JGR Oceans

RESEARCH ARTICLE

10.1029/2025JC023213

Coastal Marine Carbon and Air-Sea Fluxes Quantified From pH Sensors on an Extended AUV Deployment

Emily M. Hammermeister[1,2], Cathy Wimart-Rousseau[2], Stathys Papadimitriou[2], Pablo Trucco-Pignata[2], Edward Chaney[2], Robert Templeton[2], Alexander B. Phillips[2], and Socratis Loucaides[2]

[1]School of Ocean and Earth Sciences, University of Southampton, Southampton, UK, [2]National Oceanography Centre, Southampton, UK

Key Points:

- Autonomous observations reveal fine-scale spatiotemporal variability and seasonal processes in carbonate chemistry and air-sea CO_2 fluxes
- The Celtic Sea acts primarily as a CO_2 sink in May–June; shown by one of the region's most detailed continuous carbonate data sets collected
- Small methodological differences can reverse interpretation of CO_2 flux direction, especially under high winds

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

E. M. Hammermeister,
emily.hammermeister@noc.ac.uk

Citation:

Hammermeister, E. M., Wimart-Rousseau, C., Papadimitriou, S., Trucco-Pignata, P., Chaney, E., Templeton, R., et al. (2026). Coastal marine carbon and air-sea fluxes quantified from pH sensors on an extended AUV deployment. *Journal of Geophysical Research: Oceans*, *131*, e2025JC023213. https://doi.org/10.1029/2025JC023213

Received 28 JUL 2025
Accepted 25 JUL 2026

Author Contributions:

Conceptualization: Emily M. Hammermeister, Edward Chaney, Robert Templeton, Alexander B. Phillips, Socratis Loucaides
Data curation: Emily M. Hammermeister, Stathys Papadimitriou, Edward Chaney
Formal analysis: Emily M. Hammermeister
Funding acquisition: Robert Templeton, Alexander B. Phillips
Investigation: Emily M. Hammermeister

Abstract For the first time, the Autosub Long Range (ALR) completed a fully autonomous, long-distance (2,000 km) scientific mission, delivering new insight into coastal carbonate dynamics and air-sea CO_2 fluxes. Equipped with a suite of oceanographic sensors, including a Lab-on-Chip (LOC) pH sensor and a Sea-Bird SeaFET pH sensor, the mission generated nearly 50,000 high-resolution pH measurements, providing one of the most detailed continuous coastal carbonate data sets collected to date in the region. We evaluated the adjustment of the SeaFET reference potential (k_0), testing both the co-deployed LOC sensor and neural network estimates as reference pH. Before correction, the LOC and SeaFET sensors showed close agreement ($\Delta pH_T = 0.013 \pm 0.009$), which improved to $\Delta pH_T = 0.00004 \pm 0.007$ after LOC-based k_0 adjustment. Both sensors diverged from model estimates, indicating reduced ability of models to resolve fine-scale coastal variability and reinforcing the need for direct in situ observations. Total alkalinity (TA) was derived from salinity-based relationships and model predictions, and paired with pH (SeaFET, LOC, and modeled) to estimate the partial pressure of CO_2 (pCO_2), which ranged 263–598 ± 27 µatm. Resulting air-sea CO_2 fluxes ranged −17.0 to 7.1 ± 1.09 mmol m^{-2} d^{-1}, with the Celtic Margin acting as a net CO_2 sink in May–June of 2022. pCO_2 and CO_2 flux proved sensitive to subtle pH differences, but less so to TA estimates. Our findings demonstrate the critical role of high-resolution autonomous observations in quantifying coastal carbonate dynamics and CO_2 fluxes, capturing processes and variability that are largely unresolved by ship-based surveys or global models.

Plain Language Summary For the first time, the Autosub Long Range (ALR), a robotic underwater vehicle, completed a long-distance (2,000 km) scientific mission without ship support. During the mission, the ALR collected ocean data using sensors that measured water temperature, oxygen, and seawater acidity (pH). The mission generated nearly 50,000 high-resolution pH measurements, providing one of the most detailed continuous coastal carbonate data sets collected in this region. We evaluated how well two onboard pH sensors agreed with each other and with model-based estimates, and tested an applied correction method. Before and after correction, the sensors closely agreed, showing that autonomous platforms can gather high-quality chemical measurements that resolve coastal marine carbon dynamics better than model estimates alone. We used the pH data, along with estimates of TA, to calculate carbon dioxide (CO_2) exchange between the ocean and atmosphere, finding that the region operates primarily as a carbon sink in May–June. Our calculations were strongly influenced by pH input but much less by alkalinity, highlighting the importance of accurate pH observations. This study shows that autonomous sensors can improve monitoring of carbon cycling in dynamic, understudied coastal waters that play a key role in the global carbon system.

1. Introduction

Carbon dioxide (CO_2) levels in the atmosphere have increased by nearly 52% since the onset of the Anthropocene, with 2024 emissions exceeding 11.4 GtC yr^{-1}, which is 0.3 GtC yr^{-1} higher than in 2023 (Friedlingstein et al., 2025). It is well defined that the ocean is a critical component of the global carbon cycle and plays a crucial role in moderating the rise in atmospheric CO_2. The global net uptake of CO_2 by the oceans, known as the ocean sink, was 2.9 ± 0.4 GtC yr^{-1} in the last decade, accounting for 26% of anthropogenic CO_2 annually (Friedlingstein et al., 2025; Quéré et al., 2009). The chemical composition of seawater is affected by the dissolution of CO_2, driving ocean acidification (OA) and altering the marine carbonate system, with lasting impacts that ripple across ecosystems, food webs, and global climate processes (Barker & Ridgwell, 2012; Caldeira & Wickett, 2003; Feely et al., 2004; Wolff, 2007). As atmospheric CO_2 continues to rise, the ocean's

Methodology: Emily M. Hammermeister, Cathy Wimart-Rousseau, Stathys Papadimitriou, Pablo Trucco-Pignata, Edward Chaney, Robert Templeton, Alexander B. Phillips, Socratis Loucaides
Project administration: Robert Templeton, Alexander B. Phillips
Resources: Cathy Wimart-Rousseau, Stathys Papadimitriou, Pablo Trucco-Pignata, Edward Chaney, Socratis Loucaides
Software: Emily M. Hammermeister, Cathy Wimart-Rousseau, Stathys Papadimitriou, Pablo Trucco-Pignata, Edward Chaney, Robert Templeton, Alexander B. Phillips, Socratis Loucaides
Supervision: Stathys Papadimitriou, Pablo Trucco-Pignata, Edward Chaney, Socratis Loucaides
Validation: Stathys Papadimitriou, Pablo Trucco-Pignata, Edward Chaney, Socratis Loucaides
Visualization: Emily M. Hammermeister, Pablo Trucco-Pignata, Socratis Loucaides
Writing – original draft: Emily M. Hammermeister
Writing – review & editing: Emily M. Hammermeister, Cathy Wimart-Rousseau, Stathys Papadimitriou, Pablo Trucco-Pignata, Edward Chaney, Robert Templeton, Alexander B. Phillips, Socratis Loucaides

capacity to buffer global climate change diminishes (Tans, 2009). Models now project a buffer capacity decline by up to 34% by 2,100, likely accelerating OA and related biogeochemical perturbations (Jiang et al., 2014). Inorganic carbon (the marine carbonate system) has been deemed an essential ocean variable by the Global Ocean Observing System; a nod to the criticality of ocean carbon observations for the assessment of the ocean carbon budget and the quantification of carbon transfer between the atmosphere and ocean (air-sea CO_2 fluxes) (Tanhua et al., 2019).

Over the past decades, high quality ship-based ocean carbon observations—driven by programs like GO-SHIP and Ships of Opportunity (SOOP), and data synthesis efforts such as the Global Ocean Data Analysis Project (GLODAP), and the Surface Ocean CO_2 Atlas (SOCAT)—have substantially advanced our understanding of ocean carbon cycling. These coordinated initiatives have provided extensive, quality-controlled data sets for global climate models quantifying air–sea CO_2 exchange and ocean carbon budgets (Khatiwala et al., 2013; McKinley et al., 2016; Schuster et al., 2013; Wanninkhof et al., 2013). However, ocean carbon data remain sparse in space and time especially in coastal regions and along continental margins where carbon cycling is poorly understood (Marrec et al., 2015; Regnier et al., 2013; Roobaert et al., 2019). Coastal areas exhibit high variability due to complex physical, biogeochemical, and anthropogenic influences, complicating efforts to scale localized observations to the global level. These highly productive regions, accounting for 15%–30% of global oceanic primary production, are characterized by enhanced air-sea CO_2 fluxes and are particularly vulnerable to anthropogenic forcings such as OA and eutrophication (Duarte et al., 2013; Gattuso et al., 1999; Marrec et al., 2015). Improving observational coverage and mechanistic understanding of coastal systems is critical for reducing uncertainties in the global carbon budget and refining climate projections and mitigation efforts. The study region for this work is the Celtic Sea (Figure 1), which is a seasonally stratified, productive northwest European shelf sea where spring blooms, benthic–pelagic coupling, and strong mixing drive highly variable carbonate dynamics and air–sea CO_2 exchange (Cai, 2011; Chen et al., 2013; Frankignoulle & Borges, 2001; Humphreys et al., 2019). Despite its ecological and economic importance, carbonate system observations in the Celtic Sea remain sparse and fragmented, historically relying on discrete cruise sampling, surface underway measurements, or predictive products (Borges & Frankignoulle, 2003; Hartman et al., 2014; Humphreys et al., 2019; Marrec et al., 2015).

Autonomous observing platforms have expanded high-resolution observations of key physical and biogeochemical observations. For example, the Biogeochemical Argo (BGC-Argo) Float Program is a global array of autonomous profiling floats equipped with sensors to measure key biogeochemical variables such as oxygen, pH, nitrate, chlorophyll, and particulate matter (Claustre et al., 2020; Roemmich et al., 2019). BGC-Argo has become a key tool, critical for tracking ocean carbon cycling, improving biogeochemical models, and supporting global carbon budget assessments. Efforts to integrate carbonate sensors onto other autonomous platforms such as Uncrewed Surface Vehicles (USVs), Autonomous Underwater Vehicles (AUVs), and gliders are also well underway. For example, Saildrones and Wave Glider USVs have been equipped with the partial pressure of CO_2 (pCO_2) systems (USVpCO_2), enabling CO_2 flux quantification (Sabine et al., 2020; Sutton & Sabine, 2023). Sensors for pH and pCO_2 have been integrated on gliders (e.g.., Hauri et al., 2024; Saba et al., 2019) including Lab-On-Chip (LOC) pH sensor for quantification of sea-atmosphere CO_2 fluxes in the North Sea and carbonate system dynamics by the Dotson Ice Shelf (Pickup et al., 2025; Possenti et al., 2021). More recently, LOC sensor prototypes measuring total alkalinity (TA) and pH enabled complete carbonate system characterization from a Long Range AUV (Hammermeister et al., 2025), marking the first time this was achieved using direct autonomous observations. Compared to ship-based surveys, autonomous platforms can sustain high-frequency measurements over extended time frames and across large spatial domains, improving ocean observation coverage while reducing reliance on ship time and associated operational costs.

The marine carbonate system is defined by four key variables: dissolved inorganic carbon, TA, pH, and pCO_2. Any two of these variables, combined with temperature, salinity, and pressure, can be used to calculate the remaining two variables through a well-established system of stoichiometric equations (Millero, 1995, 2007; Park, 1969). Currently, only pH and pCO_2 sensors are commercially available and widely used for autonomous subsurface deployment (Martz et al., 2015). However, the environmental covariance of pH and pCO_2 makes them the least desirable combination to constrain the carbonate system, as it leads to large uncertainty in the derived variables (Orr et al., 2018; Raimondi et al., 2019; Sutton & Sabine, 2023). Currently, carbonate system characterization from autonomous platforms (e.g., BGC Argo floats) relies heavily on algorithm-derived or estimates of TA paired with measurements of either pH or pCO_2. As a conservative parameter in ocean chemistry, TA

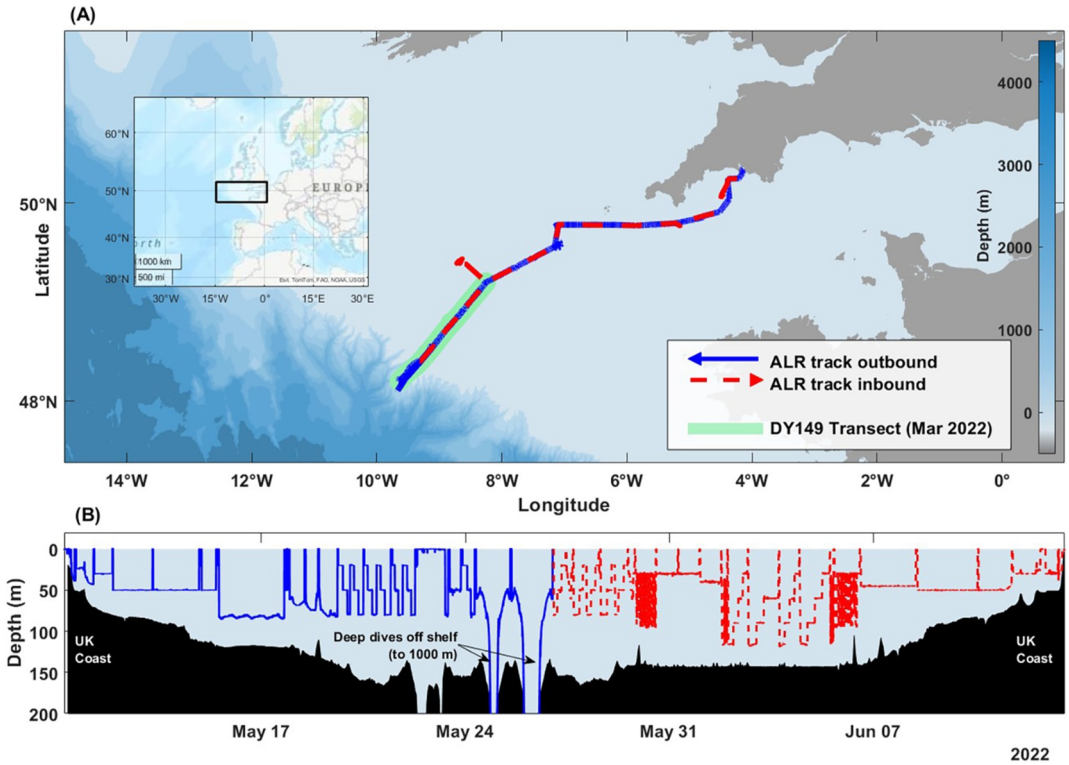

Figure 1. Long-Distance Proving Trial study site located Southeast off the coast of England in the Celtic Sea and continental margin. (a) Map of the entire deployment track of the Autosub Long Range (ALR) departing from and returning to Plymouth, UK. The green shaded area represents the DY149 ship-supported ALR transect conducted in March 2022 with ALR-2. (b) ALR transect and dive profiles plotted as depth (m) over time (mid-May to early June 2022), with seafloor bathymetry (black shading). For both (a) and (b) the outbound track of the ALR is shown as a solid blue line and the inbound track is shown as a dashed red line.

exhibits a strong linear relationship with salinity in the surface waters of the open ocean, allowing the development of salinity-based algorithms to predict TA where direct measurements are unavailable (Cai et al., 2010; Lee et al., 2006; Millero et al., 1998; Takahashi et al., 2014). The global coverage of the BGC Argo network has helped refine and validate more advanced models that incorporate temperature, oxygen, and other biogeochemical variables to estimate TA across a wide range of oceanic environments. These estimates are used with measured pH to calculate pCO_2 and its exchange with the atmosphere (Williams et al., 2017). The quality of these calculations is proportional to the uncertainty of both the TA estimation and the pH sensor measurements which can be a challenge to quantify over extended autonomous deployments. On BGC-Argo floats, pH sensor stability and accuracy is managed through frequent in situ recalibration: when the float is in deep (below 1,000 m) waters where pH is stable and well constrained (Maurer et al., 2021). However, this approach cannot be applied in shallower and more dynamic coastal environments, where pH variability and complex biogeochemical and physical processes (e.g., primary production/respiration, calcification/dissolution, air–sea CO_2 exchange, and mixing/freshwater inputs) can decouple carbonate system parameters and render model-based reference values less accurate. Consequently, monitoring pH sensor performance on autonomous platforms in coastal areas remains a significant challenge. Without reliable recalibration and validation methods in coastal zones, changes in pH sensor performance compromises the quality of derived carbonate system variables and reduces confidence in autonomous air–sea CO_2 flux estimates—especially in under-observed, highly variable environments where precise monitoring is most critical.

The work presented here addresses this challenge by investigating the use of a low-frequency (1 hr) photometric pH sensor as an onboard reference for periodic recalibration of high-frequency (1 min) electrochemical pH sensors, commonly used on profiling floats and other fast moving platforms. This approach was tested during the first shore-launched and recovered, long-range scientific mission of the Autosub Long Range (ALR) AUV. We evaluate the environmental data collected by the onboard sensors during the 35-day mission and conduct an

intercomparison of the two onboard pH sensors against community-adopted correction protocols and model estimated pH. Our analysis assesses the feasibility of sensor cross-validation for recalibration in coastal settings, examines the role of TA derivation and methodological choices, explores fine-scale spatiotemporal variability in carbonate chemistry, and evaluates the implications for determining pCO_2 and air-sea CO_2 fluxes. This work aims to demonstrate the new capabilities enabled by novel autonomous technology and represents a step toward improving the reliability and scientific value of carbonate system observations in previously under-sampled, highly dynamic coastal regions.

2. Materials and Methods

2.1. The Autosub Long-Range Autonomous Underwater Vehicle

ALR is a family of large, flight style AUVs (3.6 m length, 750 kg dry mass) developed and operated by the National Oceanography Centre (NOC), with depth ratings of either 1,500 m (ALR-1500) or 6,000 m (ALR-6000) (Roper et al., 2021). For a typical mission, they are programmed prior to launch with way points and a vertical sampling strategy (e.g., repeated profiles or depth-holds). Once deployed, the vehicle transits between way points while executing the planned dive pattern and sampling with its navigation and science payload. At the surface, the ALR GPS fixes which are communicated with the operators, and underwater it navigates primarily using Doppler Velocity Log (DVL)-aided dead reckoning, achieving accuracy <1% of distance traveled relative to distance traveled when within range of the seabed. Surface communication is achieved via Wi-Fi and acoustic links when available, with Iridium Short Burst Data (SBD®) as the primary channel for remote operations. When at the surface, operators receive information on vehicle health and mission progress and can modify the mission plan or sampling behavior. ALR AUVs use a "flight" configuration where large airfoil wings provide downward force and aft control surfaces regulate pitch and heading, enabling typical dive rates of 0.1–0.3 m s^{-1} at pitch angles of 10–30°. High energy-density lithium primary cells or lithium-ion secondary cells, combined with modest travel speeds, passive buoyancy control, and optimized power consumption of all onboard systems provide long endurance capabilities. ALR-4 (an ALR-1500) was used in this deployment and is hereafter referred to as the ALR.

2.2. Autonomous Mission

As a part of the ALR's extensive development program, the Long-Distance Proving Trial (LDPT) was defined to demonstrate the endurance capabilities of the ALR, particularly in a scientific capacity. From the 10th of May to the 13th of June 2022, the ALR transited independently in a southwest direction from Plymouth's coastal waters (50°18′N, 4°8′W), past the continental shelf into open ocean (Atlantic, 48°6′N, 9°39′W) and back (Figure 1a). The deployment consisted of dives lasting roughly 4–24 hr; after each dive the ALR telemetered back decimated data and received piloting instructions for the next dive. Over the course of the deployment, the ALR covered approximately 2,000 km across 54 programmed missions, collecting more than 800 hr of continuous observations. Following 5 weeks of operation at sea without topside vessel support, the ALR returned to Plymouth under its own power, marking its longest and furthest deployment to date (as of March 2022). Throughout the deployment, the ALR performed various sampling patterns at different depths and profiling frequencies (Figure 1b) and tested its ability to follow mission instruction dependent on sampling needs (Phillips et al., 2023).

2.2.1. Sensors and Equipment

The ALR has flooded payload bays that sit forward and aft of its main pressure hull and can be configured with a wide range of oceanographic sensors in addition to its standard sensor payload. As a part of its standard payload, the ALR was equipped with a pumped CTD (SBE 52-MP) sensor, which for this mission, was equipped with an auxiliary SBE 43F dissolved oxygen (DO) polarographic membrane sensor. The CTD-DO operated at a frequency (f) of 1 Hz (1 measurement per second) and provided measurements of in situ temperature (T), salinity (S), pressure (P), and DO concentration. The forward bay of the ALR was fitted with an Lab-On-Chip pH (LOC; NOC) sensor, and a Deep SeaFET V2 (SeaFET; SeaBird Scientific) electrochemical pH sensor (Figure 2 and Table 1).

The LOC pH sensor determines pH on the total proton scale (pH_T) photometrically using purified meta-Cresol Purple as the indicator dye. It achieves a precision of 0.001 pH units associated with a 0.003 ± 0.022 accuracy relative to validation seawater samples and a combined standard measurement uncertainty of ±0.010 pH

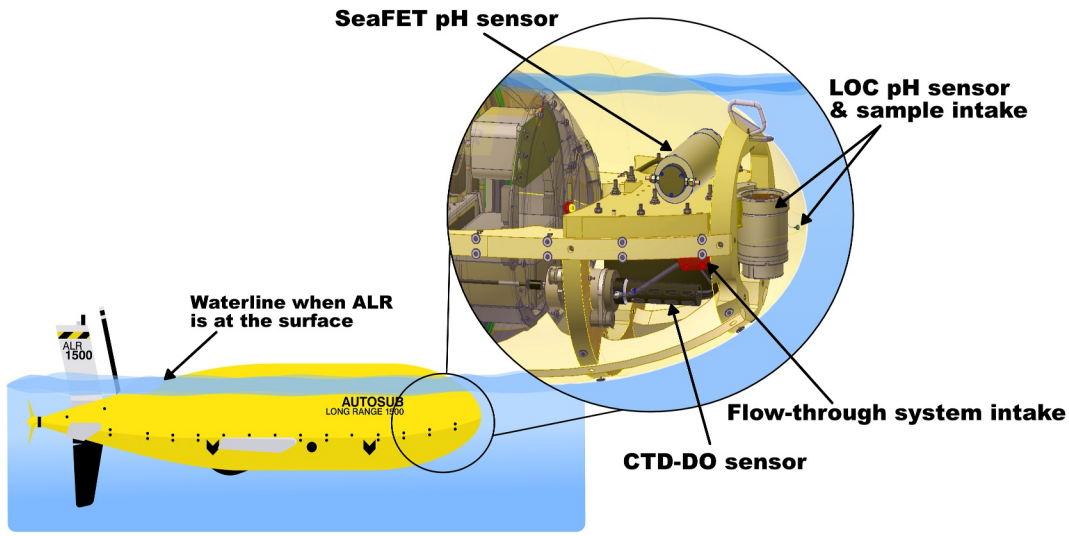

Figure 2. Schematic of Autosub Long Range (ALR) 1,500 (ALR) equipped with scientific instrumentation for the Long-Distance Proving Trial deployment. A zoomed view into the front bay of the ALR shows how the CTD-DO, each pH sensor, and sample intake systems, were situated. When positively buoyant at the surface, the ALR sits mostly submerged, with the sensor intakes approximately 0.4 m below the waterline.

units (Yin et al., 2021). The SeaFET uses a solid-state ion-sensitive field-effect transistor (ISFET) sensor and external chloride reference electrode to determine pH_T by detecting hydrogen ion concentration in seawater (based on Honeywell's Durafet ISFET technology; Johnson et al., 2016; Martz et al., 2010). The Deep SeaFET model uses only one reference electrode (external), unlike the Shallow SeaFET model, whose additional internal reference electrode with liquid junctions are not practical over the large pressure changes seen in this deployment (Bresnahan et al., 2014; Johnson et al., 2016; Scientific, 2019). The SeaFET sensor has a reported pH precision of ±0.02 pH units, which is therefore used here as an estimate of the SeaFET pH measurement uncertainty (Gonski et al., 2018; McLaughlin et al., 2017). A summary of the two pH sensors used in this deployment can be found in Table 1.

The LOC pH sensor drew power from an isolated 25.2 V 152 AH Lithium Thionyl Chloride battery pack, 1 of 30 total fitted to power the ALR. A DC-DC converter provided 12 VDC required by the sensor. The SeaFET was independently powered by its 12 D-cell alkaline internal batteries. The SeaFET sensor was fitted with a flow-cell cap and plumbed into the flow-through system shared by the CTD-DO sensors. Seawater was continuously pumped from an intake outside the ALR's fairing, 43 mm below the centerline of the vehicle (Figure 2). The LOC pH sensor sampled seawater through a 0.7 mm PEEK tube and filter (0.45 µm pore size and 2.5 cm diameter) perturbing the ALR'*S* front fairing, 30 mm below its centerline (Figure 2), which experienced no clogging or bio-fouling during the deployment. All sensors operated continuously (Table 1) throughout the ALR's deployment until its recovery. Although the ALR surfaced frequently for data transfer and navigation updates, sensor water intakes remained submerged approximately 400 mm below the sea surface (Figure 2).

Table 1
Summary of pH Sensors Used for Deployment

pH Sensor	Determination Method	Measurement Uncertainty	Measurement f (LDPT Deployment)	Depth Rating
Lab-On-Chip NOC (Southampton, UK)	Photometric, mCP reagent	0.01 pH	1 measurement/hour	6,000 m
SeaFET Sea-Bird Scientific (Seattle, WA, US)	Electrochemical, ISFET	0.02 pH	1 measurement/min	2,000 m

2.3. Sensor Data Processing

In situ T, S, and P measurements from the CTD-DO sensor were time-matched with pH sensor measurements manually post-recovery and were used to determine in situ pH_T as described in (Yin et al., 2021) for the LOC pH sensor and manufacturer's instruction (AN-99 (Sea Bird Scientific, 2023)) for the SeaFET pH sensor. More details on the calculations and calibrations of each sensor are provided in the (e.g., Table S1 in Supporting Information S1). The mixed layer depth (MLD) was determined using the threshold method with criterion $\Delta\sigma_\theta = 0.125$ kg m^{-3} where σ_θ represents potential density (de Boyer Montégut et al., 2004; Levitus, 1982; Monterey & Levitus, 1997).

2.3.1. Correction of pH_{SeaFET} Data

Although SeaFET sensors have demonstrated good measurement stability over time relative to other electrochemical and optical sensors (e.g., Bittig et al., 2018; Johnson et al., 2016), signal drifts over long autonomous deployments are unavoidable and hard to quantify. In comparison, spectrophotometric pH measurements are resistant to drift due to long-term stability of the indicator dye and principle of the measurement which is independent of changes in detector sensitivity and light source intensity (Müller & Rehder, 2018; Mosley et al., 2004; V. M. C. Rérolle et al., 2012; V. Rérolle et al., 2016). The SeaFET sensor measures in situ pH_T (pH_{SeaFET}) by detecting the voltage difference between the ISFET and the external solid-state Ag/AgCl reference electrode, which is directly exposed to seawater and sets the reference potential (k_0) using the salinity-dependent conservative concentration of chloride ions in seawater (Johnson et al., 2016; A. Dickson et al., 2007; Sea Bird Scientific, 2023). This measured potential is then transformed into pH on the total scale through calibration coefficients.

Despite careful calibration practices in the laboratory before deployment, k_0 can shift under in situ conditions. These shifts arise from the temperature dependence of the electrode slope and from intrinsic ISFET response characteristics that deviate from ideal Nernstian behavior, including temperature-dependent changes in gate-surface equilibria and external reference electrode potential (Bresnahan et al., 2014; Johnson et al., 2016, 2017; Miller et al., 2018). To combat the resulting sensor drift and reach accuracy levels required for global CO_2 system studies (Newton et al., 2015), pH_{SeaFET} data from Argo profiling floats are corrected by adjusting the sensor's k_0 (Johnson et al., 2016; Wimart-Rousseau et al., 2024). We note, however, that this approach is less robust than the full multi-point calibration procedure demonstrated by Miller et al. (2018), which uses several discrete samples to characterize sensor slope, temperature response, and reference-electrode behavior. Because no discrete bottle samples were collected prior to or during deployment, we were unable to apply the Miller et al. multipoint correction, and our interpretation reflects the limitations of a single point k_0 adjustment. Additionally, no discrete spectrophotometric pH samples, or concurrent ship-based carbonate measurements were collected along the LDPT track. Therefore, we cannot independently quantify absolute pH accuracy in situ during LDPT. Instead, we evaluate internal consistency between two sensors based on different measurement principles and treat model estimates and historical regional data sets (e.g., DY149) as contextual comparators rather than ground-truth validation.

The correction procedure developed for Argo float pH data (Maurer et al., 2021) assesses the changes in k_0 by comparing in situ pH values measured against a reference (pH_{ref}) to calculate an at-depth (typically around 1,500 dbar for Argo floats) pH anomaly. This approach was used in this study to correct the pH_{SeaFET} observations using a reference depth of 500 m adapted to account for the shallower ALR mission. As with the Argo float pH corrections, one pH_{ref} was calculated based on Empirical Seawater Property Estimation Routines from neural networks (ESPER-NN) (Carter et al., 2021). ESPER's Locally Interpolated Regression (ESPER-LIR) routine was also used to generate pH_T values for comparison but not used as a reference pH due to its similarity to the more rigorous ESPER-NN output. The ESPER-NN and ESPER-LIR routines produced pH_T estimates, hereafter labeled as pH_{NN} and pH_{LIR} respectively.

For this work, ESPER was the chosen routine for estimating pH given the following reasons:

1. ESPER's inclusion of new data from GLODAPv2.2020 data product.
2. The algorithm imitates the design of the widely regarded “Carbonate system and Nutrients concentration from hydrological properties and Oxygen using a Neural-network version B” (CANYON-B) algorithms by (Bittig et al., 2018; Sauzède et al., 2017).

3. Use of a new first-principles-based approach to estimate anthropogenic impacts on pH_T as updated from the OA adjustment presented with the locally interpolated pH regression (LIPHR) product by (Carter et al., 2018).
4. Quantifiable uncertainty estimation is returned by the routines.
5. Robust assessments, particularly in the North Atlantic region.

Additionally, Dias and Carter (2025) recently showed that ESPER-NN reproduces the GLODAPv2.2022 data product with a mean pH standard deviation of -4.59×10^{-3}, indicating that this reference can be comparable to discrete pH measurements. Similarly, Lauvset et al. (2024) reported strong correlation between neural network pH outputs and GLODAP data in some regions. Building on these results, ESPER-NN was used here as a correction reference to evaluate how model-derived estimates perform in this coastal setting. Estimates computed using both ESPER products relied on inputs *T*, *S*, DO, and depth (Carter et al., 2021), together with GPS coordinates measured by the CTD-DO and ALR systems, with the estimation date set to 2022. In this work, pH_{ref} is used as an independent reference for anomaly-based drift correction and is not intended to represent the unknown "true" pH. We correct the SeaFET pH time series because established SeaFET correction methods are available (Maurer et al., 2021). For comparison, we evaluate the use of two different pH_{ref} for the pH_{SeaFET} adjustments. Since the LOC pH sensor measurements (pH_{LOC}) were direct in situ observations with spatiotemporal crossovers occurring every ~60th SeaFET measurement, it was used as another pH_{ref}. Additionally, the LOC photometric measurements are comparatively drift-resistant and expected to provide a more stable reference. By correcting the high frequency pH_{SeaFET} measurements to the less frequent stable pH_{LOC} measurements, we maximize the strengths of each sensor, resulting in a high-resolution data set with improved reliability. The pH_{SeaFET} k_0-adjusted using pH_{LOC} will henceforth be referred to as pH_{k0-LOC} and the pH_{SeaFET} k_0-adjusted using pH_{NN} will henceforth be referred to as pH_{k0-NN}.

To determine the adjustment and overcome the offsets and drift calibration jumps that often occur during the sensor's life, the pH anomaly time series was divided into separate segments bound by breakpoint nodes determined by a cost function (Killick et al., 2012). The MATLAB function *ischange*, was used to perform the binary splitting that separated the segments to find the optimal location of an increasing number of breakpoints which is constrained by a threshold value on the mean residual. The threshold value—most commonly set to 0.005 pH units by the Argo community (Maurer et al., 2021)—specifies the sensitivity of change detection and is driven by the target accuracy of the sensor. If the threshold value is omitted or set too low, an excessive number of breakpoints will be detected and is equivalent to assuming no inherent noise (points with a few changes will be detected), which is inconsistent with known sensor behavior (Maurer et al., 2021). Next, a linear least squares fit was applied to each anomaly data series between each node, determining offset and drift values for each section (see Figure 5). Since the k_0 adjustment using pH_{LOC} was derived from observation-based pH_T with less measurements than pH_{SeaFET}, each linear rate of change segment was applied across the pH_{SeaFET} data set based on breakpoint nodes timestamps.

Conceptually, the linear rate of change should be applied to the k_0, assuming that the reference potential is the metric that may drift over time. Hence, a temperature-relation correction (TCOR) term is then applied to the determined correction value by normalizing the adjustment along the section to the temperature at which the adjustment was derived. Temperature-normalized changes in pH are calculated by multiplying the change in pH computed at depth by the ratio of the absolute temperature of the sample to the absolute temperature at reference depth. The TCOR therefore normalizes the applied adjustment to a constant deep reference temperature, ensuring that the correction reflects instrumental drift rather than thermally driven pH variability. The constant at-depth reference temperature value of 10°C is used as determined by the mean temperature recorded by the ALR below 500 m. Figure 5 in Section 3.1.2 illustrates the pH anomalies, identified breakpoints, and the segmented correction rates applied to the deployment time series.

2.3.2. Comparison of pH_T Data

The differences between pH_T values (ΔpH_T) measured by sensors and estimated from ESPER are calculated to evaluate the agreement between methods and variance within each method. Additionally, combined standard uncertainty, u_c, was calculated when comparing pH_T values. It was calculated as the positive square root of the

combined variances using Equation 1 and based on the documented uncertainties of the sensors and estimate routine outputs (JCGM et al., 2008).

$$u_c^2(y) = \sum_{i=1}^{N} u_i^2(y) \quad (1)$$

2.4. Carbonate System and Flux Calculations

2.4.1. Total Alkalinity Derivation

To constrain the marine carbonate system and consequently calculate CO_2 flux, another measurable carbonate input parameter is needed. We use four different methods to derive TA (TA_n) to compare the impact of each method on calculated pCO_2 and CO_2 flux. TA_1 was determined from salinity based on the relationship of Takahashi et al. (2014) and TA_2 was determined from salinity and temperature based on Lee et al. (2006). Other algorithms exist for TA estimation (e.g., Cai et al., 2010; Millero et al., 1998), however, the Takahashi et al. and Lee et al. relationships are best suited for the North Atlantic. TA_3 and TA_4 were determined by implementing ESPER-LIR and ESPER-NN (Carter et al., 2021) estimations respectively, using inputs T, S, DO concentration and GPS coordinates measured by the CTD-DO and ALR systems, with the estimate date set to 2022 (Table 2).

2.4.2. Calculation of pCO_2 (pH, TA)

The speciation of the carbonate system was characterized using the MATLAB (version v3.1.1) CO2SYS package (Lewis & Wallace, 1998; Orr et al., 2018; Sharp et al., 2023a; van Heuven et al., 2011) with measured pH_T and derived TA concentrations as the carbonate input parameters as well as T, S, and P from the CTD on board the ALR. Following the recommendations from (A. Dickson et al., 2007), CO2SYS computations used the dissociation constants of carbonic acid (K_1 and K_2) from (Lueker et al., 2000), K_{SO4} from (A. G. Dickson, 1990), KF from (Perez & Fraga, 1987), and total boron concentration from (Lee et al., 2010). Similar to TA, different pH_T measurements (both uncorrected and corrected) were evaluated to assess their impacts on the final pCO_2 values. Therefore, several combinations of pH_T and TA_n were used to derive seawater pCO_2 (Table 2).

The combined standard uncertainty of the resulting calculated pCO_2 from Table 1 was determined by error propagation using the CO2SYS function *errors.m* (Orr et al., 2018; Sharp et al., 2023b). The routine's default errors for dissociation constant inputs were used alongside individual parameter uncertainties (S, T, $pH_{1–3}$, $TA_{1–4}$) from table S2 in Supporting Information S1.

2.4.3. CO_2 Flux Calculations

To estimate the air-sea CO_2 fluxes from data collected by the ALR we use the bulk flux formulation:

$$CO_2 \text{ flux} = k_w \times K_0 \times \Delta pCO_2 \quad (2)$$

where CO_2 flux is the net CO_2 transfer (positive from ocean to atmosphere), k_w is the gas transfer velocity (m s^{-1}), K_0 is the solubility of CO_2 in seawater (mol m^{-3}), and $\Delta pCO_2 = pCO_{2,ocean} — pCO_{2,atm}$ is the air-sea difference in partial pressure of CO_2(µatm). As described in Section 2.4.2, xCO_2 (the mole fraction of CO_2 in dry air assuming 1 atm total pressure) was calculated using CO2SYS and filtered to include only values from when the ALR was at the surface, meaning all recorded values measured at the fixed sensor intake depth (~0.4 m below the water line). To get $pCO_{2,ocean}$ needed for flux calculations, the xCO_2 was then adjusted to be at equilibrium with surface waters in accordance with ocean CO_2 standard operating procedures (A. Dickson et al., 2007) using the following:

$$pCO_{2,ocean} = (P_{atm} - P_{H20}) \times xCO_2 \quad (3)$$

where P_{atm} = ambient atmospheric pressure (atm) and P_{H20} = vapor pressure (atm) over seawater at a given sea surface temperature (SST) and salinity (SSS), and xCO_2 = atmospheric CO_2 as mole fraction (ppm). The xCO_2 was obtained from the Penlee Point Atmospheric Observatory (PPAO; Plymouth Marine Laboratory and Yang (2017); Yang et al. (2019); Archibald et al. (2025)), part of the Western Channel Observatory, and was converted to $pCO_{2,atm}$ (µatm) according to Equation 4:

Table 2
Input Combinations of pH and Total Alkalinity Used to Calculate pCO_2

Label	pH input	TA input	pH correction
pH_1TA_1	pH_{SeaFET}	Takahashi et al. (2014)[a]	n/a
pH_1TA_2	pH_{SeaFET}	Lee et al. (2006)[b]	n/a
pH_1TA_3	pH_{SeaFET}	ESPER-LIR	n/a
pH_1TA_4	pH_{SeaFET}	ESPER-NN	n/a
pH_2TA_1	$pH_{k0\text{-}LOC}$	Takahashi et al. (2014)	pH_{SeaFET} k_0-adjusted using pH_{LOC}
pH_2TA_2	$pH_{k0\text{-}LOC}$	Lee et al. (2006)	pH_{SeaFET} k_0-adjusted using pH_{LOC}
pH_2TA_3	$pH_{k0\text{-}LOC}$	ESPER-LIR	pH_{SeaFET} k_0-adjusted using pH_{LOC}
pH_2TA_4	$pH_{k0\text{-}LOC}$	ESPER-NN	pH_{SeaFET} k_0-adjusted using pH_{LOC}
pH_3TA_1	$pH_{k0\text{-}NN}$	Takahashi et al. (2014)	pH_{SeaFET} k_0-adjusted using pH_{NN}
pH_3TA_2	$pH_{k0\text{-}NN}$	Lee et al. (2006)	pH_{SeaFET} k_0-adjusted using pH_{NN}
pH_3TA_3	$pH_{k0\text{-}NN}$	ESPER-LIR	pH_{SeaFET} k_0-adjusted using pH_{NN}
pH_3TA_4	$pH_{k0\text{-}NN}$	ESPER-NN	pH_{SeaFET} k_0-adjusted using pH_{NN}

[a]$TA = 45.30(S) + 733.00$; Location range: 40°N–55°N, 60°W–10°E.
[b]$TA = 2305 + 53.97(S - 35) + 2.74(S - 35)^2 - 1.16(T - 20) - 0.040(T - 20)^2$, Location range: Mid Atlantic Drift, 30°N–80°N.

$$pCO_{2,atm} = \left[\left(P_T - \left(\frac{RH}{100}\right) \times P_{H20}\right)\right] \times [xCO_2] \quad (4)$$

where P_T = total atmospheric pressure (in atm), RH = relative air humidity (percent), and P_{H20} = vapor pressure of water at ambient temperature (A. Dickson et al., 2007). The PPAO data set had a 15-min resolution which was linearly interpolated to match the 1-min resolution $pCO_{2,ocean}$ obtained by the sensors onboard the ALR (Figure S1 in Supporting Information S1). The solubility k_0 was computed using the T and S dependent formulation by (Weiss, 1974). The gas transfer velocity kw was parameterized following (Wanninkhof, 2014), which is defined as:

$$k_w = 0.251 \times (U_{10})^2 \times \left(\frac{Sc}{660}\right)^{-0.5} \quad (5)$$

Here U_{10} = wind speed at 10 m height (m s^{-1}), extracted from ERA5 reanalysis product (Hersbach et al., 2020), and Sc = Schmidt number. Specifically for U_{10}, the eastward and northward components of the 10 m wind were retrieved from the Climate Data Store (https://cds.climate.copernicus.eu/) at a 0.25 × 0.25 spatial resolution and hourly temporal resolution. To ensure temporal consistency, the wind time series was linearly interpolated to 1-min resolution (Figure S2 in Supporting Information S1) and time-matched to ALR surface measurements, enabling accurate flux estimations. The Schmidt number was calculated as a function of temperature in (°C) using:

$$Sc = 2116.8 - 136.25 \times T + 4.7353 \times T^2 - 0.092307 \times T^3 + 0.0007555 \times T^4 \quad (6)$$

Finally, the uncertainty of air-sea CO_2 fluxes (denoted as $u(CO_2$ flux)) was determined using first principles of error propagation of parameterized models that contain products or quotients of two or more parameters from Ellison and Williams (2012):

$$u(CO_2 \text{ flux}) = CO_2 \text{ flux} \sqrt{\left(\frac{u_{k_w}}{k_w}\right)^2 + \left(\frac{u_{K_0}}{K_0}\right)^2 + \left(\frac{u_{\Delta pCO_2}}{\Delta pCO_2}\right)^2} \quad (7)$$

where u_{k_w} is 20% (Wanninkhof, 2014), u_{K_0} is 0.20% (Ford et al., 2024; Weiss, 1974), and $u_{\Delta pCO_2}$ is equal to the propagated uncertainty of pCO_2 calculations from the surface (Table S3-a and S3-b in Supporting Information S1). Although $\Delta pCO_2 = pCO_{2,ocean} — pCO_{2,atm}$, for the purpose of this work, the uncertainty associated with ΔpCO_2 is represented only by $u(pCO_{2,ocean})$ because of the negligible contributions of $u(pCO_{2,atm})$.

3. Results

3.1. Autonomous Data Collection

Following the 5-week autonomous deployment, the ALR was recovered in mid-June near the Plymouth breakwater. While operating offshore, the ALR reached depths of up to 1,000 m, but the majority of measurements were obtained within the upper 100 m of the water column in coastal waters.

3.1.1. Biogeochemical Water Properties

Figure 3 shows the temporal and spatial trends in T, S, and DO concentration across the deployment. The full range of T (9.1–16.5°C), S (33.2–35.9 psu), and DO concentration (174–317 μmol kg^{-1}) is seen in Figures 3a, 3d and 3g. However, because the ALR spent 90% of the deployment within the top 100 m of the water column, a magnified view of T, S, and DO concentration in the water column over time is presented in Figures 3b, 3e and 3h, revealing the MLD varying between 4 and 52 m over the 5-week mission. Surface conditions (top 20 m), shown in Figures 3c, 3f and 3i indicate a steady increase of T and decrease of DO concentration over time. Salinity peaked when the ALR moved off the shelf and dropped near the coast at the mission's start and end, likely due to freshwater inputs (Figures 3d–3f).

3.1.2. pH Data

The pH sensors onboard the ALR operated the entirety of the LDPT, with neither showing evidence of hardware, software, or power malfunction. The LOC pH sensor made 858 pH_T measurements, and the SeaFET pH sensor made 49,184 pH_T measurements. Prior to any pH_{SeaFET} corrections, combined sensor pH_T measurements ranged 7.898–8.187 ($1\sigma = 0.380$, $n = 50{,}042$), with generally good agreement between the two sensors (Figure 4a). The mean uncorrected pH difference between the two sensors ($\Delta pH_{T\text{–uncorrected}}$) was 0.013 ± 0.009 pH units (1σ, $n = 858$) with 87% of $\Delta pH_{T\text{–uncorrected}}$ falling within the combined measurement uncertainty (of ±0.022 pH units) (Figure 4b). The pH_{SeaFET} required approximately 10 days to stabilize, exhibiting reduced drift and improved consistency with pH_{LOC} measurements (Figures 4a and 4b). This behavior, known as "electrode conditioning", is a well-documented characteristic of ISFET technology (Bresnahan et al., 2014; Martz et al., 2010).

After correcting pH_{SeaFET} with the k_0 adjustment using pH_{LOC} as the reference pH (see Section 2.3.1 above and Figure 5 below), all pH_T measurements from both pH sensors ranged from 7.911 to 8.200 (average $\pm 1\sigma = 0.038 \pm 0.038$, $n = 50{,}042$). As expected, the two sensors showed near-perfect agreement, with the SeaFET's electrode conditioning period effectively removed (Figure 4c). The average $\Delta pH_{T\text{–corrected}}$ approached zero and the standard deviation was improved ($1\sigma = 0.007$), with 98% of $\Delta pH_{T\text{–corrected}}$ values falling within the combined measurement uncertainty (±0.022 pH units) of the sensors (Figure 4d). The time series comparison before and after correction using pH_{NN} as the reference pH_T is not shown here due to its less effective performance and noisier appearance, but can be found in Figure S3 in Supporting Information S1.

Figure 5 illustrates the adjustment process by exhibiting the pH anomaly changes (ΔpH_T) throughout each data set (n sample) comparing pH_{SeaFET} values against pH_{LOC} measurements (Figure 5a) and pH_{NN} outputs (Figure 5b). Breakpoint nodes were identified using respective threshold values to segment the data and calculate linear rate of changes. Each threshold value was modified to preclude the determination of unrealistic breakpoint nodes. A threshold of 0.00500 was used for the pH_{LOC} data set that yielded two clear breakpoints and three segmented correction rates to be applied (Figure 5a). Given the particularly noisy anomaly series using pH_{NN}, a higher threshold value of 0.10000 was used, producing 20 breakpoints and corresponding segments (Figure 5b). Threshold values are regarded in pH units. The pH_{LOC}-based k_0 adjustment was derived from $n = 858$ observational pH_T values, while the pH_{NN}-based k_0 adjustment used $n = 49{,}184$ neural network pH_T estimates that were aligned pH_{SeaFET} timestamps ($n = 49{,}184$).

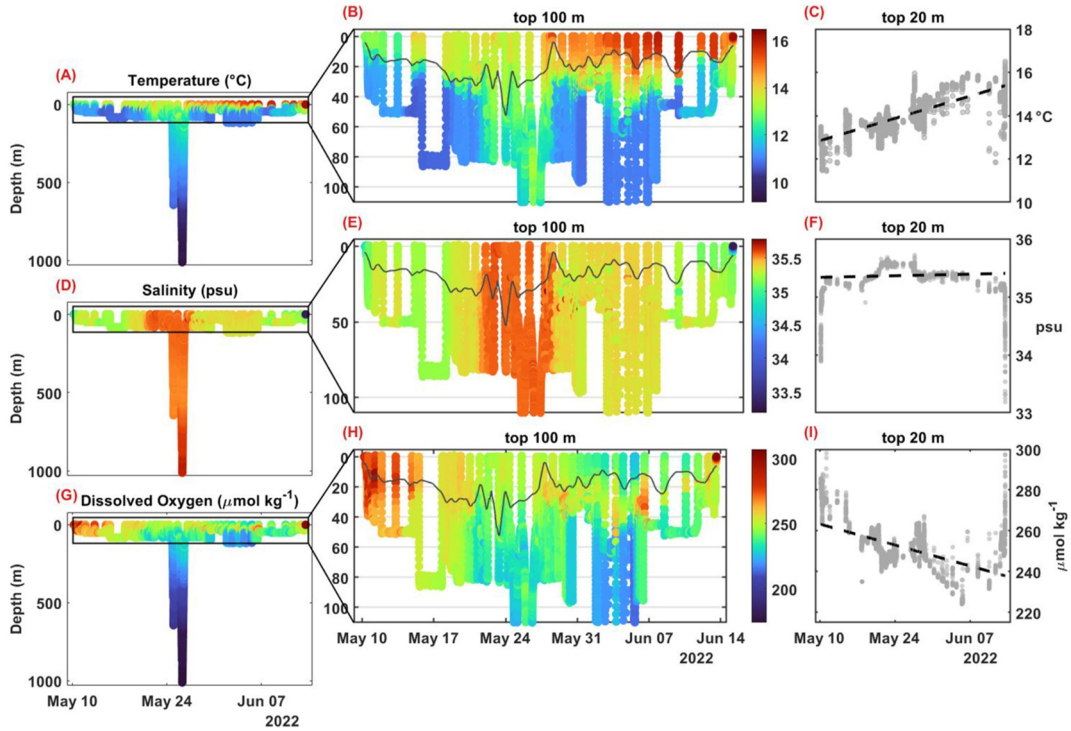

Figure 3. Autonomous high-resolution observations of temperature (°C), salinity (psu), and dissolved oxygen (DO) (μmol kg^{-1}) collected by sensors onboard the Autosub Long Range during the 5-week Long-Distance Proving Trial deployment (May–June 2022). Panels (a), (d), and (g) show full-depth profiles (to ~1,000 m) for temperature, salinity, and DO (represented by color), respectively throughout the deployment. Panels (b), (e), and (h) present a magnified view of the top 100 m, capturing most of the ALR's observational period. The black line in panels (b), (e), and (h) indicates the estimated mixed layer depth over time. Panels (c), (f), and (i) display conditions in the upper 20 m of the water column over time showing surface temporal trends over the course of the deployment.

Results from this section provide strong evidence that pH_{k0-LOC} quantifiably represents the most reliable pH determination and correction with the least associated uncertainty, and will henceforth be regarded as such. This choice does not remove the limitation that absolute pH accuracy could not be independently verified without coincident discrete samples. In the following sections, downstream products calculated from pH_{k0-LOC} (or pH_2) are emboldened (green) in figures comparing observation methods, and serve as the primary data set for scientific interpretations.

3.2. pCO_2 Data

Using various combinations of three different pH_T data sets and four different derived TA data sets, oceanic pCO_2 values were calculated during the LDPT, and ranged from 263 to 598 μatm (n = 590,208, all combinations) with average propagated uncertainty ranging ±23.4–30.4 μatm between combinations (Table S3-b in Supporting Information S1). Figure 6 shows the distribution of pCO_2 generated by each data set combination. The pCO_2 results highlight a noticeable pH input dependency, the extent of which is further supported by the matrix analysis later in this section. Notably, the highest pCO_2 values (up to 598 μatm) were predominately associated with the pH_3 (pH_{k0-NN}) input, while the lowest values were associated with the pH_2 (pH_{k0-LOC}) combination.

The pCO_2 results can be grouped according to the input pH_T used in the carbonate calculation. When calculated using pH_{SeaFET} (pH_1), the average pCO_2 seen is 413 ± 39 μatm (1σ, n = 196,736) across the deployment. In comparison, pCO_2 calculated using pH_{k0-LOC} (pH_2) yielded a lower mean value of 399 ± 39 μatm, while calculations using pH_{k0-NN} (pH_3) produced a higher average of 438 ± 38 μatm and showed a greater number of outlying values (Figure 6). We assessed the reliability of our seawater pCO_2 calculations by comparing them to available observational data sets. Within the study region, GLODAP data (1981–2019; upper 1,000 m) give a mean pCO_2 of 378.3 μatm (Olsen et al., 2016, 2017). Available SOCAT data overlapping our record (one day)

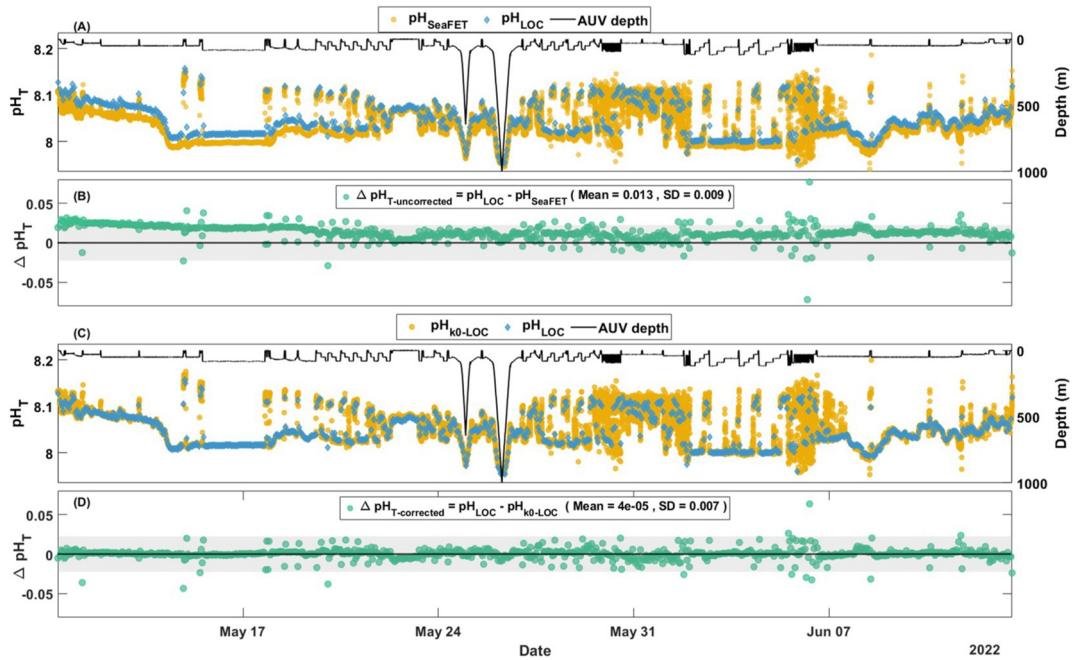

Figure 4. Time series comparison of pH measurements and offsets between SeaFET pH sensor and LOC pH sensor before and after k_0 adjustment (see Section 2.3.1). (a) In situ uncorrected pH_T data from the SeaFET sensor (yellow) and pH_{LOC} measurements (blue diamonds) shown alongside the transiting depth of the Autosub Long Range (black line). (b) Uncorrected pH difference ($\Delta pH_{T-uncorrected} = pH_{LOC} - pH_{SeaFET}$), with a mean offset of 0.013 pH units and standard deviation of 0.009 pH units. (c) SeaFET sensor pH referenced to LOC sensor pH via the k_0 adjustment (yellow), and pH_{LOC} (blue diamonds). (d) ($\Delta pH_{T-corrected} = pH_{LOC} - pH_{k0-LOC}$), with mean offset of 4×10^{-5} pH units and reduced standard deviation of 0.007 pH units. Shaded gray regions in (b) and (d) span ±0.022 pH units, representing combined sensor uncertainty u_c. The k_0-correction construction is shown in Figure 5.

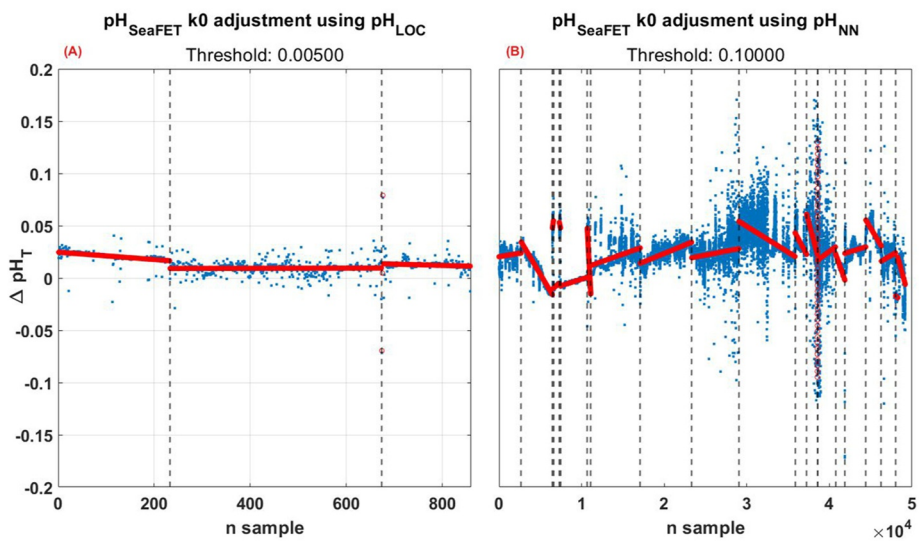

Figure 5. Correction of pH_{SeaFET} values with k_0 adjustment method using (a) pH_{LOC} and (b) pH_{NN} as the pHref. The plots represent the pH difference (ΔpH_T on *y*-axis) between each pH_{ref} and pH_{SeaFET} for each measurement taken (n sample on *x*-axis) represented by blue dots. The vertical dashed black lines indicate each instance where a breakpoint node was identified based on the defined threshold criteria (a: 0.00500, b: 0.10000). Red lines show each segmented linear correction rate to be applied between each breakpoint. The local calibration method (a) shows tighter agreement and lower variance compared to the neural network adjustment (b).

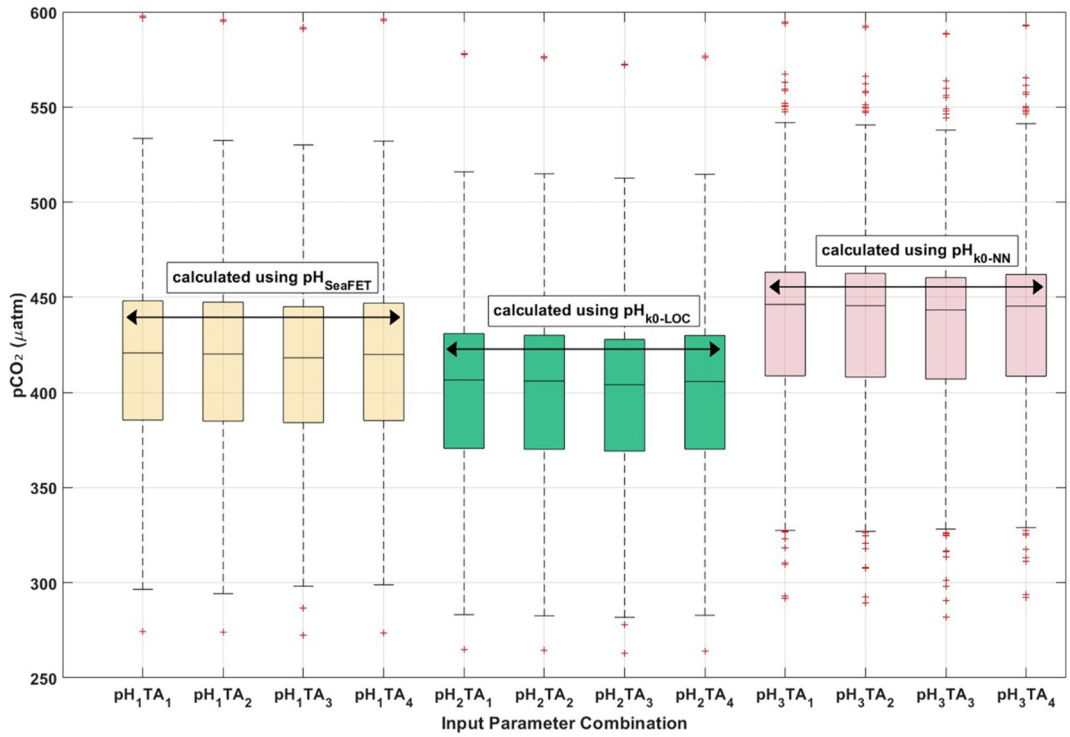

Figure 6. Comparison of calculated seawater partial pressure of CO_2 in µatm (pCO_2, *y*-axis) using different input pH and total alkalinity (TA) combinations (See Table 2). Boxplots show pCO_2 distributions grouped by pH input (1) uncorrected SeaFET sensor pH measurements (yellow), (2) k_0-adjusted pH using pH_{LOC} (pH_{k0-LOC}, emboldened green), and (3) k_0-adjusted pH using neural network-derived estimations (pH_{k0-NN}, red). For each method, four TA values (TA_1–TA_4) were paired with corresponding pH inputs to assess sensitivity. The central line in each box indicates the median value of each pCO_2 distribution, the box bounds correspond to interquartile ranges, whiskers extend to data within 1.5 times the interquartile range, and red crosses indicate outliers.

show mean fCO_2 of 393.3 ± 4.1 µatm, consistent with our estimates, and broader SOCAT coverage (2019–2023) shows a similar regional pattern, further supporting the validity of our approach (Bakker et al., 2025).

When plotted against time (Figure 7), the pCO_2 values from each group track broadly similar temporal patterns, responding coherently to environmental changes and vehicle depth during the deployment. However, consistent offsets persist between the different pH input types. The black trace, representing ALR depth, reveals that the largest pCO_2 values and sensor divergences often coincide with periods of vertical profiling or rapid depth change, a potential indication into when sensor calibrations or correction becomes most impactful. Notably, pH_{k0-NN}-based pCO_2 values remain persistently higher throughout the mission, amplified during deeper dives, while pH_{k0-LOC}-based values are tend to be lower (Figure 7). These offsets, while systematic, also appear to diverge slightly at certain timescales, suggesting both sensor-specific differences and variable sensitivity to environmental conditions.

A matrix analysis compared the mean pCO_2 differences across various input combinations, revealed that TA has an overall contribution the derived pCO_2 rather marginal relative to the dominant impact of pH input. When the pH_T is held constant for calculating pCO_2, results vary by a maximum of 3 µatm. In contrast, holding TA constant while varying pH_T leads to pCO_2 differences up to 42 µatm (Figure S5 in Supporting Information S1).

Figure 8 shows the temporal and spatial trends of pH_{k0-LOC} and pCO_2 calculated using pH_{k0-LOC} and TA_1 (Lee et al., 2006). The full range of pH_{k0-LOC} (7.911–8.200) and pCO_2 (265–578 µatm) is seen in Figures 8a–8d, and a magnified (top 100 m) view of pH_{k0-LOC} and pCO_2 is seen in Figures 8b and 8e. Throughout the water column, as depth increases pH_T decreases and pCO_2 increases. Surface values (top 20 m), shown in Figures 8c and 8f for both parameters are variable but indicate a slight decrease in pH_T and increase pCO_2 over time. Notably, there is a drop

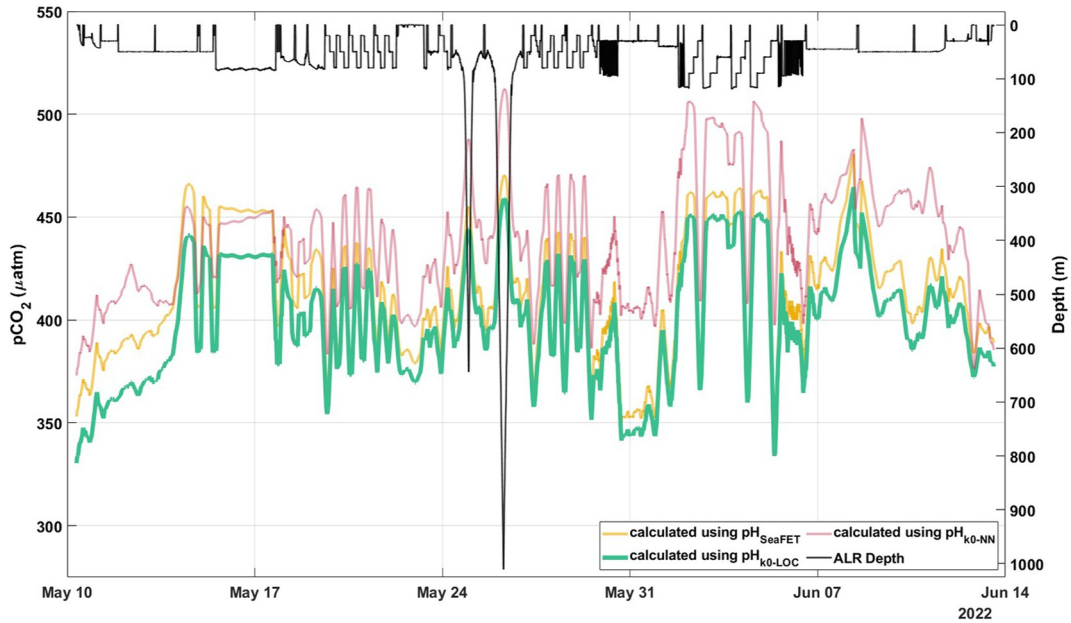

Figure 7. Time series of partial pressure of CO_2 in µatm (pCO_2, left y-axis) grouped by pH input used in carbonate calculations: pH_{SeaFET} (yellow), LOC k_0-corrected (pH_{k0-LOC}, emboldened green), and neural network k_0 corrected pH (pH_{k0-NN}, red). All pCO_2 shown here were calculated using TA_1 as the other carbonate input parameter. The Autosub Long Range transect depth (black, right y-axis) is overlaid.

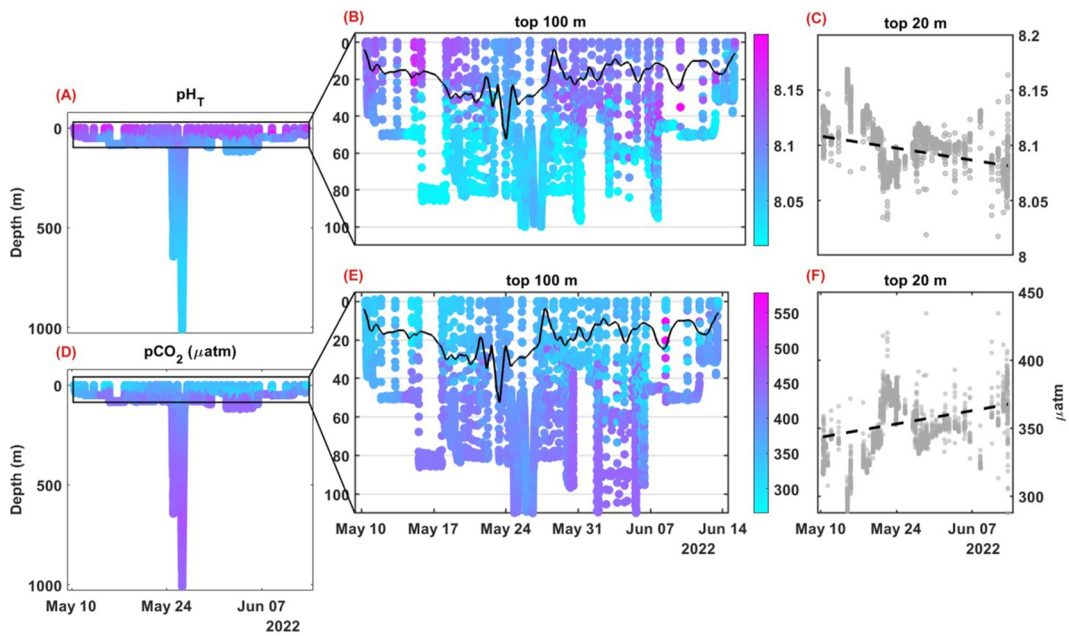

Figure 8. Autonomous high-resolution observations of pH_T and pCO_2 (µatm) during the 5-week Long-Distance Proving Trial deployment (May–June 2022). Plotted pH_T measurements (panels a–c) are from the SeaFET pH sensor k_0-corrected to the LOC pH sensor (pH_{k0-LOC}). Plotted pCO_2 values (panels d–e) are calculated using pH_{k0-LOC} and TA_1 as input parameters. Panels (a) and (d) show full-depth profiles (to ~1,000 m) of pH_T and pCO_2 respectively (represented by color), throughout the deployment. Panels (b) and (e) present magnified views of the top 100 m. The black line in panels (b) and (e) indicates the estimated mixed layer depth over time. Panels (c) and (f) display parameter values in the upper 20 m of the water column over time.

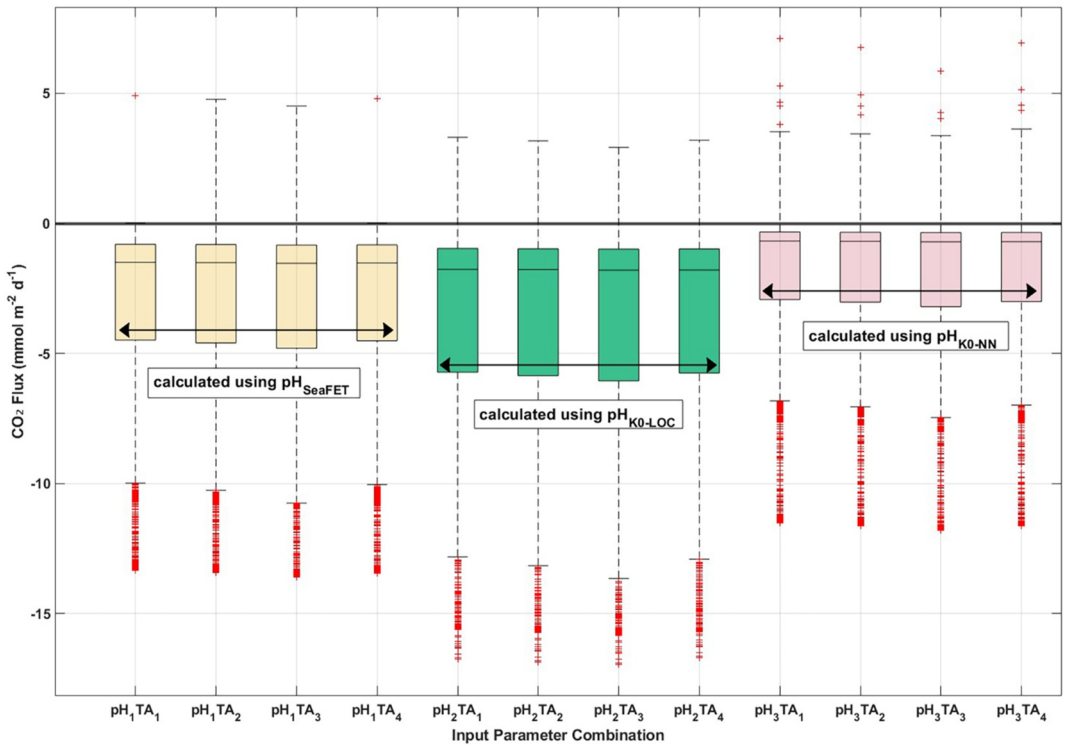

Figure 9. Comparison of calculated CO_2 flux (mmol m^{-2} d^{-1}, *y*-axis) using different input pH and total alkalinity (TA) combinations. Boxplots show CO_2 flux distributions grouped by pH input (1) uncorrected SeaFET sensor pH measurements (yellow), (2) k_0-adjusted pH using pH_{LOC} (pH_{k0-LOC}, emboldened green), and (3) k_0-adjusted pH using neural network-derived estimations (pH_{k0-NN}, red). For each method, four TA values (TA_1–TA_4; defined in Table 2) inputs were paired with corresponding pH inputs to assess sensitivity. The central line in each box indicates the median value of the distribution, boxes represent inter-quartile ranges, whiskers extend to data within 1.5 times the inter-quartile range, and red crosses indicate outliers. All flux calculations were done using only measurements from when the Autosub Long Range was at the surface (water intake 0.4 m below the waterline).

in surface pH_T and peak in surface pCO_2 around 24 May 2022, which coincides with the deepest MLD, when the ALR is furthest from shore and crossing continental shelf break (Figures 1 and 8).

3.3. CO_2 Flux

Building on the pCO_2 estimates, CO_2 fluxes were calculated to evaluate air–sea dynamics during the LDPT. Given these flux calculations are derived directly from surface pCO_2 values, they exhibit similar dependencies on input parameter combinations. During the LDPT, CO_2 fluxes ranged from -17.0 to 7.1 ± 1.1 mmol m^{-2} d^{-1} median uncertainty ($n = 43{,}152$, all combinations), depending on pH_T and TA data sets used (Table S3-b in Supporting Information S1). By convention, negative fluxes indicate uptake of CO_2 by the ocean (i.e., the ocean is a carbon sink), while positive fluxes indicate CO_2 outgassing from the ocean to the atmosphere. Figure 9 shows the distribution of air-sea CO_2 fluxes determined using various pCO_2-derived input combinations. Consistent with the pCO_2 results, flux estimates cluster primarily according to the pH_T data set used. When using pH_1 (SeaFET), the mean flux was -3.5 mmol m^{-2} d^{-1} ($1\sigma = 3.6$, $n = 3{,}596$). Using pH_2 (k_0-LOC), resulted in a more negative mean flux of -4.3 mmol m^{-2} d^{-1} ($1\sigma = 4.3$), while pH_3 (k_0-NN), produced a less negative mean of -2.4 mmol m^{-2} d^{-1} ($1\sigma = 2.9$). Across all methods, the mean air-sea CO_2 fluxes were negative, indicating that the region acted as a net CO_2 sink during the deployment period. While we present method-specific ranges, we do not propagate uncertainty from the input variables (pH, TA, wind speed, atmospheric CO_2 into the final flux estimates. As such, the variability across methods should be considered a minimum bound of potential uncertainty in calculated air-sea CO_2 flux.

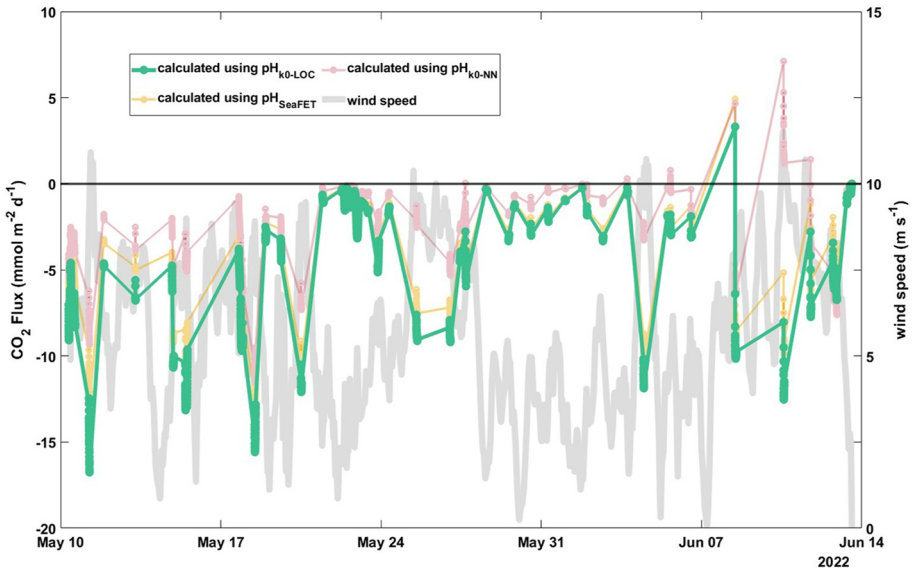

Figure 10. Time series of air–sea CO_2 fluxes (mmol m^{-2} d^{-1}, left *y*-axis) grouped by pH input to calculations: pH_{SeaFET} (yellow), LOC k_0-corrected (pH_{k0-LOC}, emboldened green), and neural network k_0 corrected pH (pH_{k0-NN}, red). All CO_2 fluxes shown here used TA_1 as the other carbonate input parameter when calculated. All flux calculations were done using only measurements from when the Autosub Long Range was at the surface (water intake 0.4 m below the waterline). Wind speed in m s^{-1} (gray line, right *y*-axis) is overlaid.

When the different flux estimates (grouped by pH input as defined in Figure 9 above) are plotted over time alongside wind speed, variability stemming from differences in input parameters and environmental factors such as wind can be observed. While wind does not directly cause divergence between methods, it modulates the magnitude of the flux response and can highlight where methodological differences become most impactful. Figure 10 illustrates the noticeable impact of high wind speed period on air-sea CO_2 fluxes variability, with periods of elevated wind speed often coinciding with amplified discrepancies between flux estimates (Figure 10). In some cases, flux calculated using pH_{SeaFET} and pH_{k0-LOC} show different responses to increased wind than flux calculated using pH_{k0-NN}. Notably, during the week of June 07–14, fluxes derived from pH_{SeaFET} and pH_{k0-LOC} exhibit different responses to wind variability compared to those based on pH_{k0-NN}, including a marked divergence in flux direction, underscoring the sensitivity of flux calculations to both sensor corrections and external forcing.

Given the high variability and wide range of values in the flux data, the performed matrix analysis compared the median fluxes from each input combination, thereby mitigating the influence of extreme values on the overall interpretation (Figure 10, Figure S6 in Supporting Information S1). When pH is held constant across different TA inputs, median flux differences within approximately 0.04 mmol m^{-2} d^{-1} are observed. Conversely, keeping TA constant while varying pH input induces differences that can exceed 1.11 mmol m^{-2} d^{-1} (Figure S6 in Supporting Information S1). This difference represents approximately 25% of the average flux magnitude, which could meaningfully alter seasonal or annual budget estimates and interpretations when extrapolated over time and space. These results, consistent with the pCO_2 analysis, underscore the implications of selecting and validating input sources in carbonate chemistry, but particularly with pH in CO_2 flux calculations.

4. Discussion

4.1. Autonomous Coastal Observations

The LDPT deployment demonstrates not only the robust endurance of the ALR and onboard sensors to sample seawater properties across coastal and open ocean waters, but more critically, its ability to autonomously monitor the marine carbonate system with built-in quality control. This represents a key step toward increasing availability of comprehensive carbonate measurements of coastal regions, building on and complementing ongoing monitoring efforts. While remote human oversight is still needed for mission programming and ALR position

monitoring, this approach also substantially reduces the carbon footprint and operational costs compared to traditional ship-based surveys (Phillips et al., 2023).

Over the 5-week deployment, continuous high-resolution profiles and transects of temperature, salinity, DO, and pH throughout the water column were successfully recorded. These measurements revealed distinct seasonal and spatial patterns across the Celtic continental margin as this deployment occurred during a season of heightened coastal productivity, when biogeochemical conditions shift rapidly over short timescales and small spatial scales. For instance, surface-layer DO concentration declined steadily as temperature increased, suggesting biological consumption likely linked to the waning spring bloom (Figure 3i). Spatial salinity trends highlighted the transition between coastal and offshore domains, with a shift from the fresher coastal waters to saltier offshore conditions, though no evident temporal salinity trends were detected (Figures 3d, 3e and 3f). Added temporal resolution is especially important when placed in the context of traditional ship-based surveys. Between 19 and 30 March 2022 (two months prior to the LDPT deployment), the ALR6000 and CTD sampling surveyed the same transect (Figure 1) on the Celtic shelf margin supported by the Royal Research Ship, *RRS Discovery*, during expedition DY149.

Within the upper 40 m of the water column, TA values from the 2022 LDPT deployment closely align with those from the 2022 DY149 deployment, with all TA estimates falling with the DY149 range of 2,325–2,345 μmol kg^{-1} (Hammermeister et al., 2025). This agreement reinforces the conservative behavior of TA in this region (Hartman et al., 2014) and supports the use of the chosen TA parameterizations for this work. The pH_T and pCO_2 values observed during LDPT, although calculated using various methodologies, span broader distributions that extend beyond the DY149 ranges (8.03–8.08 for pH_T and 375–415 μatm for pCO_2). Such divergence is expected, given the high temporal variability of these non-conservative parameters, which are subject to pronounced seasonal variability. The LDPT deployment took place during late May, which is a period typically characterized by increasing primary production, stratification, and warming. Each of which are factors that contribute to elevated pH and decreased pCO_2 that we see in Figure 8. The LDPT pH and pCO_2 observations are consistent with prior studies from the Celtic margin (Hartman et al., 2014; Joint et al., 2001; Kitidis et al., 2012; Seguro et al., 2019), indicating that the LDPT (May 2022) values reflect expected natural seasonal shifts rather than sensor bias or error when compared to the DY149 (March 2022) values.

Beyond seasonal trajectories, diel signals could also be resolved from our data set. Figure 11 shows mean DO, pH_T (pH_{k0-LOC}), and pCO_2 (from pH_{k0-LOC} and TA_1 Lee et al. (2006)) in the upper 100 m of the water column binned by hour of the day, for the entire LDPT deployment. We focus on the upper 100 m here to capture the bulk of the measurements made during the deployment and target the photic upper ocean where diel biogeochemical signals are commonly expressed. During the day, DO concentration and pH_T peak, while pCO_2 drops, indicating biological control though photosynthesis and respiration, supported by previous reports during the Celtic Sea spring bloom (Seguro et al., 2019). The error bars in Figure 11 represent one standard deviation of each mean, revealing the wide variability among all three parameters. This is an unsurprising reflection of the variation seen from aggregated data across the entire deployment's spatial domain (2,000 km), where we see other seasonal and spatial trends take place over the 35 days of observation. Impressively, a diel signal still emerges through that variability, reinforcing the presence of biological and/or physical forcing in the top 100 m of the Celtic Margin. In other words, this underscores how daily cycles can still drive short-term carbonate dynamics, even when spatial and temporal-scale variability is high. This is a signal uniquely captured by this type of autonomous, high-frequency deployment, as ship-based surveys and models would have likely not resolved it at this scale.

These results provide rare, high-resolution evidence of how seasonal and short-term productivity transitions modulate carbonate chemistry in shelf seas, resolving processes at timescales inaccessible to ships. By extending beyond the static ship-based snapshot of March, the ALR revealed the onset of late-spring carbonate changes, underscoring the importance of autonomous observations to close seasonal gaps in carbon cycle assessments. Short-term AUV observations capture the natural seasonal trajectory beyond ship-based "snapshots," which often miss the dynamics between expeditions. This speaks directly to the importance of autonomous deployments to constrain seasonal carbon budgets in continental margins for better understanding of the tightly coupled biological and physical processes that drive coastal carbonate variability.

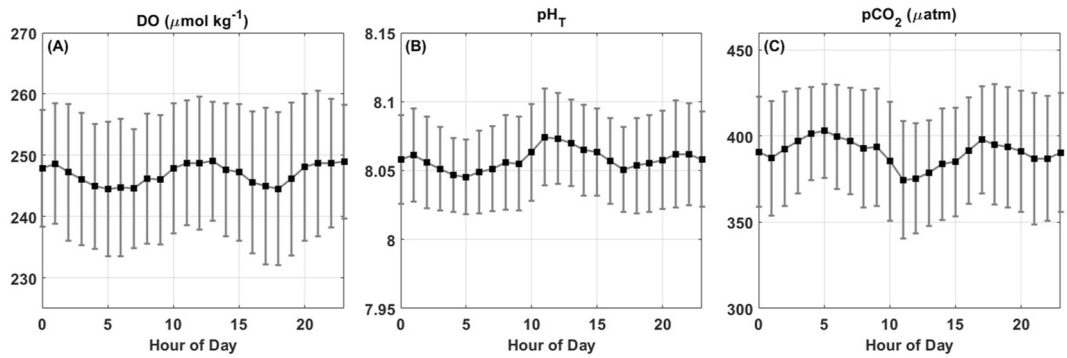

Figure 11. Diel cycling of (a) dissolved oxygen concentration (μmol kg^{-1}), (b) pH_T, and (c) pCO_2 (μatm) in the upper 100 m of the water column across the 5-week Long-Distance Proving Trial deployment. Each panel shows mean values binned by hour of the day, with the error bars representing one standard deviation. pH_T is pH_{k0-LOC}, and pCO_2 is calculated from pH_{k0-LOC} and TA_1.

4.2. Dual pH Sensors for Coastal Monitoring

The co-deployment of two distinct pH sensors (LOC and SeaFET) onboard the ALR provided value in two distinct ways. Prior to any correction, the independent agreement between the LOC and SeaFET sensors demonstrates consistency between fundamentally different measurement principles under the same environmental conditions. Subsequently, the Lab-on-Chip sensor was treated explicitly as a reference to correct SeaFET drift via k_0 adjustment, enabling a stable, high-frequency pH record. In this latter step, the sensors are no longer independent in order to generate a corrected, stable, high-frequency pH time series that is anchored to accurate photometric measurements. This deployment offered a rare opportunity to cross-examine sensor performance during long-haul seagoing conditions, something particularly important for autonomous platforms where ship-based validation is not possible, or where global models may under-perform. Despite operating on fundamentally different measurement principles, the two pH sensors showed stronger agreement with each other (pre-correction) than with model-based pH estimates. An important limitation is that the LDPT deployment lacked discrete sample validation. Therefore, while LOC–SeaFET agreement supports a stable and internally consistent pH record and strengthens confidence in resolved spatiotemporal variability, a small absolute offset in deployment-wide pH cannot be fully excluded.

Model-derived pH outputs (pH_{LIR}, pH_{NN}) were also assessed as alternatives to co-sensor calibration, particularly in scenarios where only a single sensor is available. However, while these models can effectively mitigate sensor drift, they remain constrained by their underlying assumptions and training data. Moreover, their performances tend to be better under the stable conditions typically found in the deep ocean, where platforms like the BGC-Argo array regularly recalibrate against known deepwater chemistry. In contrast, the heterogeneous and rapidly changing biogeochemical gradients of shallow shelf seas, which were mostly sampled during this mission, pose significant challenges to these models, often limiting their predictive skill and transferability. Coastal waters lack the quasi-conservative behavior seen in deep open ocean waters, and the input-output relationships captured by neural networks and linear regressions trained on deep data sets may not hold in these regimes. That two (fundamentally) different pH sensors converged more closely with each other than with model estimates is not only a case for redundancy, but also gives clear evidence that carbonate variability in shelf seas exceeds the resolving capacity of current models. Additionally, the pH estimate routines used in this work relied heavily on DO as a key input parameter, making their output inherently sensitive to the accuracy and calibration of the oxygen sensor itself. The reduced skill of globally trained data-driven models in dynamic coastal environments reinforces the motivation for the dual-sensor design and highlights the importance of direct in situ autonomous measurements for constraining the coastal marine carbonate system.

4.2.1. pH Comparisons

To quantify and assess the impact of the k_0 adjustment procedure, comparisons were made between sensor-measured pH_T, corrected pH_T, and model-estimated pH_T. Figure 12 represents the distribution of average $\Delta pH_T \pm$ first standard deviation (1σ) for each comparison. Both before and after correction, the two onboard pH

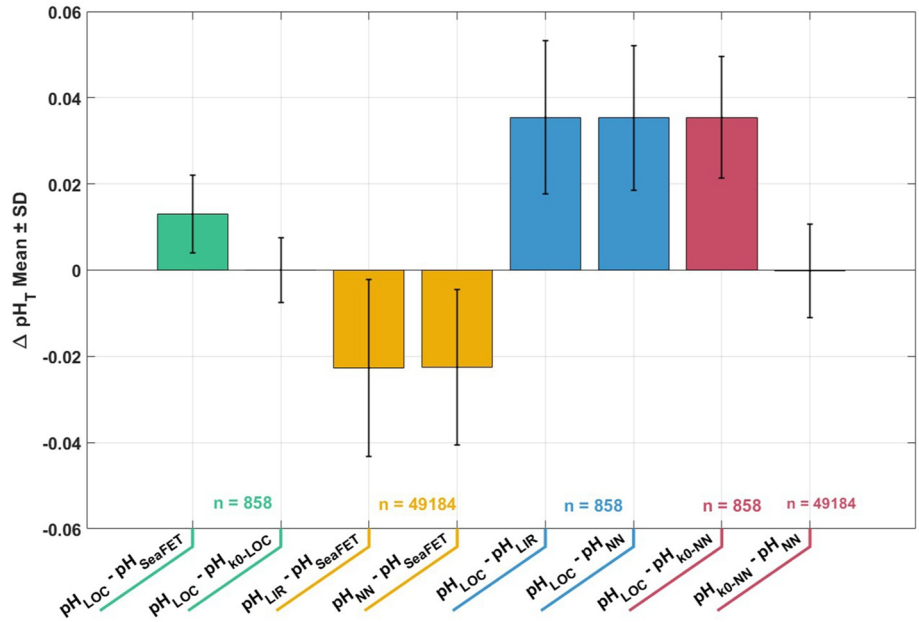

Figure 12. Bar chart with error bars of mean difference (± standard deviation) of pH_T (ΔpH_T) between different measurement methods and/or model estimations. Grouped by colors based on comparison parameters: between the two onboard sensors before and after k_0 correction based on LOC as pH_{ref} (green), original pH_{SeaFET} values compared to both output model estimates (yellow), pH_{LOC} values compared to both output model estimates (blue) and comparison of pH_{k0-NN} to pH_{LOC} and to pH_{NN} which was used as pH_{ref} to get pH_{k0-NN} (red). Sample size (n) for each comparison is shown. Error bars represent 1 standard deviation.

sensors showed the best agreement and lowest variability relative to ESPER estimates (LIR and NN). Interestingly, the uncorrected pH_{SeaFET} aligned more closely with both pH_{LIR} and pH_{NN} (both $\mu = -0.023$) than pH_{LOC} did (both $\mu = 0.035$), but this came with the highest variation of all comparisons ($1\sigma = 0.020/0.018$ vs. $1\sigma = 0.018/0.017$ for pH_{LIR} and pH_{NN} respectively) (Figure 12). The two ESPER-derived pH_T estimate data sets were nearly identical to each other, yielding indistinguishable mean results when compared to the same pH sensor measurement, although pH_{LIR} comparisons consistently showed a slightly greater variability of 0.002 higher standard deviation compared to pH_NN. Correction routines applied using both pH_{LOC} and pH_{NN} successfully adjusted the pH_{SeaFET} resulting in corrected values (pH_{LOC} and pH_{NN}) with rounded mean ΔpH_T values of 0. Notably, the distribution of the pH_{k0-NN} comparison has more variability ($1\sigma = 0.011$) than the pH_{k0-LOC} comparison ($1\sigma = 0.007$). The closer agreement between pH_{LOC} and pH_{SeaFET} compared to either versus the model-derived values reinforces that unresolved short-term and fine-scale processes dominate carbonate dynamics in coastal waters. These dynamics are poorly represented in models, which helps explain why shelf carbonate systems remain one of the least constrained components of the global carbon budget. Independent comparisons against GLODAP indicate larger near-surface discrepancies, reinforcing reduced NN skill in coastal waters (Text S3 and Figure S4 in Supporting Information S1). Yet, our findings are consistent with previous observations of fine-scale pH variability in shelf and coastal waters (Carstensen & Duarte, 2019; Sutton et al., 2019). Except for regionally trained neural networks, as demonstrated by Osborne et al. (2024), these rapid changes are poorly resolved and captured by empirical or neural-network approaches (e.g., Hauri et al., 2013). Similarly, sensor-based pH measurements from autonomous platforms (e.g.,Wimart-Rousseau et al., 2024) demonstrate higher fidelity to local dynamics than numerical or neural-network models alone. These studies reinforce the need for high-frequency, in situ observations to complement model-derived estimates, as reflected in our dual-sensor SeaFET measurements.

Across all comparisons, ΔpH_T indicated that the LOC sensor consistently measured a higher pH_T than the SeaFET and the ESPER estimates, and the SeaFET sensor, in turn, reported higher pH_T values than the ESPER models (Figure 12). These observations reinforce the value of in situ sensors, which captured fine-scale spatiotemporal pH variations more accurately than model-based estimates. Micro-scale pH patterns were detected by the onboard sensors but less reflected by model estimates, evident by the much better agreement

between the two sensors than either sensor with the models. This divergence highlights a key limitation of data-driven models in dynamic environments and speaks to the importance of direct, high-frequency measurements, especially in coastal monitoring or, when interpreting related parameters such as pCO_2 and CO_2 flux. Together, these comparisons demonstrate that autonomous dual-sensor observations do more than provide redundancy: they reveal heterogeneous and dynamic shelf carbonate systems, and demonstrate why models alone cannot yet deliver reliable constraints on coastal carbon fluxes.

4.2.2. Efficacy of the k_0 Correction Approach

A central aspect of this study was the adaptation and implementation of the k_0 correction to adjust the SeaFET pH sensor measurements using either a co-sensor (LOC pH sensor) or neural network-derived estimates (ESPER-NN) as the reference. In this context, pH_{ref} is used as an independent reference for anomaly-based k_0 corrections rather than an assumed "true" pH. pH_{LOC} is the favorable reference given its photometric, stable measurement with lower documented uncertainty, while pH_{NN} provides a community-accepted modeled estimate in the absence of a co-sensor. The correction method effectively addressed the initial electrode conditioning period typically experienced by the SeaFET sensors, and brought its measurements into near-perfect alignment with the corresponding pH_{ref}. However, comparisons also revealed that k_0-corrected SeaFET data based on pH_{NN} exhibited greater variability ($1\sigma = 0.011$) than corrections based on pH_{LOC} ($1\sigma = 0.007$). This difference stems from the noise and uncertainty associated with model-generated pH estimates, particularly in a shallow coastal setting. The k_0 correction was originally designed for deep-ocean profiling floats equipped with ISFET pH sensors, and using model estimates for pH_{ref}, thus benefiting from stable environmental conditions and better model performances at depth. Our work, however, shows this correction approach can be successfully applied to monitoring marine carbon in the coastal ocean when adapted to a co-sensor whose performance is more reliable than model output in such a setting. In the standard BGC-Argo implementation, model-derived pH serves as the only available k_0 reference, and is evaluated at depth where environmental variability is lower and model performance is more reliable (Maurer et al., 2021). This approach is less appropriate for shallow coastal deployments where model-derived pH may be less reliable (Figure S3, S4 and Text S3 in Supporting Information S1). Using the co-deployed LOC pH sensor as the k_0 adjustment reference offers a practical, traceable, and robust correction strategy for pH observation, and represents the type of advancements the carbon community is actively seeking for autonomous coastal deployments.

Although pH_{LOC} had far fewer data points than pH_{SeaFET}, this did not inhibit the correction's performance. The k_0 approach leveraged the complementary strengths of both onboard sensors by using the more power-intensive, high-precision measurements from the LOC sensor intermittently as "stabilizers" for the more frequent, lower-power measurements from the SeaFET. Beyond improving individual measurements, the k_0 correction enabled nearly 50,000 reliable pH observations throughout the full deployment, providing one of the most detailed continuous coastal carbonate records collected to date in this region. The resulting data set captures fine-scale spatiotemporal variability in pH, offering unprecedented opportunities to resolve the processes driving carbonate dynamics on short timescales but also potentially on biologically mediated processes, given the direct links between pH, marine carbon chemistry, and ecosystem function. For the deployment we describe here, the LOC sensor was operated at an hourly measurement frequency which may have been excessive based on the relatively stable performance of the SeaFET during the duration of the deployment. Although the behavior of individual electrochemical sensors cannot easily be predicted in the long term, a much less reference measurement frequency (i.e., once per day) could be adopted to mimic the correction frequency on profiling floats for example, to conserve endurance. This reinforces the advantage of co-deploying sensors for deriving corrections, while also demonstrating that neural network estimates can substitute for a second sensor when needed, albeit with slightly reduced precision. Taken together, this combination of dual sensors and k_0 adjustment of one sensor to the companion sensor output represents a step-change in observational capacity for coastal carbonate system research, particularly in dynamic shelf-sea environments where traditional ship-based monitoring cannot resolve rapid biogeochemical changes.

4.3. Air-Sea CO_2 Fluxes

Air-sea CO_2 flux estimates from the LDPT autonomous deployment indicate that during late spring of 2022, the Celtic Sea acted predominately as a net sink of atmospheric CO_2 into the ocean (Figure 10). These findings are consistent with previous studies also reporting the Celtic Sea as a net sink for atmospheric CO_2, which is common

for this type of marginal sea (Laruelle et al., 2014; Mathis et al., 2024). For example, Marrec et al. (2015) estimated seasonal air-sea CO_2 flux cycles in the Western English Channel using both observed and algorithm-derived pCO_2 values and found that the region acts a sink for atmospheric CO_2 during spring, with fluxes reaching -0.3 mol C m^{-2} month $^{-1}$, in agreement with our findings. Using a three-dimensional coupled biogeochemical-hydrodynamic model (MIRO-CO_2 & CO) implemented in the English Channel and the Southern Bight of the North Sea, Gypens et al. (2011) reported that this area acts as a sink for atmospheric CO_2 from December to June (-0.6 to -3.5 ± 4.6 mmol C m^{-2} d^{-1}), with spring fluxes ranging from -1.2 to -20.2 mmol C m^{-2} d^{-1} and substantially stronger uptake in nearshore waters. The spike in pCO_2 just before May 24 in Figures 8d–8f, coincident with the ALR's transition off the continental shelf (Figure 1), corresponds to a period of near air-sea equilibrium (CO_2 flux approaching zero) shown in Figure 10. This strongly corroborates the findings of Humphreys et al. (2019) who found the Celtic Margin to exhibit greater net air-sea CO_2 flux than the open ocean, aligning with the increasing strength of biogeochemical stratification that we also clearly observe in Figures 3 and 8 as our observations move away from the open ocean. Direct observational data of the carbonate system in the Celtic Sea remains scarce, with most prior flux estimates derived from models or algorithms (Hammermeister et al., 2025; Humphreys et al., 2019; Marrec et al., 2015; Wakelin et al., 2012). High-resolution data sets like the one presented here help fill this capability gap and anchor modeled flux magnitudes to direct in situ evidence, which is shown to be crucial for coastal carbon budget models.

Air-sea flux calculations are inherently sensitive to the accuracy of input carbonate chemistry parameters, particularly surface pCO_2, which depends strongly on the quality of pH data used to derive it when not measured directly. Despite robust correction strategies and convergence of pH data sets, calculated CO_2 fluxes still showed variability. Importantly, this variability sometimes resulted in apparent disagreements in flux direction. However, the propagated uncertainty analysis indicates that these events occur during near-equilibrium conditions, when ΔpCO_2 overlaps its uncertainty envelope and flux direction is therefore not always statistically resolvable (Figure S8 in Supporting Information S1).

As shown in Figure 9, flux values segregate more by pH data set than by TA, underscoring the critical role of pH observations for flux interpretation. Further complicating matters is the influence of wind speed in modulating air–sea exchange. A critical aspect of this sensitivity lies in the non-linear nature of carbonate system thermodynamics. CO_2 flux is computed from the gradient between surface water pCO_2 and atmospheric pCO_2, scaled by wind speed–dependent transfer velocity. Flux direction is therefore treated as resolved only when $\Delta pCO_2 > u(\Delta pCO_2)$, otherwise direction is interpreted as ambiguous within uncertainty (Figure S8 in Supporting Information S1). Because transfer velocity is multiplicative, even slight differences in calculated pCO_2 ($\pm$0.01–0.02 pH units) can result in disproportionately large changes in calculated flux, especially under high wind conditions.

Shown in Figure 10, the clearest example of the amplifying role of the wind occurs between June 7–14, when wind speeds reach up to $\sim$12 m s^{-1}. During this period, the flux estimates diverge substantially: fluxes based on pH_{k0-LOC} suggest the ocean remains a sink, absorbing nearly 15 mmol m^{-2} d^{-1} whereas those based on pH_{k0-NN} approach weak outgassing of nearly 8 mmol m^{-2} d^{-1}. However, the coincident ΔpCO_2 (calculated using pH_{k0-NN}) time series shows that this interval is characterized by near-equilibrium conditions, where ΔpCO_2 approaches zero and overlaps with the $u(\Delta pCO_2)$ value, so the sign difference should be interpreted as a period of sign ambiguity rather than an unequivocal switch from sink to source (Figure S8 in Supporting Information S1). Such directional sensitivity under near-equilibrium and high wind conditions has major implications, such as altering the interpretation of ocean–atmosphere carbon exchange during this period. This finding is fundamentally important for interpreting both regional and global flux estimates, as it suggests that some existing flux products may misclassify short-term sink/source transitions, particularly in coastal and shelf environments.

An additional factor influencing CO_2 flux determination is the actual depth at which "surface" carbonate parameters are measured. While this deployment sampled at just 0.4 m below the surface, many instrumented platforms contributing to SOCAT, such as SOOP, measure pCO_2 from depths of approximately 3–5 m. However, near-surface vertical gradients in pCO_2 and pH_T can develop rapidly due to stratification, diurnal heating, and biological activity, particularly in coastal or shallow systems (Ho & Schanze, 2020; Morgan et al., 2025). To quantify the possible effect of this sampling-depth mismatch, we applied a simple perturbation test to the baseline surface data set using imposed offsets of $\pm$0.01 and $\pm$0.02 pH units. These perturbations shifted mean calculated surface-water pCO_2 by -19.0 to $+19.9$ µatm and mean air-sea CO_2 flux by -1.0 to $+1.1$ (max values -4.6 to

+4.8) mmol m^{-2} d^{-1} relative to baseline, (See Table S4 and Figure S9 in Supporting Information S1). Although this test is illustrative rather than a direct reconstruction of the true vertical gradient during LDPT, it demonstrates that differences of only a few meters in sampling depth may introduce systematic offsets when comparing autonomous and ship-based CO_2 flux products, and challenges assumptions used in current global climatologies. This underscores the need for carbonate measurements as close to the air–sea interface as possible, because flux uncertainty depends not only on sensor precision, but on whether the sampled water represents the surface layer driving gas exchange.

Therefore, in autonomous deployments, where real-time corrections and discrete validation may be absent, sampling, selection, and correction of pH input becomes paramount. Conservative interpretation of flux results, particularly during high wind intervals, should account for these interactions. As Figure 10 shows, agreement among methods is best preserved during lower wind periods, suggesting that high-frequency wind monitoring is an essential co-variable in assessing confidence levels of derived flux products. As highlighted by Gandoin and Garza (2024) and Nickford et al. (2024), discrepancies between in situ wind observations and reanalysis products like ERA5 (used in this work) can introduce substantial uncertainty in air-sea CO_2 flux estimates, particularly in coastal and upwelling regions. Their study found that ERA5 tends to underestimate the highest wind speed events (>10 m s^{-1}), leading to systematic underestimations of CO_2 flux magnitude. These biases likely contribute to the amplified flux differences observed in our data under high winds. If ERA5 underestimates high wind events, then k_w (and therefore CO_2 flux) is biased low, dampening the magnitude of exchange. While this does not change the sign of ΔpCO_2, the added uncertainty $u(k_w)$ may mask flux amplitude, and mischaracterize air-sea carbon exchange dynamics.

The broader implication is that pH uncertainty cannot be considered in isolation when interpreting CO_2 flux. Instead it acts as a non-linear driver whose influence is modulated and amplified by environmental factors such as wind speed or fine-scale surface gradients which can transform small sensor or model disagreements into large differences in flux magnitude, and even direction. This dynamic interplay complicates efforts to identify sink/source transitions and may lead to contradictory assessments of the ocean's role in short-term carbon exchange, particularly in regions with variable weather and shallow bathymetry where mixed-layer dynamics respond quickly to physical forcing. Taken together, these findings underscore why high-resolution autonomous carbonate observations are essential for robust quantification of air–sea CO_2 fluxes and downstream products like pCO_2, especially in dynamic coastal seas that remain a major source of uncertainty in the global carbon cycle. By capturing fine-scale temporal and vertical variability, autonomous platforms provide critical insight into processes that ship-based surveys and coarse models cannot resolve.

5. Conclusion

This mission is a first of its kind deployment. The LDPT is the first fully autonomous deployment of the ALR equipped with scientific biogeochemical sensors—including dual pH instrumentation—to perform continuous, long-duration, and long-distance carbonate system monitoring. The deployment achieved the longest autonomous pH acquisition to date by the ALR and generated nearly 50,000 reliable pH observations, providing one of the most detailed coastal carbonate data sets collected in this region. Crucially, the deployment demonstrated the success of dual sensor operation, which not only improved data quality but enabled high-frequency capture of fine-scale spatiotemporal variability. These data reveal how tightly coupled biological and physical processes drive carbonate chemistry in shelf seas, capturing seasonal transitions and vertical gradients that are typically missed by ship-based surveys. Small differences in pH were shown to significantly affect calculated CO_2 fluxes, including reversals in flux direction under high wind conditions, emphasizing that coarse models may misclassify source or sink status in dynamic coastal systems. Future work could pair these high-resolution observations with satellite chlorophyll and regional hydrography to contextualize bloom timing and water-mass variability, and extend the air–sea exchange analysis by comparing CO_2 and O_2 fluxes and refining regional T–S relationships.

Altogether, this confirms the value of our approach as a powerful tool for determining air-sea gas fluxes, with clear potential applications beyond CO_2. Overall, this mission highlights the central role of autonomous, high-resolution observations in constraining short-term carbon dynamics, air–sea gas fluxes, and broader coastal carbon cycle processes. This study paves the way for future high-resolution biogeochemical monitoring as well as for long-endurance deployments to fill critical observational gaps, improving our understanding of OA, carbon

cycling, and air-sea exchange, and providing essential data to support climate modeling that contribute to scientific policy making and mitigation efforts.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

All sensor data are available from The University of Southampton Data Repository via DOI: 10.5258/SOTON/D3716.

Acknowledgments
The authors would like to thank the Marine Autonomous Robotic Systems team and others that made the LDPT possible with their tireless commitment to a successful shore vehicle launch and recovery, mission piloting, and any and all engineering, mechanical, or software needs. The development and LDPT of the ALR1500 was supported by the Industrial Strategy Challenge Fund/Natural Environment Research Council's (ISCF/NERC's) Oceanids Programme and NERC National Underpinning Autonomy Capability funding and through the INSPIRE DTP. We would also like to thank Dr. Mingxi Yang and the Penlee Point Atmospheric Observatory at Plymouth Marine Laboratory for providing atmospheric CO_2 data which were supported by the NERC ACSIS (Grant NE/N018044/1) and MOYA (Grant NE/N015932/1) projects. Finally, as always, the authors thank Swift (D.F.A), Director B. Leyert, Donna, Tim, Wika, and Ruthie B.

References

Archibald, A. T., Sinha, B., Russo, M. R., Matthews, E., Squires, F. A., Abraham, N. L., et al. (2025). Data supporting the north atlantic climate system integrated study (ACSIS) programme, including atmospheric composition; oceanographic and sea-ice observations (2016–2022); and output from ocean, atmosphere, land, and sea-ice models (1950–2050). *Earth System Science Data*, *17*(1), 135–164. https://doi.org/10.5194/essd-17-135-2025

Bakker, D. C. E., Alin, S. R., Aramaki, T., Barbero, L., Bates, N., Gkritzalis, T., et al. (2025). Surface ocean co2 atlas database version 2025 (socatv2025). *NOAA National Centers for Environmental Information*. https://doi.org/10.25921/648f-fv35

Barker, S., & Ridgwell, A. (2012). Ocean acidifcation. *Nature Education Knowledge*, *3*.

Bittig, H. C., Körtzinger, A., Neill, C., van Ooijen, E., Plant, J. N., Hahn, J., et al. (2018). Oxygen optode sensors: Principle, characterization, calibration, and application in the ocean. *Frontiers in Marine Science*, *4*, 429. https://doi.org/10.3389/fmars.2017.00429

Borges, A. V., & Frankignoulle, M. (2003). Distribution of surface carbon dioxide and air-sea exchange in the English Channel and adjacent areas. *Journal of Geophysical Research*, *108*(C5), 3140. https://doi.org/10.1029/2000JC000571

Bresnahan, P. J., Martz, T. R., Takeshita, Y., Johnson, K. S., & LaShomb, M. (2014). Best practices for autonomous measurement of seawater pH with the Honeywell Durafet. *Methods in Oceanography*, *9*, 44–60. https://doi.org/10.1016/j.mio.2014.08.003

Cai, W. J. (2011). Estuarine and coastal ocean carbon paradox: Co 2 sinks or sites of terrestrial carbon incineration? *Annual Review of Marine Science*, *3*(1), 123–145. https://doi.org/10.1146/annurev-marine-120709-142723

Cai, W.-J., Hu, X., Huang, W.-J., Jiang, L.-Q., Wang, Y., Peng, T.-H., & Zhang, X. (2010). Alkalinity distribution in the western North Atlantic Ocean margins. *Journal of Geophysical Research*, *115*(C8). https://doi.org/10.1029/2009JC005482

Caldeira, K., & Wickett, M. E. (2003). Anthropogenic carbon and ocean pH. *Nature*, *425*(6956), 365. https://doi.org/10.1038/425365a

Carstensen, J., & Duarte, C. M. (2019). Drivers of ph variability in coastal ecosystems. *Environmental Science & Technology*, *53*(8), 4020–4029. https://doi.org/10.1021/acs.est.8b03655

Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y.-Y., et al. (2021). New and updated global empirical seawater property estimation routines. *Limnology and Oceanography: Methods*, *19*(12), 785–809. https://doi.org/10.1002/lom3.10461

Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., & Takeshita, Y. (2018). Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. *Limnology and Oceanography: Methods*, *16*(2), 119–131. https://doi.org/10.1002/lom3.10232

Chen, C.-T. A., Huang, T.-H., Chen, Y.-C., Bai, Y., He, X., & Kang, Y. (2013). Air–sea exchanges of co2 in the world's coastal seas. *Biogeosciences*, *10*, 6509–6544. https://doi.org/10.5194/bg-10-6509-2013

Claustre, H., Johnson, K. S., & Takeshita, Y. (2020). Observing the global ocean with Biogeochemical-Argo. *Annual Review of Marine Science*, *12*, 23–48. https://doi.org/10.1146/ANNUREV-MARINE-010419-010956/CITE/REFWORKS

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, *109*(C12), 1–20. https://doi.org/10.1029/2004JC002378

Dias, L. M., & Carter, B. R. (2025). Pyesperv1.0.0: A python implementation of empirical seawater property estimation routines (espers). *Geoscientific Model Development*, *18*(20), 7275–7295. https://doi.org/10.5194/gmd-18-7275-2025

Dickson, A., Chris, S., & Christian, J. R. (2007). *Guide to best practices for ocean CO2 measurements* (Vol. 3). Guide to Best Practices for Ocean CO2 Measurements.

Dickson, A. G. (1990). Standard potential of the reaction: $AgCl(s) + 12H_2(g) = Ag(s) + HCl(aq)$, and and the standard acidity constant of the ion HSO_4^- in synthetic sea water from 273.15 to 318.15 K. *The Journal of Chemical Thermodynamics*, *22*(2), 113–127. https://doi.org/10.1016/0021-9614(90)90074-Z

Duarte, C. M., Hendriks, I. E., Moore, T. S., Olsen, Y. S., Steckbauer, A., Ramajo, L., et al. (2013). Is ocean acidification an open-ocean syndrome? Understanding anthropogenic impacts on seawater pH. *Estuaries and Coasts*, *36*(2), 221–236. https://doi.org/10.1007/s12237-013-9594-3

Ellison, S. L. R. & Williams, A. (Eds.) (2012)., *Eurachem/citac guide: Quantifying uncertainty in analytical measurement* (3rd ed.). Eurachem. Retrieved from www.eurachem.org

Feely, R. A., Sabine, C. L., Lee, K., Berelson, W., Kleypas, J., Fabry, V. J., & Millero, F. J. (2004). Impact of anthropogenic CO2 on the CaCO3 system in the oceans. *Science*, *305*(5682), 362–366. https://doi.org/10.1126/science.1097329

Ford, D. J., Blannin, J., Watts, J., Watson, A. J., Landschützer, P., Jersild, A., & Shutler, J. D. (2024). A comprehensive analysis of air-sea co2 flux uncertainties constructed from surface ocean data products. *Global Biogeochemical Cycles*, *38*(11), e2024GB008188. https://doi.org/10.1029/2024GB008188

Frankignoulle, M., & Borges, A. V. (2001). European continental shelf as a significant sink for atmospheric carbon dioxide. *Global Biogeochemical Cycles*, *15*(3), 569–576. https://doi.org/10.1029/2000GB001307

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., et al. (2025). Global carbon budget 2024. *Earth System Science Data*, *17*(3), 965–1039. https://doi.org/10.5194/essd-17-965-2025

Gandoin, R., & Garza, J. (2024). Underestimation of strong wind speeds offshore in ERA5: Evidence, discussion and correction. *Wind Energ. Sci.*, *9*(8), 1727–1745. https://doi.org/10.5194/wes-9-1727-2024

Gattuso, J.-P., Allemand, D., & Frankignoulle, M. (1999). Photosynthesis and calcification at cellular, organismal and community levels in coral reefs: A review on interactions and control by carbonate chemistry. *American Zoologist*, *39*(1), 160–183. https://doi.org/10.1093/icb/39.1.160

Gonski, S. F., Cai, W.-J., Ullman, W. J., Joesoef, A., Main, C. R., Pettay, D. T., & Martz, T. R. (2018). Assessment of the suitability of durafet-based sensors for ph measurement in dynamic estuarine environments. *Estuarine, Coastal and Shelf Science*, *200*, 152–168. https://doi.org/10.1016/j.ecss.2017.10.020

Gypens, N., Lacroix, G., Lancelot, C., & Borges, A. V. (2011). Seasonal and inter-annual variability of air–sea co2 fluxes and seawater carbonate chemistry in the southern north sea. *Progress in Oceanography*, *88*(1–4), 59–77. https://doi.org/10.1016/j.pocean.2010.11.004

Hammermeister, E. M., Papadimitriou, S., Arundell, M., Ludgate, J., Schaap, A., Mowlem, M. C., et al. (2025). New capability in autonomous ocean carbon observations using the autosub long-range AUV equipped with novel pH and total alkalinity sensors. *Environmental Science & Technology*, *59*(14), 7129–7144. https://doi.org/10.1021/acs.est.4c10139

Hartman, S. E., Humphreys, M. P., Kivimäe, C., Woodward, E. M. S., Kitidis, V., Mcgrath, T., et al. (2014). Seasonality and spatial heterogeneity of the surface ocean carbonate system in the northwest european continental shelf. *Progress in Oceanography*, *177*, 101909. https://doi.org/10.1016/j.pocean.2018.02.005

Hauri, C., Gruber, N., Vogt, M., Doney, S. C., Feely, R. A., Lachkar, Z., et al. (2013). Spatiotemporal variability and long-term trends of ocean acidification in the california current system. *Biogeosciences*, *10*(1), 193–216. https://doi.org/10.5194/bg-10-193-2013

Hauri, C., Irving, B., Hayes, D., Abdi, E., Kemme, J., Kinski, N., & McDonnell, A. M. P. (2024). Expanding seawater carbon dioxide and methane measuring capabilities with a seaglider. *Ocean Science*, *20*(5), 1403–1421. https://doi.org/10.5194/os-20-1403-2024

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Ho, D. T., & Schanze, J. J. (2020). Precipitation-induced reduction in surface ocean pCO2: Observations from the eastern tropical Pacific ocean. *Geophysical Research Letters*, *47*(15), e2020GL088252. https://doi.org/10.1029/2020GL088252

Humphreys, M. P., Achterberg, E. P., Hopkins, J. E., Chowdhury, M. Z. H., Griffiths, A. M., Hartman, S. E., et al. (2019). Mechanisms for a nutrient-conserving carbon pump in a seasonally stratified, temperate continental shelf sea. *Progress in Oceanography*, *177*, 101961. https://doi.org/10.1016/j.pocean.2018.05.001

JCGM, BIPM, IEC, IFCC, ILAC, ISO, & OIMIL. (2008). Evaluation of measurement data-guide to the expression of uncertainty in measurement. *JCGM*, *100*, 2008. Retrieved from www.bipm.org

Jiang, Z.-P., Tyrrell, T., Hydes, D. J., Dai, M., & Hartman, S. E. (2014). Variability of alkalinity and the alkalinity-salinity relationship in the tropical and subtropical surface ocean. *Global Biogeochemical Cycles*, *28*(7), 729–742. https://doi.org/10.1002/2013GB004678

Johnson, K. S., Jannasch, H. W., Coletti, L. J., Elrod, V. A., Martz, T. R., Takeshita, Y., et al. (2016). Deep-sea DuraFET: A pressure tolerant pH sensor designed for global sensor networks. *Analytical Chemistry*, *88*(6), 3249–3256. https://doi.org/10.1021/acs.analchem.5b04653

Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., et al. (2017). Biogeochemical sensor performance in the SOCCOM profiling float array. *Journal of Geophysical Research: Oceans*, *122*(8), 6416–6436. https://doi.org/10.1002/2017JC012838

Joint, I., Wollast, R., Chou, L., Batten, S., Elskens, M., Edwards, E., et al. (2001). Pelagic production at the Celtic Sea shelf break. *Deep Sea Research Part II: Topical Studies in Oceanography*, *48*(14–15), 3049–3081. https://doi.org/10.1016/S0967-0645(01)00032-7

Khatiwala, S., Tanhua, T., Fletcher, S. M., Gerber, M., Doney, S. C., Graven, H. D., et al. (2013). Global ocean storage of anthropogenic carbon. *Biogeosciences*, *10*(4), 2169–2191. https://doi.org/10.5194/bg-10-2169-2013

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Kitidis, V., Hardman-Mountford, N. J., Litt, E., Brown, I., Cummings, D., Hartman, S., et al. (2012). Seasonal dynamics of the carbonate system in the Western English Channel. https://doi.org/10.1016/j.csr.2012.04.012

Laruelle, G. G., Lauerwald, R., Pfeil, B., & Regnier, P. (2014). Regionalized global budget of the co2 exchange at the air-water interface in continental shelf seas. *Global Biogeochemical Cycles*, *28*(11), 1199–1214. https://doi.org/10.1002/2014gb004832

Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., et al. (2024). The annual update glodapv2.2023: The global interior ocean biogeochemical data product. *Earth System Science Data*, *16*(4), 2047–2072. https://doi.org/10.5194/essd-16-2047-2024

Lee, K., Kim, T.-W., Byrne, R. H., Millero, F. J., Feely, R. A., & Liu, Y.-M. (2010). The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. *Geochimica et Cosmochimica Acta*, *74*(6), 1801–1811. https://doi.org/10.1016/j.gca.2009.12.027

Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., et al. (2006). Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. *Geophysical Research Letters*, *33*(19). https://doi.org/10.1029/2006GL027207

Levitus, S. (1982). *Climatological atlas of the world ocean* (Vol. 13). US Department of Commerce, National Oceanic and Atmospheric Administration.

Lewis, E. R., & Wallace, D. W. R. (1998). Program developed for CO2 system calculations. https://doi.org/10.15485/1464255

Lueker, T. J., Dickson, A. G., & Keeling, C. D. (2000). Ocean pCO2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: Validation based on laboratory measurements of CO2 in gas and seawater at equilibrium. *Marine Chemistry*, *70*(1–3), 105–119. https://doi.org/10.1016/S0304-4203(00)00022-0

Marrec, P., Cariou, T., Macé, E., Morin, P., Salt, L. A., Vernet, M., et al. (2015). Dynamics of air–sea CO2 fluxes in the northwestern European shelf based on voluntary observing ship and satellite observations. *Biogeosciences*, *12*(18), 5371–5391. https://doi.org/10.5194/bg-12-5371-2015

Martz, T. R., Connery, J. G., & Johnson, K. S. (2010). Testing the Honeywell durafet® for seawater pH applications. *Limnology and Oceanography: Methods*, *8*(5), 172–184. https://doi.org/10.4319/lom.2010.8.172

Martz, T. R., Daly, K. L., Byrne, R. H., Stillman, J. H., & Turk, D. (2015). Technology for ocean acidification research: Needs and availability. *Oceanography*, *28*(2), 40–47. https://doi.org/10.5670/oceanog.2015.30

Mathis, M., Lacroix, F., Hagemann, S., Nielsen, D. M., Ilyina, T., & Schrum, C. (2024). Enhanced co2 uptake of the coastal ocean is dominated by biological carbon fixation. *Nature Climate Change*, *14*(4), 373–379. https://doi.org/10.1038/s41558-024-01956-w

Maurer, T. L., Plant, J. N., & Johnson, K. S. (2021). Delayed-mode quality control of oxygen, nitrate, and pH data on SOCCOM biogeochemical profiling floats. *Frontiers in Marine Science*, *8*, 683207. https://doi.org/10.3389/fmars.2021.683207

McKinley, G. A., Pilcher, D. J., Fay, A. R., Lindsay, K., Long, M. C., & Lovenduski, N. S. (2016). Timescales for detection of trends in the ocean carbon sink. *Nature*, *530*(7591), 469–472. https://doi.org/10.1038/nature16958

McLaughlin, K., Dickson, A., Weisberg, S. B., Coale, K., Elrod, V., Hunter, C., et al. (2017). An evaluation of isfet sensors for coastal ph monitoring applications. *Regional Studies in Marine Science*, *12*, 11–18. https://doi.org/10.1016/j.rsma.2017.02.008

Miller, C. A., Pocock, K., Evans, W., & Kelley, A. L. (2018). An evaluation of the performance of sea-bird scientific's seafet™ autonomous ph sensor: Considerations for the broader oceanographic community. *Ocean Science*, *14*(4), 751–768. https://doi.org/10.5194/os-14-751-2018

Millero, F. J. (1995). Thermodynamics of the carbon dioxide system in the oceans. *Geochimica et Cosmochimica Acta*, *59*(4), 661–677. https://doi.org/10.1016/0016-7037(94)00354-O

Millero, F. J. (2007). The marine inorganic carbon cycle. *Chemical Reviews*, *107*(2), 308–341. https://doi.org/10.1021/CR0503557

Millero, F. J., Lee, K., & Roche, M. (1998). Distribution of alkalinity in the surface waters of the major oceans. *Marine Chemistry*, *60*(1–2), 111–130. https://doi.org/10.1016/S0304-4203(97)00084-4

Monterey, G. I., & Levitus, S. (1997). *Seasonal variability of mixed layer depth for the world ocean* (Vol. 14). NOAA atlas NESDIS.

Morgan, S., Wong, S., Byrne, T., Comeau, A., Ward, B., Barry, M., & Atamanchuk, D. (2025). Wave glider-based measurements and corrections of near-surface pCO2 gradients in the coastal ocean. *Global Biogeochemical Cycles*, *39*(7), e2024GB008396. https://doi.org/10.1029/2024GB008396

Mosley, L. M., Husheer, S. L. G., & Hunter, K. A. (2004). Spectrophotometric ph measurement in estuaries using thymol blue and m-cresol purple. *Marine Chemistry*, *91*(1–4), 175–186. https://doi.org/10.1016/j.marchem.2004.06.008

Müller, J. D., & Rehder, G. (2018). Metrology of pH measurements in brackish waters—part 2: Experimental characterization of purified meta-cresol purple for spectrophotometric pHT measurements. *Frontiers in Marine Science*, *5*, 177. https://doi.org/10.3389/fmars.2018.00177

Newton, J., Feely, R., Jewett, E., Williamson, P., & Mathis, J. (2015). *Global ocean acidification observing network: Requirements and governance plan*. UNESCO IOC. Retrieved from https://research-portal.uea.ac.uk/en/publications/global-ocean-acidification-observing-network-requirements-and-gov

Nickford, S., Palter, J. B., & Mu, L. (2024). The importance of contemporaneous wind and pCO2 measurements for regional air-sea CO2 flux estimates. *Journal of Geophysical Research: Oceans*, *129*(6), e2023JC020744. https://doi.org/10.1029/2023JC020744

Olsen, A., Key, R. M., Lauvset, S. K., Kozyr, A., Tanhua, T., Hoppema, M., et al. (2017). *Global ocean data analysis project, version 2 (GLODAPv2) (NCEI accession 0162565)*. NOAA National Centers for Environmental Information.

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2016). The global ocean data analysis project version 2 (GLODAPv2) – An internally consistent data product for the world ocean. *Earth System Science Data*, *8*(2), 297–323. https://doi.org/10.5194/essd-8-297-2016

Orr, J. C., Epitalon, J. M., Dickson, A. G., & Gattuso, J. P. (2018). Routine uncertainty propagation for the marine carbon dioxide system. *Marine Chemistry*, *207*, 84–107. https://doi.org/10.1016/j.marchem.2018.10.006

Osborne, E., Xu, Y.-Y., Soden, M., McWhorter, J., Barbero, L., & Wanninkhof, R. (2024). A neural network algorithm for quantifying seawater ph using biogeochemical-argo floats in the open gulf of Mexico. *Frontiers in Marine Science*, *11*, 1468909. https://doi.org/10.3389/fmars.2024.1468909

Park, P. K. (1969). Oceanic CO2 system: An evaluation of ten methods of investigation. *Limnology and Oceanography*, *14*(2), 179–186. https://doi.org/10.4319/lo.1969.14.2.0179

Perez, F. F., & Fraga, F. (1987). Association constant of fluoride and hydrogen ions in seawater. *Marine Chemistry*, *21*(2), 161–168. https://doi.org/10.1016/0304-4203(87)90036-3

Phillips, A. B., Templeton, R., Roper, D., Morrison, R., Pebody, M., Bagley, P. M., et al. (2023). Autosub Long Range 1500: A continuous 2000 km field trial. *Ocean Engineering*, *280*, 114626. https://doi.org/10.1016/j.oceaneng.2023.114626

Pickup, D. D., Bakker, D. C. E., Heywood, K. J., Glassup, F., Hammermeister, E., Stammerjohn, S. E., et al. (2025). Cold lenses in the amundsen sea: Impacts of sea ice formation on subsurface ph and carbon. *EGUsphere*, 1–23. https://doi.org/10.5194/egusphere-2025-2441

Plymouth Marine Laboratory, & Yang, M. (2017). Penlee point atmospheric observatory: Meteorological and chemical observations 2014–present. *Centre for Environmental Data Analysis*. Retrieved from https://catalogue.ceda.ac.uk/uuid/8f1ff8ea77534e08b03983685990a9b0. Accessed: 7 November 2024.

Possenti, L., Humphreys, M. P., Bakker, D. C. E., Cobas-García, M., Fernand, L., Lee, G. A., et al. (2021). Air-sea gas fluxes and remineralization from a novel combination of pH and O2 sensors on a glider. *Frontiers in Marine Science*, *8*, 696772. https://doi.org/10.3389/fmars.2021.696772

Quéré, C. L., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., et al. (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, *2*(12), 831–836. https://doi.org/10.1038/ngeo689

Raimondi, L., Matthews, J. B. R., Atamanchuk, D., Azetsu-Scott, K., & Wallace, D. W. R. (2019). The internal consistency of the marine carbon dioxide system for high latitude shipboard and in situ monitoring. *Marine Chemistry*, *213*, 49–70. https://doi.org/10.1016/j.marchem.2019.03.001

Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F. T., Gruber, N., Janssens, I. A., et al. (2013). Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nature Geoscience*, *6*(8), 597–607. https://doi.org/10.1038/ngeo1830

Rérolle, V., Ruiz-Pino, D., Rafizadeh, M., Loucaides, S., Papadimitriou, S., Mowlem, M., & Chen, J. (2016). Measuring pH in the Arctic ocean: Colorimetric method or SeaFET? *Methods in Oceanography*, *17*, 32–49. https://doi.org/10.1016/j.mio.2016.05.006

Rérolle, V. M. C., Floquet, C. F. A., Mowlem, M. C., Connelly, D. P., Achterberg, E. P., & Bellerby, R. R. G. J. (2012). Seawater-pH measurements for ocean-acidification observations. *TrAC, Trends in Analytical Chemistry*, *40*, 146–157. https://doi.org/10.1016/j.trac.2012.07.016

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., et al. (2019). On the future of Argo: A global, full-depth, multi-disciplinary array. *Frontiers in Marine Science*, *6*, 439. https://doi.org/10.3389/fmars.2019.00439

Roobaert, A., Laruelle, G. G., Landschützer, P., Gruber, N., Chou, L., & Regnier, P. (2019). The spatiotemporal dynamics of the sources and sinks of CO2 in the global coastal ocean. *Global Biogeochemical Cycles*, *33*(12), 1693–1714. https://doi.org/10.1029/2019GB006239

Roper, D., Harris, C. A., Salavasidis, G., Pebody, M., Templeton, R., Prampart, T., et al. (2021). Autosub Long Range 6000: A multiple-month endurance AUV for deep-ocean monitoring and survey. *IEEE Journal of Oceanic Engineering*, *46*(4), 1179–1191. https://doi.org/10.1109/JOE.2021.3058416

Saba, G. K., Wright-Fairbanks, E., Chen, B., Cai, W.-J., Barnard, A. H., Jones, C. P., et al. (2019). The development and validation of a profiling glider deep ISFET-based pH sensor for high resolution observations of coastal and ocean acidification. *Frontiers in Marine Science*, *6*, 664. https://doi.org/10.3389/fmars.2019.00664

Sabine, C., Sutton, A., McCabe, K., Lawrence-Slavas, N., Alin, S., Feely, R., et al. (2020). Evaluation of a new carbon dioxide system for autonomous surface vehicles. *undefined*, *37*(8), 1305–1317. https://doi.org/10.1175/JTECH-D-20-0010.1

Sauzède, R., Bittig, H. C., Claustre, H., de Fommervault, O. P., Gattuso, J.-P., Legendre, L., & Johnson, K. S. (2017). Estimates of water-column nutrient concentrations and carbonate system parameters in the global ocean: A novel approach based on neural networks. *Frontiers in Marine Science*, *4*, 128. https://doi.org/10.3389/fmars.2017.00128

Schuster, U., McKinley, G. A., Bates, N., Chevallier, F., Doney, S. C., Fay, A. R., et al. (2013). An assessment of the atlantic and arctic sea–air CO2 fluxes, 1990–2009. *Biogeosciences*, *10*(1), 607–627. https://doi.org/10.5194/bg-10-607-2013

Scientific, S.-B. (2019). Isfet ph sensors: Internal and external references. Retrieved from https://www.seabird.com/news/seafet-v2-internal-external-comparison

Sea Bird Scientific. (2023). Calculating pH from ISFET pH sensors (Tech. Rep. No. Application Note 99). In *Sea-Bird scientific*. Retrieved from https://242648281.fs1.hubspotusercontent-na2.net/hubfs/242648281/Resources/Application%20Notes/AN99-CalculatepH.pdf

Seguro, I., Marca, A. D., Painting, S. J., Shutler, J. D., Suggett, D. J., & Kaiser, J. (2019). High-resolution net and gross biological production during a celtic sea spring bloom. *Progress in Oceanography*, *177*, 101885. https://doi.org/10.1016/j.pocean.2017.12.003

Sharp, J. D., Pierrot, D., Humphreys, M. P., Epitalon, J.-M., Orr, J. C., Lewis, E. R., & Wallace, D. W. (2023a). Co2sysv3 for matlab. *Zenodo*. https://doi.org/10.5281/zenodo.7552554

Sharp, J. D., Pierrot, D., Humphreys, M. P., Epitalon, J.-M., Orr, J. C., Lewis, E. R., & Wallace, D. W. (2023b). CO2SYSv3 for MATLAB. https://doi.org/10.5281/ZENODO.7552554

Sutton, A. J., Feely, R. A., Maenner-Jones, S., Musielwicz, S., Osborne, J., Dietrich, C., et al. (2019). Autonomous seawater pco2 and ph time series from 40 surface buoys and the emergence of anthropogenic trends. *Earth System Science Data*, *11*(1), 421–439. https://doi.org/10.5194/essd-11-421-2019

Sutton, A. J., & Sabine, C. L. (2023). Emerging applications of longstanding autonomous ocean carbon observations. *Oceanography*, *36*, 148–155. https://doi.org/10.5670/OCEANOG.2023.209

Takahashi, T., Sutherland, S. C., Chipman, D. W., Goddard, J. G., Ho, C., Newberger, T., et al. (2014). Climatological distributions of pH, pCO2, total CO2, alkalinity, and CaCO3 saturation in the global surface ocean, and temporal changes at selected locations. *Marine Chemistry*, *164*, 95–125. https://doi.org/10.1016/j.marchem.2014.06.004

Tanhua, T., McCurdy, A., Fischer, A., Appeltans, W., Bax, N., Currie, K., et al. (2019). What we have learned from the framework for ocean observing: Evolution of the global ocean observing system. *Frontiers in Marine Science*, *6*, 436048. https://doi.org/10.3389/fmars.2019.00471

Tans, P. (2009). An accounting of the observed increase in oceanic and atmospheric CO2 and the outlook for the future. *undefined*, *22*(4), 26–35. https://doi.org/10.5670/OCEANOG.2009.94

van Heuven, S., Pierrot, D., Rae, J., Lewis, E., & Wallace, D. W. R. (2011). CO2SYS v 1.1, matlab program developed for CO2 system calculations.

Wakelin, S. L., Holt, J. T., Blackford, J. C., Allen, J. I., Butenschön, M., & Artioli, Y. (2012). Modeling the carbon fluxes of the northwest european continental shelf: Validation and budgets. *Journal of Geophysical Research*, *117*(C5). https://doi.org/10.1029/2011JC007402

Wanninkhof, R. (2014). 6). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography: Methods*, *12*(6), 351–362. https://doi.org/10.4319/lom.2014.12.351

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., et al. (2013). Global ocean carbon uptake: Magnitude, variability and trends. *Biogeosciences*, *10*(3), 1983–2000. https://doi.org/10.5194/bg-10-1983-2013

Weiss, R. F. (1974). Carbon dioxide in water and seawater: The solubility of a non-ideal gas. *Marine Chemistry*, *2*(3), 203–215. https://doi.org/10.1016/0304-4203(74)90015-2

Williams, N. L., Juranek, L. W., Feely, R. A., Johnson, K. S., Sarmiento, J. L., Talley, L. D., et al. (2017). Calculating surface ocean pCO2 from biogeochemical argo floats equipped with pH: An uncertainty analysis. *Global Biogeochemical Cycles*, *31*(3), 591–604. https://doi.org/10.1002/2016GB005541

Wimart-Rousseau, C., Steinhoff, T., Klein, B., Bittig, H., & Körtzinger, A. (2024). Technical note: Assessment of float ph data quality control methods – A case study in the subpolar northwest atlantic ocean. *Biogeosciences*, *21*(5), 1191–1211. https://doi.org/10.5194/bg-21-1191-2024

Wolff, G. (2007). Sarmiento, j. l. & gruber, n. 2006. ocean biogeochemical dynamics.xiii + 503 pp. princeton, woodstock: Princeton university press. price £48.95 (hard covers). 0 691 01707 7. *Geological Magazine*, *144*(6), 1034. https://doi.org/10.1017/S0016756807003755

Yang, M., Bell, T. G., Brown, I. J., Fishwick, J. R., Kitidis, V., Nightingale, P. D., et al. (2019). Insights from year-long measurements of air–water CH_4 and CO_2 exchange in a coastal environment. *Biogeosciences*, *16*(5), 961–978. https://doi.org/10.5194/bg-16-961-2019

Yin, T., Papadimitriou, S., Rérolle, V. M. C., Arundell, M., Cardwell, C. L., Walk, J., et al. (2021). 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*(21), 14968–14978. https://doi.org/10.1021/acs.est.1c03517

References From the Supporting Information

Johnson, K. S., Plant, J. N., Maurer, T. L., & Takeshita, Y. (2023). *Processing BGC-Argo ph data at the DAC level*. Ifremer.

Khoo, K. H., Ramette, R. W., Culberson, C. H., & Bates, R. G. (1977). Determination of hydrogen ion concentrations in seawater from 5 to 40. degree.C: Standard potentials at salinities from 20 to 45%. *Analytical Chemistry*, *49*(1), 29–34. https://doi.org/10.1021/ac50009a016

Liu, X., Patsavas, M. C., & Byrne, R. H. (2011). Purification and characterization of meta-cresol purple for spectrophotometric seawater ph measurements. *Environmental Science and Technology*, *45*(11), 4862–4868. https://doi.org/10.1021/es200665d

Millero, F. J. (1982). The effect of pressure on the solubility of minerals in water and seawater. *Geochimica et Cosmochimica Acta*, *46*(1), 11–22. https://doi.org/10.1016/0016-7037(82)90286-1

Millero, F. J. (1983). The estimation of the pkha of acids in seawater using the pitzer equations. *Geochimica et Cosmochimica Acta*, *47*(12), 2121–2129. https://doi.org/10.1016/0016-7037(83)90037-6