Supplement of Ocean Sci., 22, 1609–1623, 2026
https://doi.org/10.5194/os-22-1609-2026-supplement

Ocean Science Open Access EGU

Supplement of

Obtaining accurate, high-frequency and long-term seawater pH data by using coupled lab-on-chip and optode sensing technologies

Anthony J. Lucio et al.

Correspondence to: Anthony J. Lucio (anthony.lucio@noc.ac.uk)

S1. Chemical and physical drivers

The chemical (salinity and dissolved oxygen) and physical (tidal height and temperature) drivers of seawater pH for this 6-month study are captured below in Figure S1.

The tidal height data is provided in Figure S1a and varies from ca. 2 m to 6 m. There are approximately two peaks and two valleys within any given day and in any given month during the study period, which corresponds to semidiurnal tides. The two data gaps visible within phase 1 were unfortunately a result of spent batteries.

The seawater temperature (Figure S1b) was relatively constant around 20 °C during the initial three summer months (phase 1) but showed a relatively linear decrease towards 8 °C through autumn to winter (phase 2).

The salinity (Figure S1c) was largely around 30 psu for the first four months but showed more variability in the final two months. A small volume of each of the discrete co-samples collected from the test site were measured with a commercial conductivity meter (WTW Cond 3110) and probe (WTW TetraCon 325) to provide independent verification of the salinity, which are overlaid in Figure S1c (yellow filled circles). We suspect that the data gaps during phase 1 present within the salinity data are a result of stagnant seawater within the conductivity flow cell of the CTD section of the SeapHOx sensor, which has been reported previously (Bresnahan et al., 2021). The external pump was activated during each measurement for 5 seconds, yet due to the high biofouling environment we discovered several instances of seawater not fully turning over until after significant time delays. While data was still being collected it was not representative of the environment, and this resulted in several large portions of data not being accurate. As a result of this, the salinity and temperature data are provided by the SeapHOx CTD during phase 1 and then from an independent C-T sensor (microCAT; Sea-Bird Scientific) during phase 2.

The dissolved oxygen (DO) data shown in Figure S1d shows that levels remained within ca. 4 mL/L to 6.5 mL/L during the 6-month trial, and we encountered two periods of oxygen supersaturation within phase 1.

Finally, the seawater pH is shown in Figure S1e from the LOC sensor along with lab validated co-samples overlaid as filled yellow circles. The LOC data agree with previous results from the same deployment site (Yin et al., 2021).

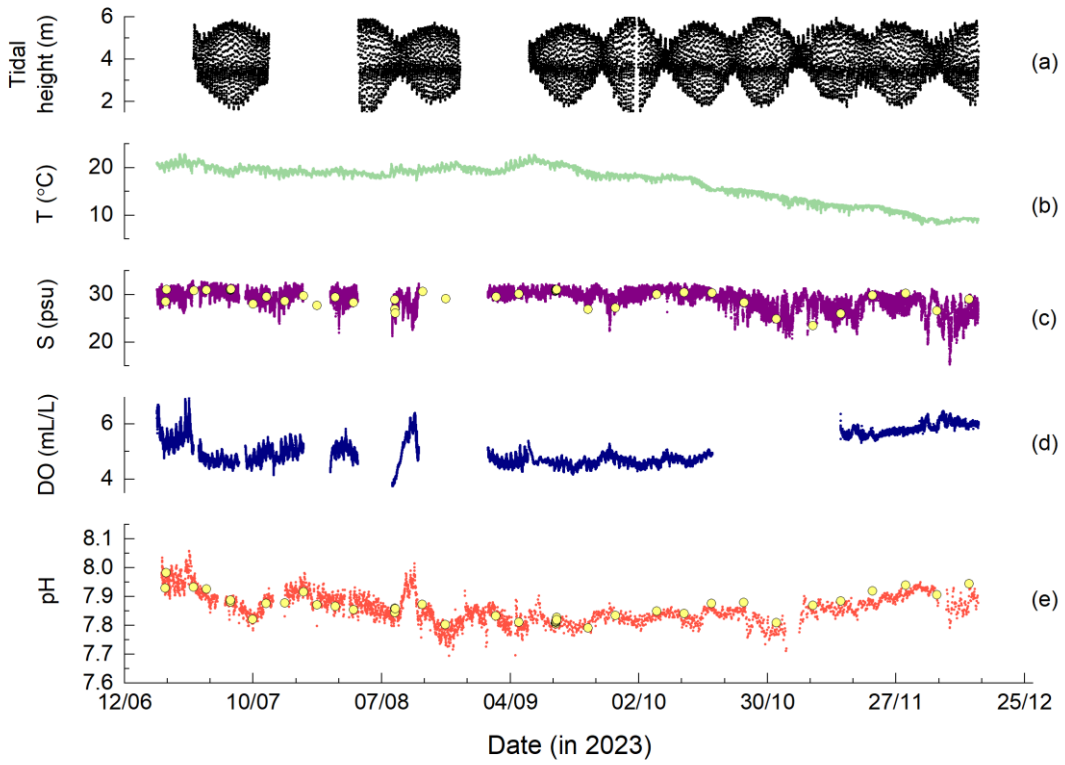
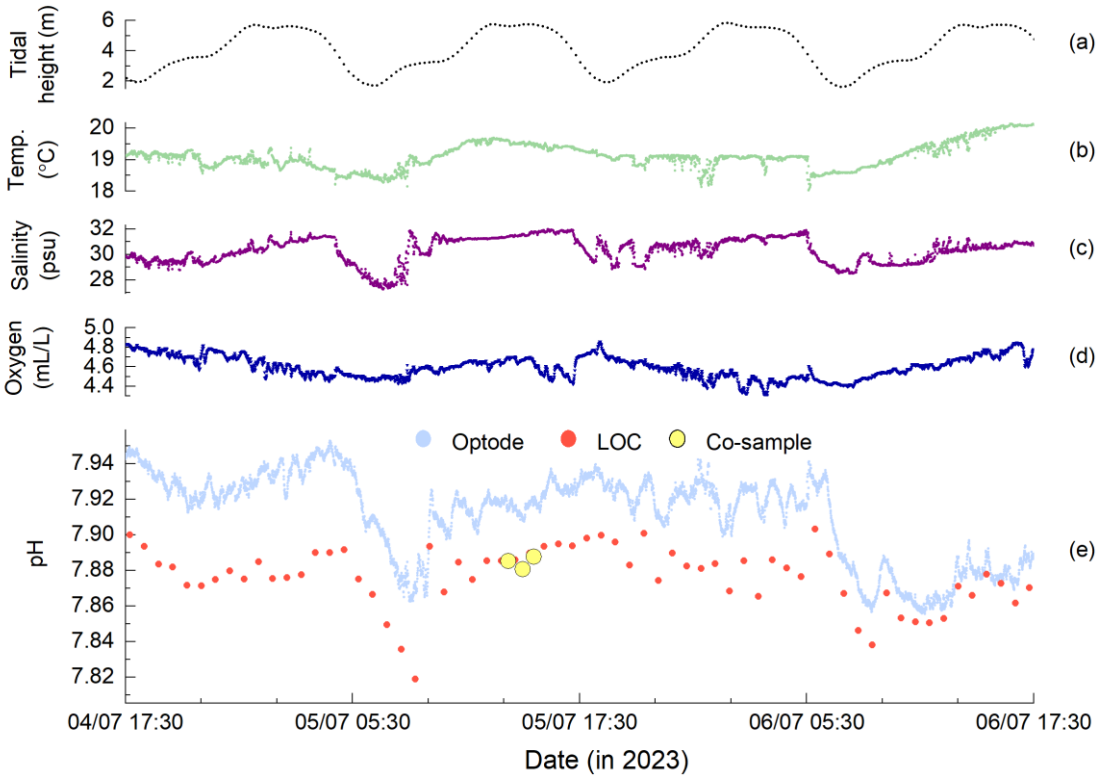


Figure S1. Overlay of chemical and physical drivers of seawater pH captured during 6-month pH trials test: (a) depth, (b) temperature, (c) salinity, (d) dissolved oxygen, and (e) LOC pH. The filled yellow circles in (c) are discrete co-sample salinity measurements done with a lab-based independent conductivity probe/meter and the filled yellow circles in (e) are discrete co-samples measured in the laboratory.

A representative example in Figure S2, shows a zoomed-in portion of Figure S1. As can be seen, in this instance, two tidal cycles occurring in the evening hours drove significant changes in the salinity and this in turn was reflected in the resulting pH measured. For reference, the optode pH data is also overlaid in Figure S2 to further illustrate this point. The temperature and dissolved oxygen signals were not as affected in this instance.



35 **Figure S2. A zoomed-in section of Figure S1showing 48 hours of data.**

S2. Zoom-in on optode drift during phase 1

A zoom-in on the pH data from Figure 2 (main text), provided in Figure S3 below, shows that the optode pH signal is following the LOC pH signal during a period of oxygen supersaturation (ca. August 12 to 16th). At this point the optode signal has drifted to an offset between 0.2 and 0.3 relative to the LOC pH data.

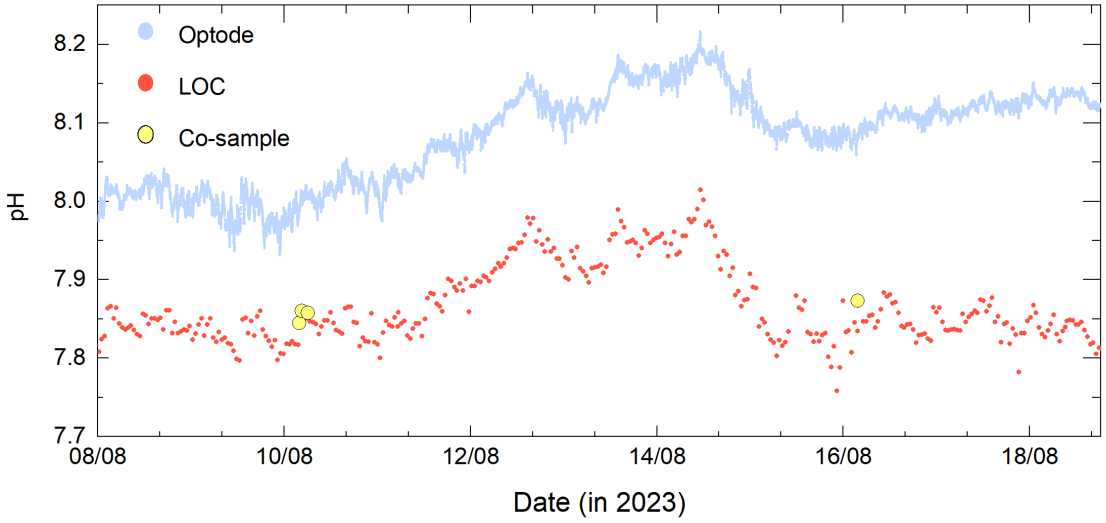


Figure S3. Zoom-in showing the LOC and optode signals during mid-August 2023. Discrete co-samples are also provided (filled yellow circles).

Furthermore, a second zoom-in, provided in Figure S4 below, shows that even during the period of significant signal drift in early-September 2023, where the optode is now offset between 0.6 and 0.8 relative to LOC pH data, the optode signal still follows the trends of that from the LOC sensor. This is an interesting finding given that at, this point, >200,000 measurements have been performed with this optode pH cap and there is significant bioaccumulation on the sensors. While the optode data is severely drifting away from the discrete co-sample data, we demonstrate that it can be corrected using the less frequent yet accurate LOC pH sensor data. It is worth noting that previous results utilising an optode-based pH sensor observed similar periodic fluctuations in pH data within a pier in Southampton waters (Staudinger et al., 2018). Relative to the LOC data, during this period in Figure S4 we estimate a peak drift rate of nearly 0.040/day for the optode-based sensor near the end of phase 1. As mentioned previously, we intentionally let the sensor continue to drift towards unrealistic pH values during this period to access its performance. Prior to mid-August, when the significant signal drift commenced, the drift rate was ≤0.012/day.

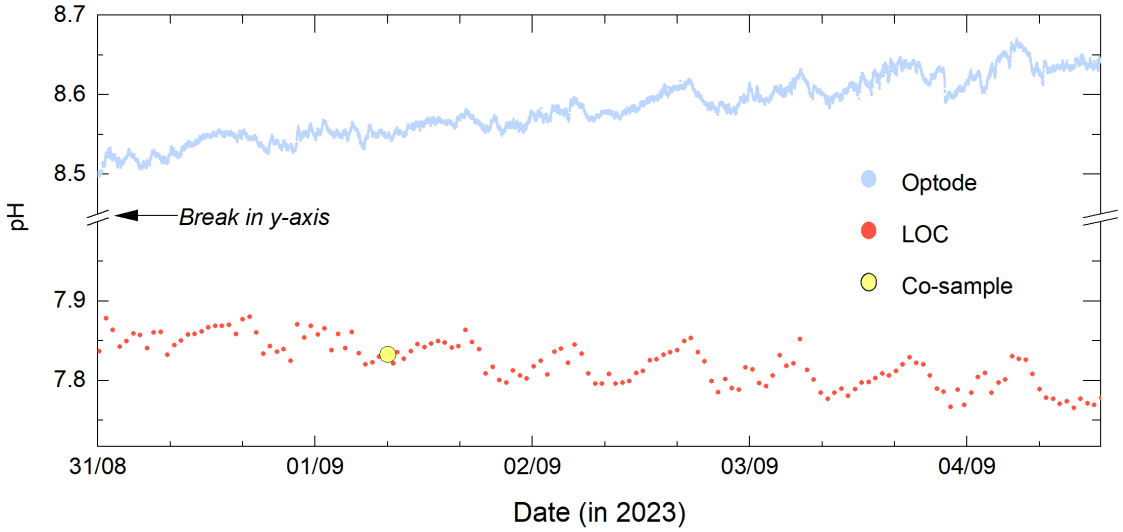


Figure S4. Zoom-in showing the LOC and optode signals during early-September 2023. Discrete co-samples are also provided (yellow filled circles). Note: labelled there is a break in the y-axis from pH 8.00 to 8.45 for visual clarity.

60

S3. Further details on the fitting method

Examination of different data correction methods

We have examined four different fitting methods, and exemplar results are highlighted in Table S1 with respect to a 24-hr correction frequency and Table S2 shows a more frequent 6-hr correction. Specifically, we have compared a multi-point moving average, multi-point PCHIP, a multi-point line fit, a y-axis (pH) offset correction to our current method that is a 2-point line fit, and an enhanced offset correction that also fits to long-term residual drift (i.e., over 3 months). Performance was evaluated using the mean bias, standard deviation and root mean square error (RMSE) of the differences between corrected optode measurements and the discrete co-samples measured in the laboratory. While some alternative approaches produced comparable results, none consistently improved both $\bar{x}$ and 1σ metrics simultaneously. As can be seen in Table S1, the current method (2-point line fit) used in this work yields the lowest RMSE. The multi-point moving average results are not significantly worse, but including more data points in the fitting method requires more LOC sampling (i.e., more reagent consumption, power consumption and waste generation), which will be dictated by the deployment conditions as to whether this is feasible. In particular, more complex interpolation methods tended to increase systematic bias or introduce signs of overfitting. Therefore, this demonstrates that the 2-point line fit used within this study can produce sufficiently corrected optode data while minimising the LOC sampling requirements.

Table S1. Indicates the correction method type and the resulting performance from the whole dataset, relative to discrete lab-validated co-samples. Reported are the mean sensor error ($\bar{x}$ ΔpH), standard deviation of the error (1σ ±ΔpH), and the root mean square error (RMSE ΔpH). The number of samples for determining these metrics were n = 44. Tabulated data are representative of a 24-hr correction interval.

Correction type	$\bar{x}$ ΔpH	1σ ±ΔpH	RMSE ΔpH
2-point line (current method)	-0.011	0.017	0.020
Moving average 3-points	-0.011	0.018	0.021
Moving average 8-points	-0.012	0.019	0.022
Offset	-0.006	0.025	0.026
Long-term residual + offset	-0.008	0.025	0.026

To probe this one step further we have examined fitting methods akin to collecting LOC data points once every six hours to see how more frequent LOC data collection impacts the results. We have kept this comparison to be done within a 24-hr window, and the results are highlighted in Table S2. Here we can see that a multi-point line fit using four data points within a 24-hour window (e.g., a line fit with data points at 0 hr, 8 hr, 16 hr and 24 hr) produces the best fit when compared to the other methods. Nonetheless, the resulting sensor performance, as estimated by the RMSE, is effectively as good as the less frequent 2-point line fit result from Table S1. Therefore, the same result can be achieved by using less frequent in-situ LOC sampling to correct high-frequency optode data and is the chosen correction method for this work.

Table S2. Indicates the correction method type and the resulting performance from the whole dataset, relative to discrete lab-validated co-samples. Reported are the mean sensor error ($\bar{x}$ ΔpH), standard deviation of the error (1σ ±ΔpH), and the root mean square error (RMSE ΔpH). The number of samples for determining these metrics were n = 44. Tabulated data are representative of a 6-hr correction interval.

Correction type	$\bar{x}$ ΔpH	1σ ±ΔpH	RMSE ΔpH
Multi 4-point line	-0.013	0.017	0.021
2-point line (current method)	-0.018	0.019	0.026
PCHIP 4-point	-0.016	0.020	0.026
Offset	-0.018	0.021	0.028

Effect of linear fit starting point

In the fitting protocol outlined within this work, the very first LOC data point collected on 20/06/2023 at 14:20:54 is used as the starting point after which all other fitting intervals are defined (e.g., 6-hr, 24-hr, 168-hr, etc. from this starting point). Anchoring to this starting point could bias the data fitting, and to examine this we looked at the impact this would have on the results. The performance is computed by finding the nearest datetime-matched corrected optode data point to the 44 triplicate discrete co-sample data points and calculating the mean error ($\bar{x}$), the standard deviation of the error (1σ), and the root mean square error (RMSE). The mean error (i.e., $\bar{x}$) represents systematic bias (offset) relative to the reference measurements, while the standard deviation (1σ) reflects the spread of residual differences and therefore the measurement precision. The RMSE incorporates both the systematic bias and the random variability and therefore provides a combined estimate of total measurement error. This is shown below in Figure S5 for the 24-hr correction interval by using the first 200 LOC data points as the starting point; for reference the 200th LOC data point occurred on 24/06/2023 at 22:31:00. We can see that the performance (relative to discrete co-samples) shows little dependence on the starting index point (the average standard deviation is ±0.017), and all starting indices produce a difference within a ±0.05 threshold. Therefore, for consistency the very first LOC data point has been used in all subsequent data fitting examples.

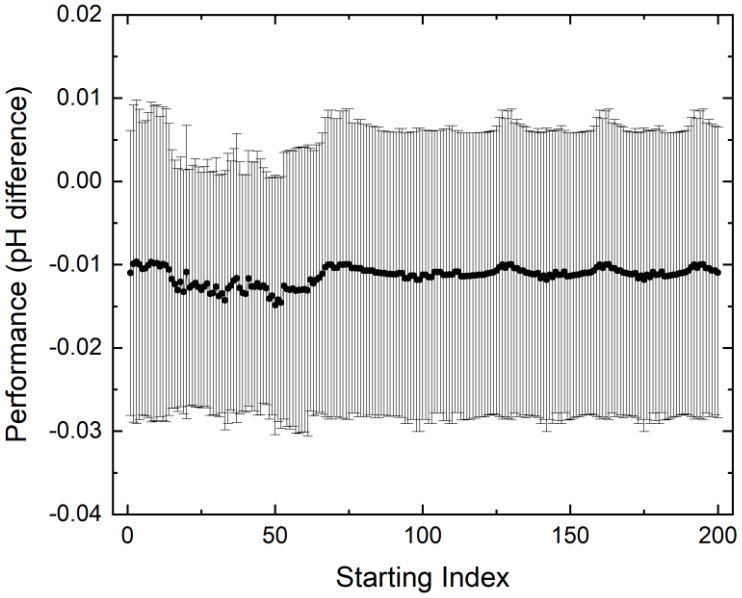


Figure S5. Exemplar plot of the sensor performance as a function of the LOC data point starting index. The symbols represent the $\bar{x}$ ΔpH and the error bars are the 1σ ±ΔpH. A 24-hr correction interval was modelled here.

S4. pH sensor comparison

This deployment included two spectrophotometric LOC pH sensors (NOC), each with an optode-based pH sensor (PyroScience GmbH) attached to it and were both deployed side-by-side. For brevity, we chose to show data only from one of each of these sensors within the main text (i.e., denoted as LOC1 and optode-1 here within the supplementary information sections) and have included the second combined dataset results here (within sections S4 and S5) for clarity.

LOC1-LOC2

The data from LOC1 has been used throughout the main text but the below discussion demonstrates that either sensor could be used. The performance of the LOC sensors relative to one another is highlighted in Figure S6. The pH data shows good agreement over the entire 6-month deployment (bottom panel). Within the first month of phase 2 (i.e., from 21/09/2023 to 09/10/2023) the LOC pH sensors showed the longest sustained difference i.e., absolute difference of 0.034 ± 0.013 (216 data points in common), yet the mean and standard deviation was 0.007 ± 0.020 for the entire data set (3,108 data points in common). This is also shown via a difference (ΔpH) plot (top panel), where the grey solid line shows the mean error (difference), and the grey dashed lines represent the standard deviation of the error. With respect to the SeapHOx and AquapHOx-L-pH sensors that have manufacturer reported performance of ±0.050, most of the ΔpH data points fall within this range (i.e., only 103 data points or ca. 3% are beyond this threshold). The below data demonstrates that, when deployed in the field, the spectrophotometric LOC pH sensors return reproducible data from sensor to sensor. These pH sensors have been commercialised and have since been deployed around the globe.

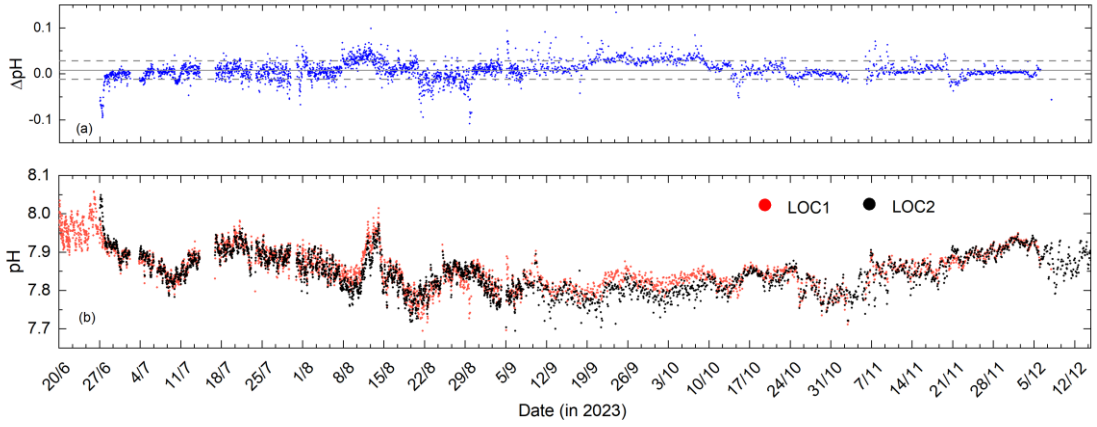


Figure S6. LOC pH sensor performance relative to one another. Top panel (a) shows a difference plot between the two LOC pH sensor datasets. The grey solid line represents the mean error, and the grey dashed lines represent the standard deviation of the error. Bottom panel (b) shows both LOC pH sensor datasets overlaid for the 6-month deployment period.

SeapHOx-LOC1 comparison

The LOC1 pH sensor performance relative to the SeapHOx pH sensor is shown in Figure S7 for the 6-month dataset. As mentioned in the main text, there are a few gaps without data from the SeapHOx pH sensor that was likely a result of stagnant seawater not flushing properly within the pumped section of the sensor (during phase 1) due to significant biofouling and secondly due to spent batteries (during phase 2). Nonetheless, the two datasets are in good agreement, and we find a difference of -0.022 ± 0.023 for the entire dataset (3,182 data points in common). This is also shown via a difference plot (top panel), where the grey solid line shows the mean error (difference), and the grey dashed lines represent the standard deviation of the error. Similar to the above LOC1-LOC2 comparison, most of the ΔpH data points fall within a ±0.050 threshold (i.e., only 247 data points or ca. 8% are beyond this threshold). The below comparison provides independent verification of the LOC pH sensor performance.

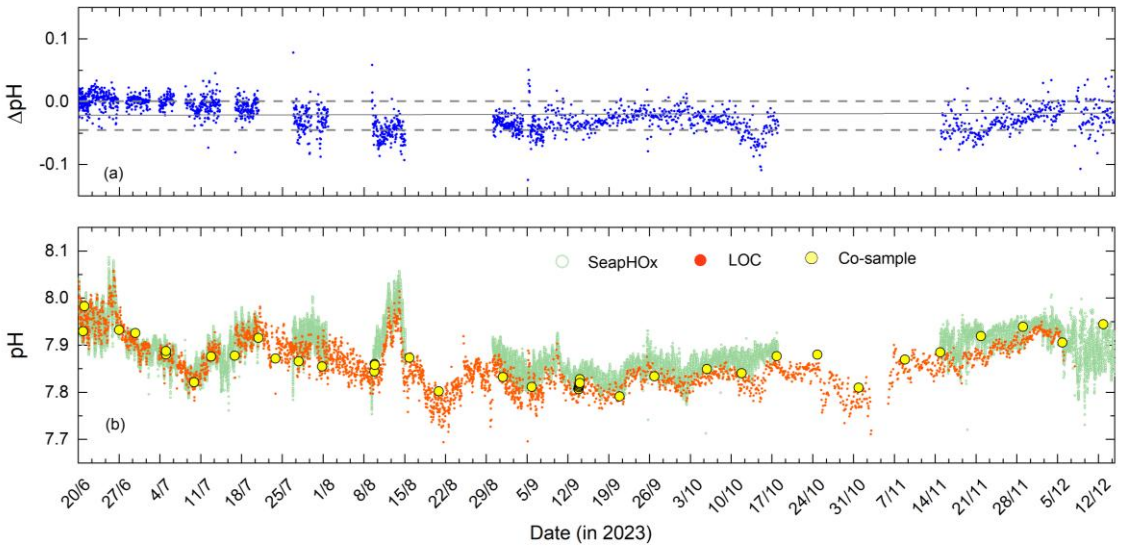


Figure S7. SeapHOx-LOC1 performance relative to one another. Top panel (a) shows a difference plot between the two datasets. The grey solid line represents the mean error, and the grey dashed lines represent the standard deviation of the error. Bottom panel (b) shows both SeapHOx and LOC pH datasets overlaid for the 6-month deployment period, in addition to lab-validated co-samples.

Overall, we find that the performance of the SeapHOx pH sensor relative to validation samples to be well within its reported specification (i.e., ±0.050) from the manufacturer (refer to Table S3). The two LOC pH sensors are also within this range and in agreement with prior work reporting a difference of +0.003 ± 0.022 (n = 47) at this same test site (Yin et al., 2021). The LOC pH sensors that are based on a spectrophotometric method achieved comparable performance as the widely utilised SeaFET/SeapHOx sensors that are based on an electrochemical measurement method. Therefore, the LOC pH data can be used to correct other sensor data that may exhibit signal drifts/offsets.

Table S3. Sensor performance ($\bar{x}$ΔpH mean error, 1σ ΔpH standard deviation of error, and RMSE ΔpH) for the SeapHOx and LOC sensors used in this study. The performance is reported relative to the discrete lab-validated co-sample data. Note: n is the number of samples.

Sensor	$\bar{x}$ ΔpH	1σ ±ΔpH	RMSE ΔpH	n
SeapHOx	+0.014	0.015	0.021	38
LOC1	-0.015	0.014	0.021	28
LOC2	-0.022	0.018	0.028	29

S5. Optode-2 performance

As previously mentioned, during phase 2 a second optode pH sensor (denoted optode-2 within the supplementary section) was deployed on 19/09/2023 alongside the already deployed second LOC pH sensor (LOC2). Optode-2 was an AquapHOx-L-pH (PyroScience GmbH) sensor identical to the first one deployed during phase one. Optode-2 was, however, equipped with an anti-fouling sensor protection cage and an anti-fouling sensor protection cap that are available from the manufacturer (PyroScience GmbH). These two copper-alloy mesh cages prevented biofouling on the sensing portion of the pH sensor. Example pictures (Figure S8) after ca. three weeks of use at the testing site show the ability of the anti-fouling copper-alloy mesh to minimise bioaccumulation on the sensing portion of the optode. The optode manufacturer suggests that the anti-fouling protected optode sensors experience less signal drift over time compared to the analogous optode sensor without anti-fouling protection (accessed 03/11/2025: https://www.pyroscience.com/en/applications/white-papers/anti-biofouling).

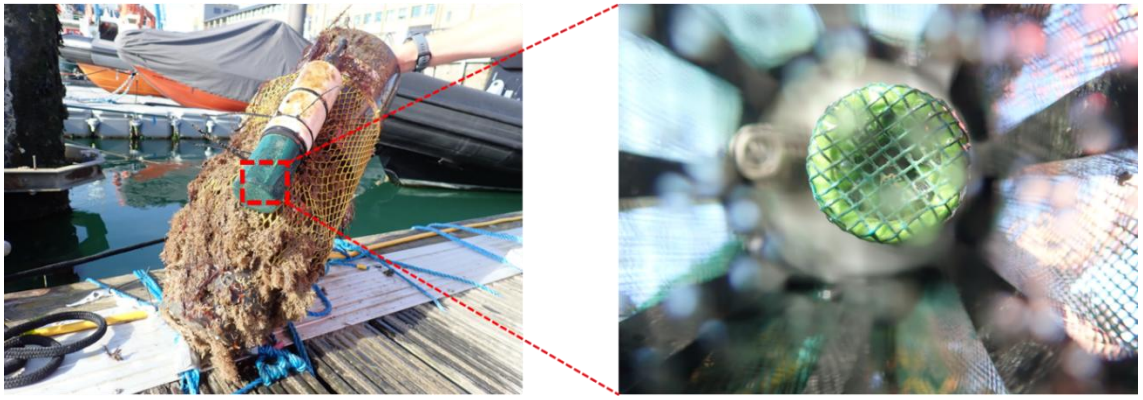


Figure S8. Exemplar pictures of optode-2 anti-fouling sensor protection cage (left) and a zoom inside showing the anti-fouling protection cap (right). Optode-2 was attached to LOC2 pH sensor. Pictures taken on 09/10/2023.

The raw response of optode-2 is shown overlaid in Figure S9, alongside the corresponding raw optode-1, LOC2 and discrete co-sample data. As can be seen, after four days optode-2 quickly drifted to a relatively constant offset of +0.124 ± 0.024 (1,890 data points in common) relative to LOC2 data (for reference it is a smaller offset of +0.091 ± 0.024 (n = 19) relative to co-sample data) over the course of phase 2. For the month period between 23/09/2023 and 23/10/2023 the two optode pH sensors were in reasonable agreement with an absolute difference of 0.019 ± 0.020 (7,805 data points in common) and both exhibited a positive pH offset relative to discrete co-samples. For two weeks between 07/11/2023 and 20/11/2023 the optode pH sensors showed the longest sustained difference i.e., absolute difference of 0.125 ± 0.036 (3,615 data points in common), wherein the optode-1 signal fell in line with the LOC2 data (i.e., a difference of 0.008 ± 0.040 with 302 data points in common) and the optode-2 signal maintained its relatively constant offset of +0.129 ± 0.018 (870 data points in common) relative to LOC2 data for the remainder of the deployment. We are unsure why the offset change happened to optode-1, however, thereafter optode-1 returned to exhibit a positive pH offset relative to LOC2 data (+0.165 ± 0.055 with 597 data points in common) for the remainder of the deployment. This level of pH offset (+0.1 to +0.4) has been observed previously with optode pH sensors

(Delaigue et al., 2025; Wirth et al., 2024). Nonetheless, as we have demonstrated in the main text, the linear rate correction can be used to improve the drifting/offset pH optode data.

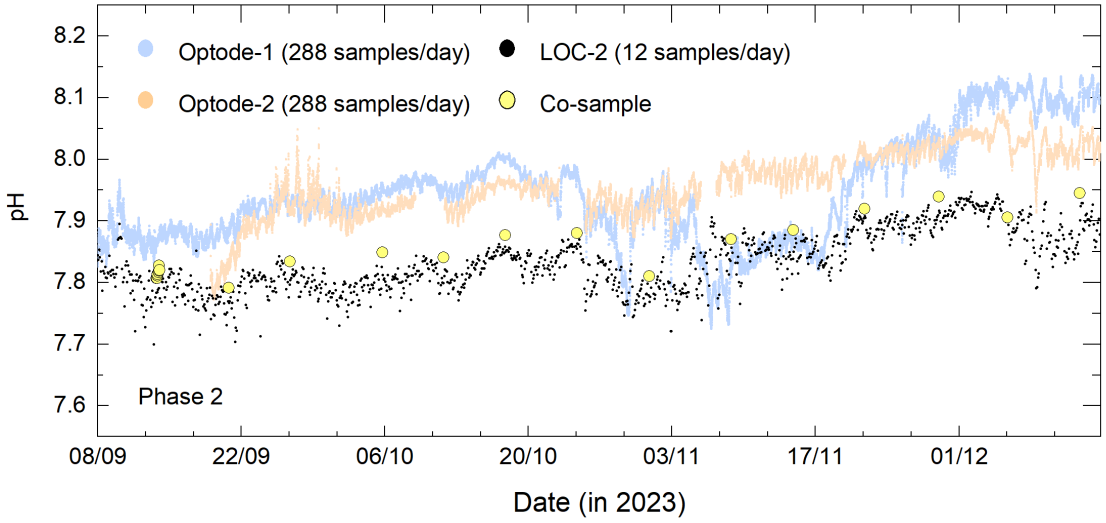


Figure S9. Overlay of optode-1 (light blue filled circles), optode-2 (light orange filled circles), LOC2 (black filled circles), and discrete lab validated co-sample (yellow filled circles) pH data during phase 2.

An overlay of the 24-hr corrected optode-2 data is provided in Figure S10a, alongside SeapHOx, LOC2 and discrete co-sample data. The performance of the optode-2 sensor throughout phase 2 is -0.039 ± 0.021 (n = 11) relative to discrete co-sample data. For reference, the same for optode-1 during phase 2 is -0.015 ± 0.014 (n = 13) relative to discrete co-sample data. The decrease in performance with optode-2 is likely a result of the parent LOC data used to undertake the corrections to the raw optode data. For example, when utilising optode-2 data that has been corrected using LOC1 data, we find its phase 2 performance improves to -0.016 ± 0.012 (n = 11) relative to discrete co-sample data, which is in much better agreement with the accuracy of optode-1 during phase 2 (also corrected using LOC1 data).

Zoomed-in plots below show exemplar 24-hour periods. Figure S10b shows that, at the beginning of the deployment approaching the end of September, optode-2 exhibited an offset of -0.044 ± 0.012 (288 data points in common) difference relative to the SeapHOx data. In early-October this offset increased to -0.061 ± 0.011 (288 data points in common) in Figure S10c and increased further around mid-November in Figure S10d to -0.086 ± 0.014 (271 data points in common). Towards the end of the study in early-December it reduced to -0.019 ± 0.017 (288 data points in common) difference (Figure S10e). Although the optode-2 sensor is offset relative to the SeapHOx sensor, it captured high-frequency pH fluctuations that are observed in the SeapHOx sensor, in addition to the broader trends from the LOC2 pH sensor sampling at a much lower frequency (i.e., a measurement every 2-hr).

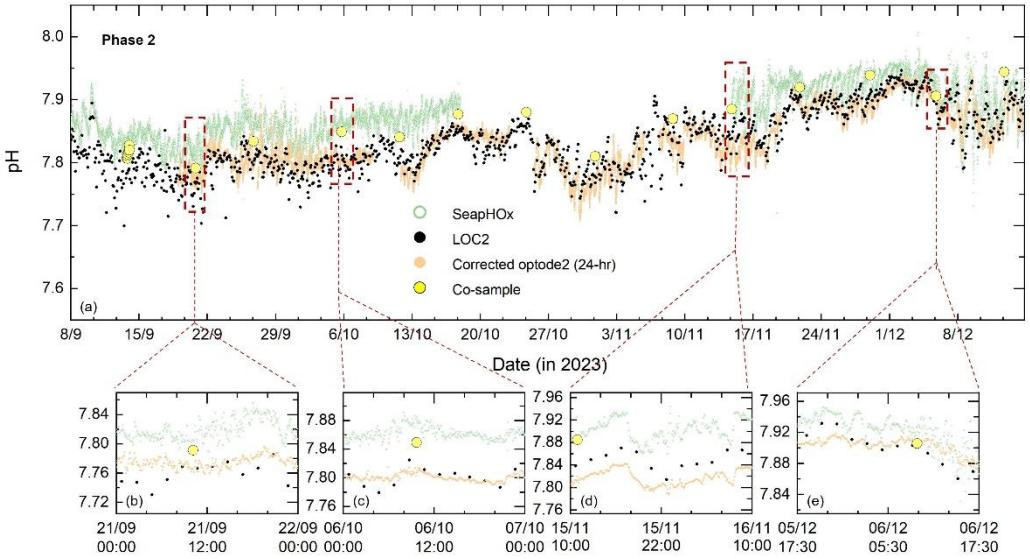


Figure S10. Top panel (a) shows the entire 3-month optode-2 series with 24-hr corrected (light orange filled circles) data with both SeapHOx (sage green unfilled circles) and LOC2 pH data (black filled circles) overlaid for comparison. Lab validated co-samples are provided for reference (yellow filled circles). Bottom panels (b) through (e) show four zoomed in snapshots within the larger overlaid dataset.

As previously mentioned, two optode pH sensors were deployed in this study. Optode-1 was deployed for six-months whereas optode-2 was deployed for the final 3-months (phase 2 only). Again, these were deployed on LOC1 and LOC2 pH sensors, respectively and during phase 2 both optodes were measuring at a frequency of every 5 minutes. The bottom panel in Figure S11b shows the corrected optode datasets (at a 24-hr correction interval) that has been corrected using its respective LOC pH sensor data. As can be seen, there are periods when the sensors are tracking each other well and then there are periods when they diverge. The mean and standard deviation between these two datasets was 0.015 ± 0.026 (22,621 data points in common). This is also shown via a difference (ΔpH) plot (top panel, Figure S11a), where the grey solid line shows the mean error (difference), and the grey dashed lines represent the standard deviation of the error. With respect to the SeapHOx and AquapHOx-L-pH sensors that have manufacturer reported performance of ±0.050, most of the ΔpH data points fall within this range (i.e., 2,166 data points or ca. 10% are beyond this threshold).

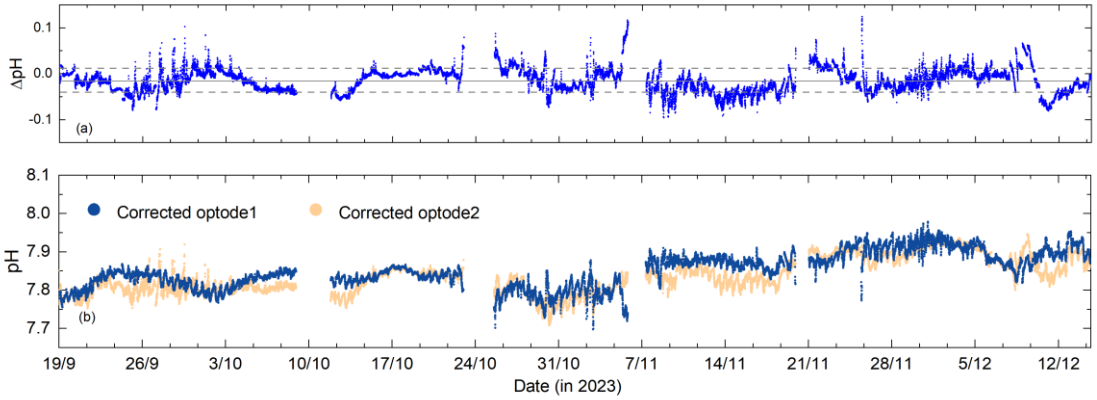


Figure S11. Optode pH sensor performance relative to one another. Top panel (a) shows a difference plot between the two optode pH sensor datasets. The grey solid line represents the mean error, and the grey dashed lines represent the standard deviation of the error. Bottom panel (b) shows both optode pH sensor datasets overlaid for the 3-month deployment period during phase 2.

References

Bresnahan, P. J., Takeshita, Y., Wirth, T., Martz, T. R., Cyronak, T., Albright, R., Wolfe, K., Warren, J. K., and Mertz, K.: Autonomous in situ calibration of ion-sensitive field effect transistor pH sensors, Limnology and Oceanography: Methods, 19, 132-144, https://doi.org/10.1002/lom3.10410, 2021.

Delaigue, L., Reichart, G. J., Qiu, L., Achterberg, E. P., Ourradi, Y., Galley, C., Mutzberg, A., and Humphreys, M. P.: From small-scale variability to mesoscale stability in surface ocean pH: implications for air–sea CO2 equilibration, Biogeosciences, 22, 5103-5121, 10.5194/bg-22-5103-2025, 2025.

Staudinger, C., Strobl, M., Fischer, J. P., Thar, R., Mayr, T., Aigner, D., Müller, B. J., Müller, B., Lehner, P., Mistlberger, G., Fritzsche, E., Ehgartner, J., Zach, P. W., Clarke, J. S., Geißler, F., Mutzberg, A., Müller, J. D., Achterberg, E. P., Borisov, S. M., and Klimant, I.: A versatile optode system for oxygen, carbon dioxide, and pH measurements in seawater with integrated battery and logger, Limnology and Oceanography: Methods, 16, 459-473, https://doi.org/10.1002/lom3.10260, 2018.

Wirth, T., Takeshita, Y., Davis, B., Park, E., Hu, I., Huffard, C. L., Johnson, K. S., Nicholson, D., Staudinger, C., Warren, J. K., and Martz, T.: Assessment of a pH optode for oceanographic moored and profiling applications, Limnology and Oceanography: Methods, 22, 805-822, https://doi.org/10.1002/lom3.10646, 2024.

Yin, T., Papadimitriou, S., Rérolle, V. M. C., Arundell, M., Cardwell, C. L., Walk, J., Palmer, M. R., Fowell, S. E., Schaap, A., Mowlem, M. C., and Loucaides, S.: A Novel Lab-on-Chip Spectrophotometric pH Sensor for Autonomous In Situ Seawater Measurements to 6000 m Depth on Stationary and Moving Observing Platforms, Environmental Science & Technology, 55, 14968-14978, 10.1021/acs.est.1c03517, 2021.