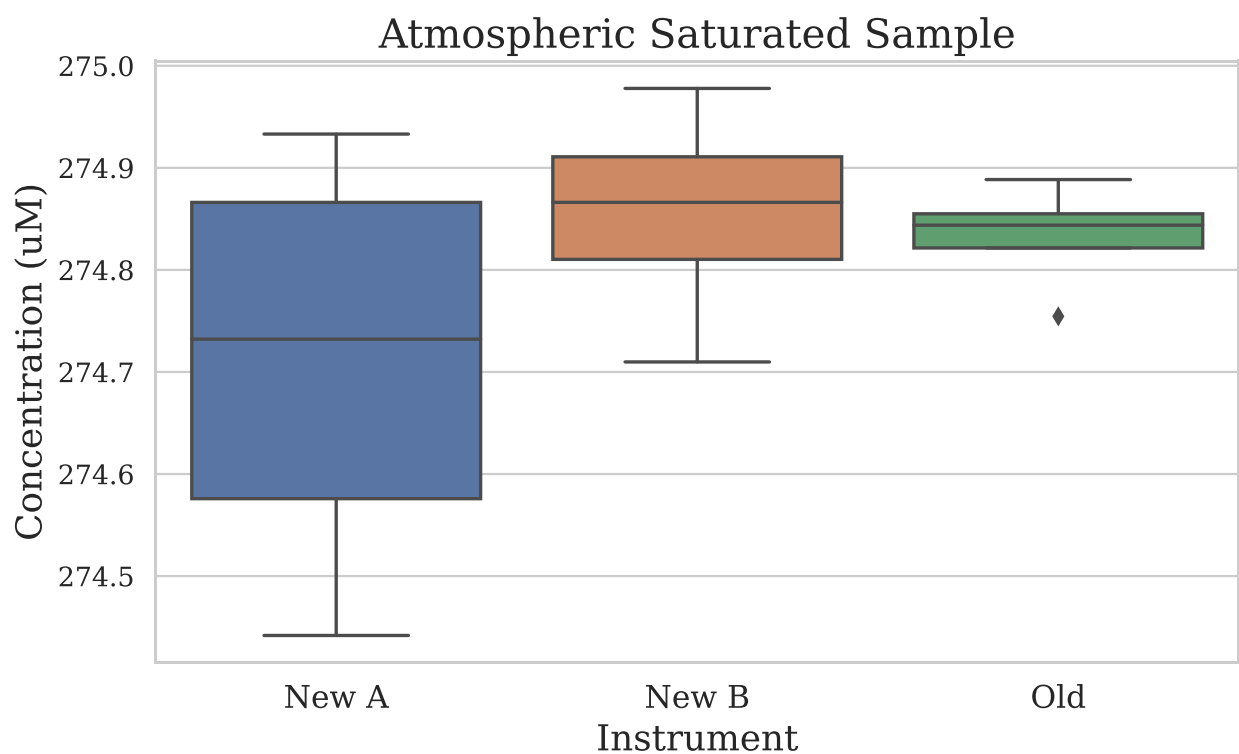


Figure 3.3.1: Boxplot style chart with the measurements made on the oxygen saturated samples created using the new rig. 12 samples in total were collected for this section, with each instrument analysing 4 of these.

3.3.2. Atmospheric Saturated Sample Boxplot (QC Control Lines)

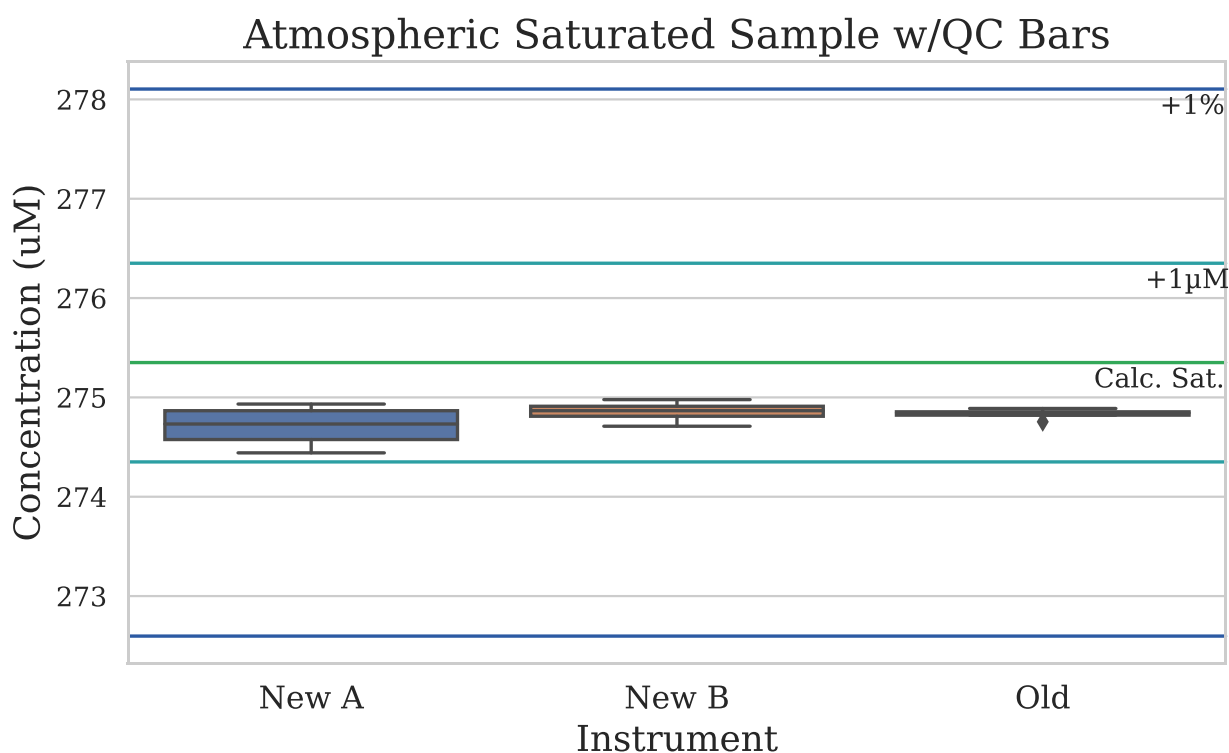


Figure 3.3.2: Boxplot of the same results from figure 3.3.1, however now depicted with quality control lines which show the theoretical oxygen saturation concentration (green), the saturated concentration $\pm 1\mu\text{M}$ (cyan) and the saturated concentration $\pm 1\%$ of the concentration (blue).

3.3.3. Descriptive Statistics

STAT	NEW A	NEW B	OLD
Mean (uM)	274.7098	274.8550	274.8327
Median (uM)	274.7322	274.8662	274.8438
Standard Deviation	0.2218	0.1116	0.0562
% RSD	0.081%	0.041%	0.020%
n	4	4	4

Table 3.3.3: Basic descriptive statistics of the saturated oxygen measurements.

3.4. Atmospheric Saturated Sample: One Instrument

This section includes the results from the repeated measurement of saturated oxygen samples collected off of the rig. Samples were measured on instrument New B, this was 12 measurements.

3.4.1. Atmospheric Saturated Sample: Instrument New B (auto-scale)

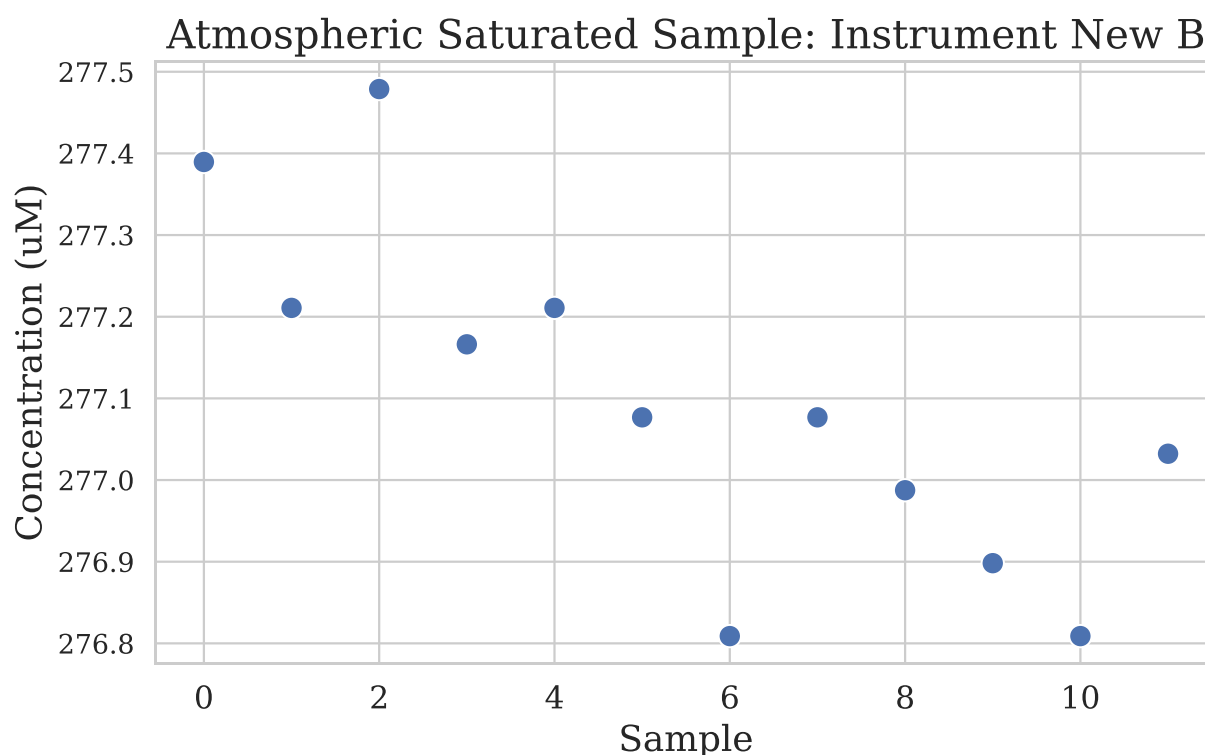


Figure 3.4.1: Scatter plot of the measurements made by instrument New B on a second set of 12 saturated oxygen atmospheric samples. The x axis is sample number in order of collection, y axis is concentration.

3.4.2. Atmospheric Saturated Sample: Instrument New B (auto-scale)

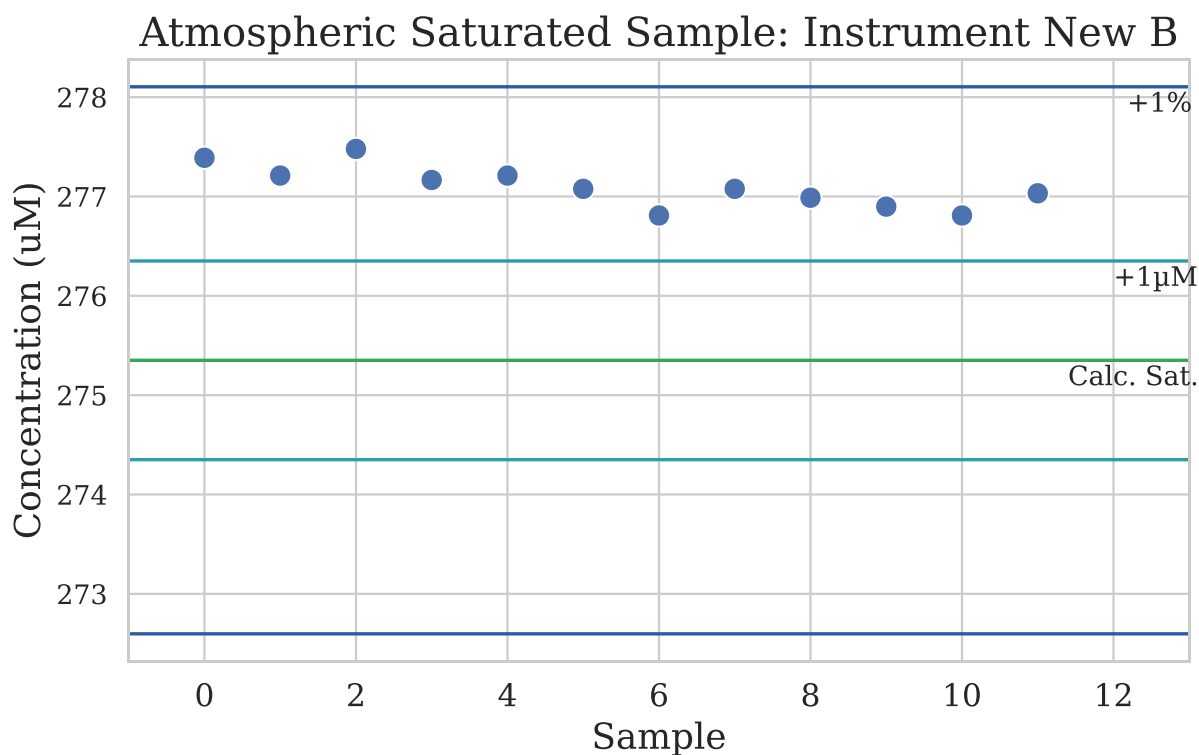


Figure 3.4.2: This scatter plot is like figure 3.4.1, however it includes quality control lines which show the theoretical oxygen saturation concentration (green), the saturated concentration $\pm 1\mu\text{M}$ (cyan) and the saturated concentration $\pm 1\%$ of the concentration (blue).

3.4.3. Atmospheric Saturated Sample: Instrument New B Boxplot (QC Control Limits)

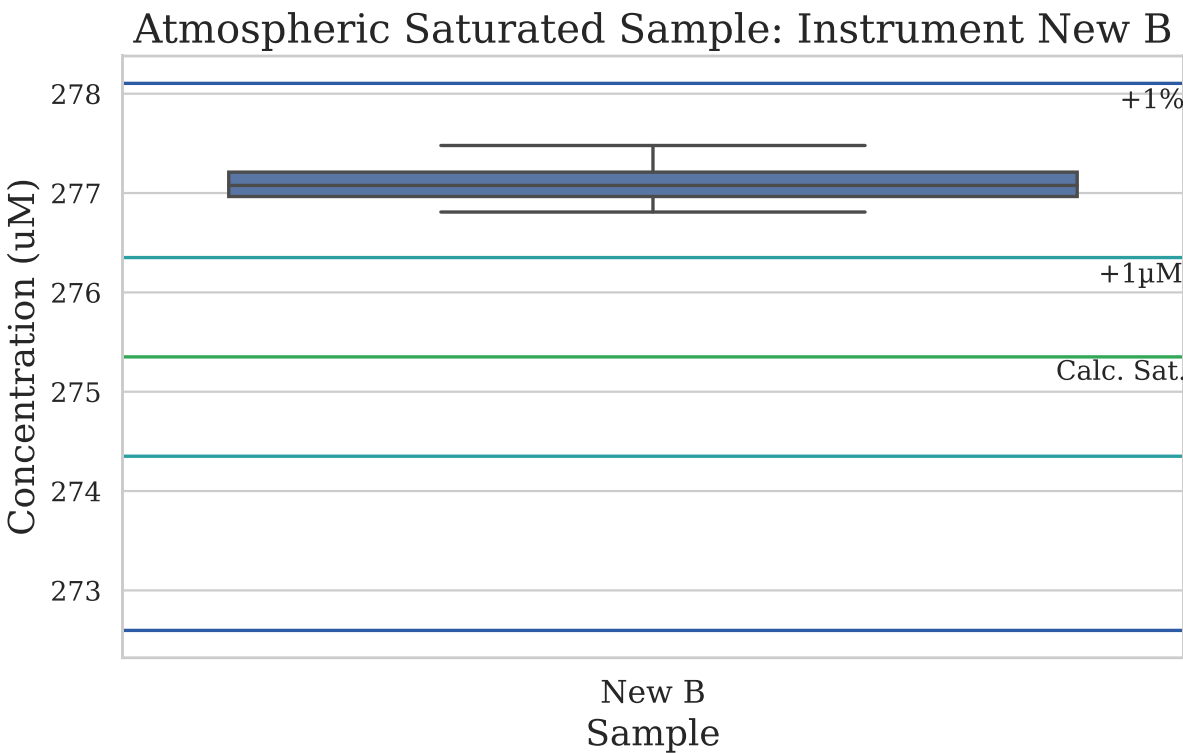


Figure 3.4.3: Very similar to figure 3.4.2, however depicted in a boxplot style to match the analysis completed on the other experiments.

3.4.4. Descriptive Statistics

STAT	NEW B
Mean (uM)	277.09
Median (uM)	277.07
Standard Deviation	0.209
% RSD	0.075%
n	12

Table 3.4.4: The descriptive statistics for the 12 atmospheric samples measured by instrument new B.

3.5. Water Profile Comparison

Duplicate samples were collected from a deployment that was used to go throughout the entire water column. Each duplicate sample was split between two instruments, New A and Old. There was also Niskin replicates as every depth, most depths contained triplicate firings. These samples were collected on deployment 1.

3.5.1. Water Profile Plot

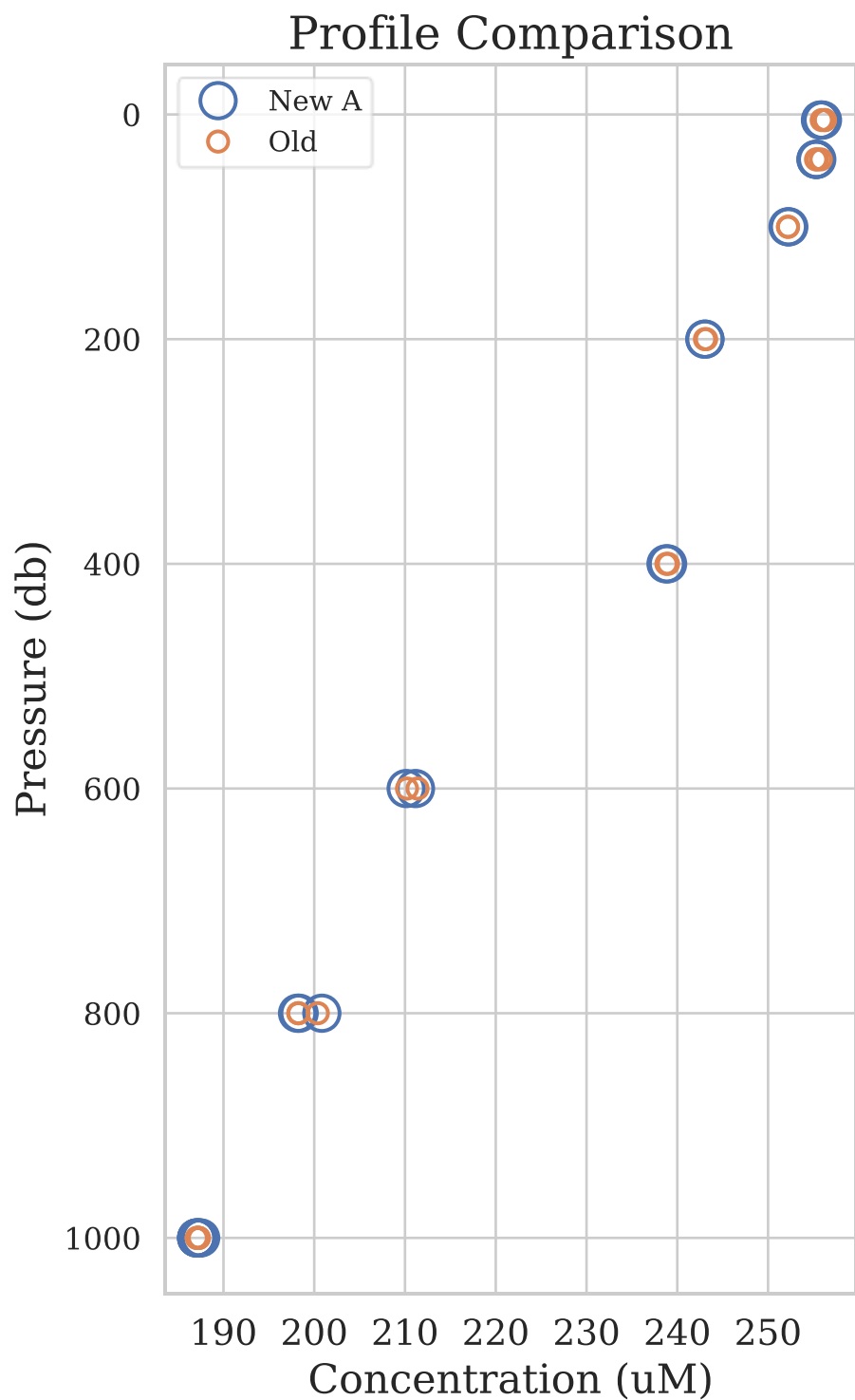


Figure 3.5.1: Profile plot where the y axis is pressure (reversed, 0 at top), x axis is concentration. The different instrument measurements are depicted by the circle colours and size, small orange circles are Old instrument and bigger blue circles are the New A instrument.

3.5.2. T-Test Comparison of Means at specific depths

DEPTH	P-VALUE	SIGNIFICANT DIFFERENCE
5	0.10	No
40	0.30	No
800	0.90	No
1000	0.86	No

Table 3.5.2: T-Tests were completed on the depths where triplicate firing occurred. This test is used to ascertain if there is a difference in the means of the two groups (in this case Instruments). None of the tests produced a significant result for the samples throughout the water column.

3.6. Repeated Deep Sample Measurement: 2

Deployment 2 involved firing all Niskin bottles at the bottom depth of 1000 meters. A dissolved oxygen sample was then taken from each niskin, this resulted in each instrument measuring 8 samples of the total 24 replicates.

3.6.1. Deployment 2 Replicates Boxplot

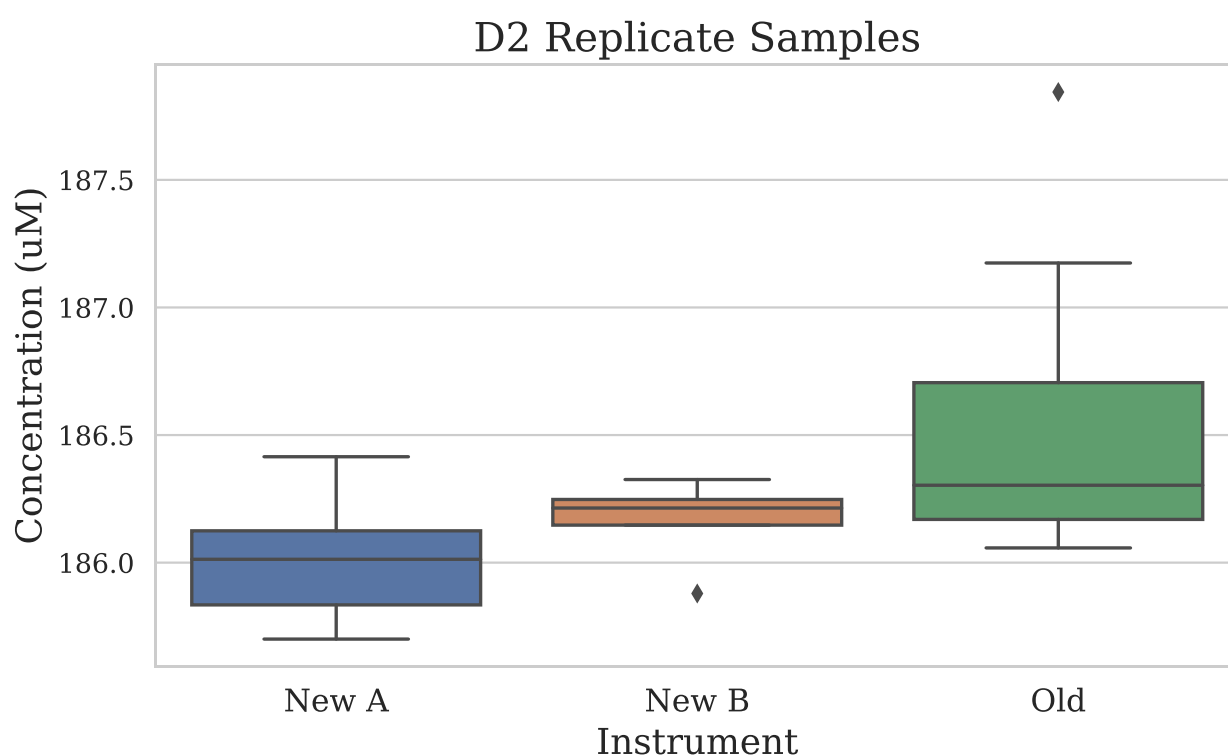


Figure 3.6.1: Boxplot style chart of the measurements made on replicates taken from deployment 2. The replicates were produced by firing every Niskin bottle at 1000 meters, a sample was then taken from each bottle. Each instrument then analysed 8 samples.

3.6.2. Descriptive Statistics

STAT	NEW A	NEW B	OLD
Mean (uM)	186.02	186.18	186.56
Median (uM)	186.01	186.21	186.30
Standard Deviation	0.245	0.136	0.627
% RSD	0.131%	0.073%	0.336%
n	8	8	8

Table 3.6.2: The descriptive statistics of the deployment 2 deep sample replicates measured by each instrument.

3.6.3. T-Test Comparison of Means

TEST COMPARISON	P-VALUE	SIGNIFICANT DIFFERENCE
-----------------	---------	------------------------

New A to Old	0.04	Yes
New B to Old	0.11	No
New A to New B	0.14	No

Table 3.6.2: T-Test comparison of the deployment 2 replicates.

3.7. Experiment-Wide Meta-Analysis

With the collection of so much data throughout many different experiments, there is the opportunity to collate the results and gain additional insight. Each instrument can be assessed from an overarching perspective irrespective of the specific in each experiment. While

3.7.1. Calculation of Mean Measurement Standard Deviation

INSTRUMENT	EXPERIMENT-WIDE MEAN SD (μM)
New A	0.204
New B	0.241
Old	0.278

Table 3.7.1: The mean experiment-wide standard deviation was calculated as an approach to understand how each instrument performed overall.

3.7.2. F-Test Comparison of Variances

The one-way F-Test did not show a significant statistic when comparing the standard deviations of the three instruments across all experiments.

STAT	RESULT	SIGNIFICANCE
F-Value	0.229	
p value	0.803	No

Table 3.7.2: Results from the one-way F-test looking at the standard deviations of all instruments from the different experiments.

4. Discussion

Testing of multiple dissolved oxygen instruments is an extremely tedious process, where attempting to have all 3 instruments operating in parallel takes a significant amount of work. This process, when typically completed with just one instrument is fine – even in the situation where issues arise – it can be managed. However, in this instance with all 3 running it becomes extremely time consuming to troubleshoot and ensure each instrument is operating to its best capability.

Initial setup and preparation for testing took a significant amount of time, mainly due to issues stemming from the new instruments and using computers that were not initially configured. This report will not go into extensive detail about each issue encountered as they were essentially teething problems but below each will be summarised.

- Failing tubing due to both broken thread and flanging
- Bugged software version (2.36e) not working correctly with 876 Dosimat
- Stirrer plate positioning incorrect
- Stirrer bar being pulled into path of light source
- Thermistor reversed plug (old thermistors don't interchange with new instruments)

Independent Iodate Standards

Results from the first experiment, the independent iodate standards, provided the very initial information that the instruments were performing in a very similar specification. While this experiment involved only the measurement of 4 standards, it still showed some level of reproducibility between the 3 instruments.

Deep Sample Replicates #1

Following on from the independent iodate standards a deep sample replicates experiment was undertaken. In this experiment, samples collected from 1000 meters were measured repeatedly. In this first part of this experiment, samples all from one Niskin were measured, this was expected to have the tightest grouping of any “real” sample replicates measured. In this test, there was no statistically significant difference between the measurements made by different instruments – a positive indicator. The variance in measurements was also comparable between instruments, further indicating that the comparability between all 3 is quite good. Additionally, 2 more Niskins had 1000-meter replicates taken to make 12 sample in total. These samples were split evenly among the instruments, so each instrument had 4 samples to measure. Results from this experiment were like the first part, however for some reason instrument New A differed ever so slightly from the other two instruments, but still had a very similar measurement variance.

Atmospheric Sample

Up next to be tested is the saturated oxygen at atmospheric air pressure sample, this was the first time the new rig had been setup and tested. As an aside, the setup seemed to work well and

allowed for the collection of an oxygen sample in the same way as from a Niskin. It will need some additional testing, most likely around the amount of volume that can be taken from the carboy while still providing a homogenous sample set.

Firstly, for the atmospheric sample, there were 12 collected and evenly distributed between all instruments. This meant 4 measurements were made by each instrument, the results well all extremely comparable – well within the 1 micromole window we'd been advised about. When the bars indicated $\pm 1\mu\text{M}$ and 1% are overlaid it becomes obvious how closely the instruments are measuring to one another. The variance from the instrument New A measurements appear large when looking at the auto-scaled chart. However, this is just due to the scale of the y axis. Curiously in this set of measurements the Old instrument posted the tightest group of measurements, with a standard deviation of $0.05\mu\text{M}$. It does appear as though there is more resolution in the output value in the newer instruments. The older instrument seems to “step” ever so slightly more, perhaps due to less than half of the end point *dots* being recorded.

In the second round of testing using the atmospheric sample, 12 were collected and only measured by instrument New B. This was a closer investigation of the precision associated with these samples – it also gave a great view of the reproducibility on the new instrument. In this case the instrument New B had a very tight grouping of the sample measurements. The accuracy of the measurement was a big off, giving a result slightly higher than expected. Nothing had changed on the instrument between the previous atmospheric sample measurements. The only change was that the rig had been refilled and potentially the aquarium air stone was too low below the water level and ended up over-saturating the sample. This will require a small bit of investigation, I read somewhere the over-saturation can occur even when the stone is only 5 centimetres below the water level. Further investigation on the sampled volume needs to be undertaken, throughout the 12 samples taken there was a slight downward trend in the dissolved oxygen concentration. Though, this is perhaps more likely due to the aquarium stone coming out of the water and not continually over saturating the sample.

Profile Comparison

To verify that the new oxygen instruments responded to changing concentrations in the same way as the old a water column of samples was compared. In this experiment duplicate samples were taken out of each Niskin and one of the duplicates was measured by either the Old or New A instrument. There was also depth replicates, where multiple Niskins were fired at the same depth. This allowed T-Tests to be completed as a statistical approach of comparison.

None of the T-Tests found that there was a significant difference between the measurements made on the Old and New A instruments. Statistically speaking the calibration produced by either instrument would be the same.

Deep Sample Replicates #2

For further assessing the precision of the instruments more sample replicates were collected. These sample replicates were taken from deployment 2, where every Niskin bottle was fired at the

cast bottom depth of 1000 meters. One sample was collected from every Niskin and distributed equally to all three instruments.

In this experiment there was no statistically significant difference between instrument New B and the Old instrument. Indicating the two groups of measurements match. Curiously, instrument new A was significantly different to the Old instrument, but this was only true when a high outlier measurement was kept. When this measurement is removed from the set of Old instrument results, there is no significant difference between New A and Old.

Experiment-Wide Meta-Analysis

By using all the data collected and the relevant groupings, further analysis and understanding can be gleaned about instrument performance. This approach does not look at the specific results from each experiment, but rather takes a holistic methodology to assess the overall precision of an instrument. Perhaps the most promising result from the entire investigation, a statistical F-Test found there to be no significant difference in the precision of each instrument. This is a strong indicator that there is no difference in precision of the instruments.

5. Appendix

5.1. CTD deployment locations

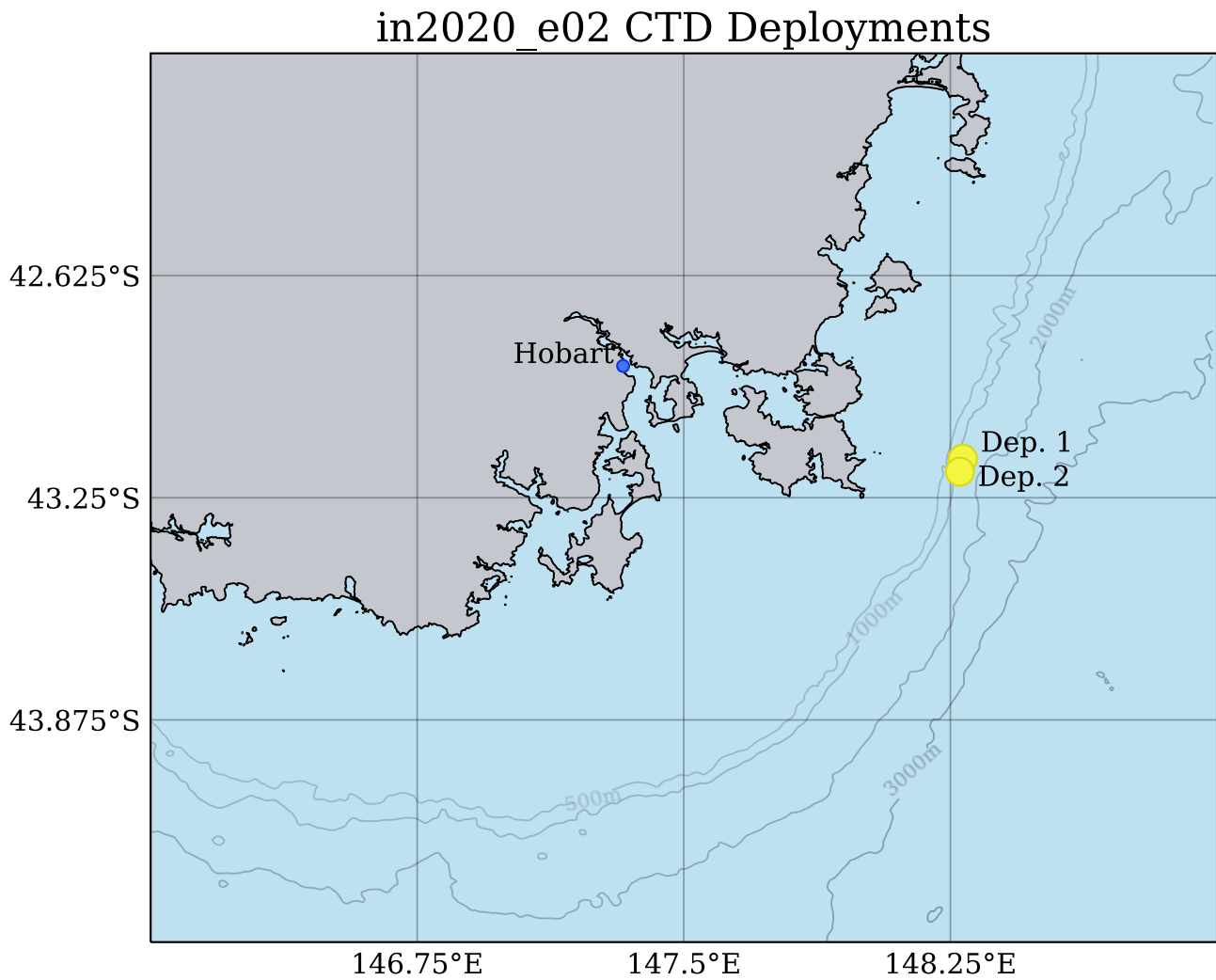


Figure 5.1: Shown are the CTD deployment locations used for the dissolved oxygen instrument experiments on in2020_e01.