Check for updates

Long-term intercomparison of two pCO_2 instruments based on ship-of-opportunity measurements in a dynamic shelf sea environment

Vlad A. Macovei [1]*, Yoana G. Voynova [1], Meike Becker [2,3] Jack Triest,[4,5] Wilhelm Petersen[1]

[1]Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Geesthacht, Germany
[2]Geophysical Institute, University of Bergen, Bergen, Norway
[3]Bjerknes Centre for Climate Research, Bergen, Norway
[4]Kongsberg Maritime Contros GmbH, Kiel, Germany
[5]Present address: 4H-JENA Engineering GmbH, Kiel, Germany

Abstract

The partial pressure of carbon dioxide (pCO_2) in surface seawater is an important biogeochemical variable because, together with the pCO_2 in the atmosphere, it determines the direction of air–sea carbon dioxide exchange. Large-scale observations of pCO_2 are facilitated by Ships-of-Opportunity (SOOP-CO$_2$) equipped with underway measuring instruments. The need for expanding the observation capacity and the challenges involving the sustainability and maintenance of traditional equilibrator systems led the community toward developing simpler and more autonomous systems. Here we performed a comparison between a membrane-based sensor and a showerhead equilibration sensor installed on two SOOP-CO$_2$ between 2013 and 2018. We identified time- and space-adequate crossovers in the Skagerrak Strait, where the two ship routes often crossed. We found a mean total difference of 1.5 ± 10.6 μatm and a root mean square error of 11 μatm. The pCO_2 values recorded by the two instruments showed a strong linear correlation with a coefficient of 0.91 and a slope of 1.07 (± 0.14), despite the dynamic nature of the environment and the difficulty of comparing measurements from two different vessels. The membrane-based sensor was integrated with a FerryBox system on a ship with a high sampling frequency in the study area. We showed the strength of having a sensor-based network with a high spatial coverage that can be validated against conventional SOOP-CO$_2$ methods. Proving the validity of membrane-based sensors in coastal and continental shelf seas and using the higher frequency measurements they provide can enable a thorough characterization of pCO_2 variability in these dynamic environments.

The most recent global carbon budget study estimated the atmospheric carbon dioxide concentration growth for the 2009–2018 period at 2.3 ± 0.01 ppm yr^{-1} (Friedlingstein et al. 2019). This increase would have been larger had the ocean not taken up 2.5 ± 0.6 Pg C yr^{-1} over the same period (Friedlingstein et al. 2019). Overall since the industrial revolution, the ocean has taken up approximately a quarter of the anthropogenic carbon emitted (Le Quéré et al. 2018). While the global ocean is on average a carbon sink, there is substantial seasonal and regional variability (Takahashi et al. 2009). This heterogeneity is particularly evident in coastal seas and continental shelves, where low-latitude regions are generally considered to be carbon sources to the atmosphere, while mid-latitude regions are carbon sinks (Borges et al. 2006; Cai et al. 2006). On a global scale, this has led to continuous efforts to quantify the contribution of the coastal ocean to atmospheric carbon dioxide uptake. Estimates have ranged from 1 Pg C yr^{-1}, when the concept of the "continental shelf pump" was first introduced (Tsunogai et al. 1999), to 0.21 Pg C yr^{-1} with more recent estimations (Laruelle et al. 2010). A key limitation in accurately assessing this value is the limited number of observations available for determining air–sea CO$_2$ fluxes (Roobaert et al. 2019).

Ships-of-Opportunity equipped with instruments measuring carbonate system parameters (SOOP-CO$_2$) have been used since the 1990s to complement the limited observational capacity of scientific research cruises and fixed-point observatories in oceanic regions around the world (Cooper et al. 1998; Lüger et al. 2004; Chierici et al. 2006). Often making use of repeating commercial vessel routes, they provide a cost-effective way to observe the surface ocean at a relatively large temporal resolution and spatial coverage (Jiang et al. 2019).

*Correspondence: vlad.macovei@hzg.de

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

Measurements of carbon parameters collected by SOOP-CO$_2$ have been used for scientific studies ranging from small, continental shelf scale (Dumousseaud et al. 2010; Jiang et al. 2013) to large, ocean basin scale (Schuster and Watson 2007; Olsen et al. 2008), as well as for long-term investigations (Fröb et al. 2019; Wanninkhof et al. 2019; Macovei et al. 2020). Every year there are more observations of the partial pressure of carbon dioxide (pCO$_2$) in seawater collected into quality-controlled databases such as the Surface Ocean Carbon Atlas (SOCAT), which grows with every published version (Bakker et al. 2018). Traditional showerhead equilibrator-style pCO$_2$ measuring instruments are relatively large, expensive, and require a lot of power. Their general complexity means trained technicians are needed and a strict set of rules assuring that suitable on-board operations are followed. Furthermore, the use of standard gases for calibration, as recommended in the standard operating procedures, complicates the logistics of their installation (Dickson et al. 2007).

There is still potential to expand data collection through the use of commercial ships, but the challenges of initial installation, logistics, and data quality remain. With these challenges in mind, instruments with semipermeable membranes for dissolved gas equilibration, which are autonomous, and easy to install and operate, have been developed in the past 10 years (Byrne 2014). The technical design of these systems requires less frequent and less costly maintenance and often does not require standard gases for calibration during deployment. While showerhead-equilibrator systems with calibration gases capable of claimed accuracies better than 2 μatm remain the "gold standard" in marine observations, membrane-based instruments are being produced and continuously developed by several companies. Therefore, intercomparison exercises which can validate the accuracy of commercial pCO$_2$ sensors alongside established shipboard instruments have become important for the scientific community (Tamburri et al. 2011). Such studies are valuable for maintaining high standards in ocean observations and are considered essential deliverables in large international projects with a focus on environmental observational infrastructures (González-Dávila et al. 2016). Intercomparison studies have been performed with instruments running in parallel in the laboratory, on the same ship, on the same mooring, or even on different ships at nearby locations. The results from these studies suggest that membrane-based sensors have high potential for achieving the high standard of accuracy required by the scientific community (Körtzinger et al. 1996; Jiang et al. 2014; Lorenzoni et al. 2017; Laakso et al. 2019; Arruda et al. 2020).

FerryBoxes are automated instrument packages usually installed on commercial ships that make use of the regular shipping routes to provide surface seawater measurements. Current active routes are found in the North, Baltic, Norwegian, and Mediterranean Seas (www.ferrybox.org). The setup of FerryBox systems involves a water inlet from where

seawater is pumped into an array of sensors. A basic FerryBox includes sensors for temperature, salinity, turbidity and chlorophyll- *a* (Chl- *a*) fluorescence, and a GPS receiver for position control (Petersen 2014). The system is computer-controlled and data is transmitted directly to shore, which gives it a high degree of autonomy and makes it ideal to use on SOOP-CO$_2$. The compact nature and modular functioning of the autonomous carbon sensors allows simple integration with existing FerryBox systems. The basic FerryBoxes have been measuring since the start of the century (Petersen et al. 2018) and more recently, membrane-based pCO$_2$ and alkalinity sensors have been integrated on some of the routes (Voynova et al. 2019).

One of the main regions covered by the FerryBox community is the North Sea. Continental shelf seas receive significant inorganic nutrient inputs from surrounding terrestrial sources, which makes them highly productive regions important in the global carbon cycle (Gattuso et al. 1998). The North Sea is no exception—with a relatively short residence time (Otto et al. 1990) and strong circulation connection to the North Atlantic, it is considered a major sink for atmospheric carbon dioxide (Thomas et al. 2004) and an efficient carbon export pathway below the permanent thermocline (Thomas et al. 2005). Will these characteristics be maintained in the context of increasing atmospheric pCO$_2$ and sea surface temperatures? Clargo et al. (2015), for example, demonstrated that the North Sea has already experienced years when the sum of the processes leading to carbon outgassing exceeded the carbon uptake and transformed it into a carbon source. Their study, however, calculated seawater pCO$_2$ from dissolved inorganic carbon and total alkalinity measurements, a method which usually has limited temporal and spatial coverage, especially in shelf seas like the North Sea, which experiences large pCO$_2$ gradients (Salt et al. 2013). To properly evaluate whether a trend exists and to further investigate and quantify the regional variability in coastal sinks and sources, a need arises for direct high frequency and long-term pCO$_2$ measurements, such as those obtained from underway instruments (Omar et al. 2010, 2019; Laruelle et al. 2018). Traditionally, equilibrator-based instruments have been used, but recently, membrane-based sensors are becoming more common. Having long time series of both systems makes the North Sea an excellent fit for studying the benefit of supplementing equilibrator-based observations with validated, high-resolution membrane-based observations.

This study performs a crossover investigation between pCO$_2$ measurements taken by membrane sensors on FerryBox-equipped SOOP-CO$_2$ and measurements taken by a SOOP-CO$_2$ equipped with a conventional equilibrator-based instrument (Pierrot et al. 2009). Unlike previous intercalibration studies, this study is done on two long-term data sets in a dynamic coastal environment, where not only background environmental characteristics are changing, but also sensors were swapped and recalibrated multiple times between 2013 and

2018. Based on the validated data set, the advantages of the higher temporal resolution of the FerryBox measurements are explored. This work contributes toward the community need for intercomparison studies and, by validating a new data set, increases our observational capacity, in line with operational oceanography recommendations (Davidson et al. 2019).

Materials and procedures

Instrument descriptions

Oceanographic data were collected via a flow-through FerryBox (Petersen 2014) installed aboard the cargo vessel (CV) *Lysbris Seaways* (DFDS Seaways), traveling in the North Sea since 2007. While the vessel changed routes between 2007 and 2018, in this study we only examine data part of the route between (1) Immingham, UK, (2) Moss/Halden, Norway, and (3) Zeebrugge, Belgium. The FerryBox contains instruments that measure a variety of oceanographic variables, but the main focus of this study is on the partial pressure of carbon dioxide instrument—HydroC CO$_2$-FT (formerly Kongsberg Maritime Contros GmbH, Kiel, Germany; now 4H-Jena Engineering GmbH, Kiel, Germany; accuracy of ± 1%). Temperature and salinity sensors (Falmouth Scientific, Cataumet, Massachusetts, U.S.A. and Teledyne RD Instruments, Poway, California, U.S.A.), oxygen optode (Aanderaa Instruments, Xylem Analytics, Germany), and pH electrode (Sensortechnik Meinsberg, Xylem Analytics, Germany) are also used to aid interpretation of data.

The HydroC CO$_2$-FT *p*CO$_2$ instrument has been operated on the *Lysbris* line since late 2013. It measures the concentration of carbon dioxide in a stream of wet air using a nondispersive infrared (NDIR) detector following equilibration through a semipermeable silicone membrane. From here on, this instrument is referred to as MBS (membrane-based sensor). While the instrument takes a reading every second, only the 20-s averages are recorded. These averages were processed according to manufacturer recommendations (Fietzek et al. 2013). No response time correction was applied since the τ_{63} response time factor (time it takes for the measurement to reach 63% of the final value) is between 70 and 120 s for the HydroC sensors (Fiedler et al. 2013; Fietzek et al. 2013) and this is small enough to not interfere with the crossover selection (details below). Anti-biofouling measures were taken during every port visit. The instrument automatically entered a wash cycle where sulfuric and oxalic acids were flushed through the system for about 5 min at a pH of 2. The intake lines were then rinsed with freshwater and freshwater was kept inside the lines until the next journey.

A deployment was defined as the time from when a sensor was initially installed on the ship until it was replaced. In all cases but one, the sensor had been recently calibrated. For each deployment before 2017, the precalibration information from the manufacturer data

processing sheet was used for calculations. After 2017, both pre- and postcalibration coefficients were used. The NDIR detector uses a dual-beam system according to two wavelengths at which carbon dioxide is efficiently absorbed (the raw signal), and at which no absorption occurs (the reference signal). The two-beam signal was obtained by dividing the raw signal by the reference signal. No external gas standards were used, but regular zeroing during the lifetime of a sensor allowed for a correction for instrument drift to be applied. The zero drift was anomalously strong for the deployment between April and June 2015, so those results were excluded from this analysis. A linear trendline was fitted through the dual beam ratio at the times of zeroing and thus an equivalent "zero" value was calculated for each real measurement. The drift-corrected signal was then used in conjunction with the three coefficients of the calibration polynomial and the gas temperature to calculate the wet mole fraction of CO$_2$ (xCO$_2$). The correction for the pressure inside the sensor was the final step for converting the values to *p*CO$_2$. An alternative drift correction method of linear interpolation between each zeroing event was also checked. The mean difference between the calculated *p*CO$_2$ with the two methods was 0.06 ± 1.94 μatm, which suggests that the drift correction method does not affect the results. Since the start of 2017, the instruments were calibrated again at different concentrations at the end of their deployment (postcalibration) to cover concentration-dependent effects that equate to changes in the characteristics of the NDIR sensor's calibration polynomial (Fietzek et al. 2013). Using both the pre- and postcalibration coefficients, a span drift correction was applied by linearly interpolating the calibration polynomials according to the change in the zero signal during the deployment lifetime. For the two deployments when postcalibration was available, the observed span drift was approximately 1.9 μatm month^{-1}. This becomes particularly important when a sensor is deployed over a long period of time. A final correction was done by eliminating any data points outside the 200–800 μatm instrument calibration range. The MBS data, as well as the other measurements taken on the *Lysbris Seaways*, are available from the European FerryBox Database (http://ferrydata.hzg.de/).

Showerhead equilibrator-based measurements were collected on the commercial vessel M/V *Nuka Arctica* (Royal Arctic Lines). The ship operated between Ilulissaat, Greenland, and Aalborg, Denmark crossing the Skagerrak region in an eastward (westward) direction when sailing into (out of) Aalborg port. The *p*CO$_2$ measuring system (GO system model 8050, General Oceanics, U.S.A.) was located in the engine room of the vessel. Regular cleaning and exchanges of hoses, filters, and equilibrators were done during port visits to prevent biofouling. A detailed description of the setup can be found in Olsen et al. (2008). The instrument uses a NDIR CO$_2$/H$_2$O gas analyzer (LI 6262, LI-COR Biosciences, U.S.A.),

equipped with one zero and three nonzero reference gases, which are used for calibration every 3 h. Seawater carbon dioxide fugacity (fCO$_2$) is calculated according to Pierrot et al. (2009), with a final reported uncertainty of 2 μatm, although the lack of a certified water standard for fCO$_2$ makes direct assessment of the accuracy impossible. Furthermore, field deployments can introduce additional sources of error such as respiration due to biofouling in the pipes (Juranek et al. 2010) or positive bias in the measurement of the temperature of equilibration (Arruda et al. 2020). Nevertheless, the *Nuka Arctica* line is well established in the SOCAT database with 38 transects having a quality flag of A indicating a high quality crossover with another data set during each of these. An example of such a comparison is shown by Pierrot et al. (2009). The combination of the more stable measurements on the long trans-oceanic lines and the membrane-based systems on coastal lines is a good network setup for crossover comparisons in the coastal area. The CO$_2$ data are available through the Surface Ocean Carbon Atlas (www.socat.info) or the Integrated Carbon Observation System (ICOS)-Norway website (https://no.icos-cp.eu). The salinity data can be accessed via the LEGOS website (http://www.legos.obs-mip.fr/observations/sss/datadelivery/dmdata). The reported fCO$_2$ was converted to pCO$_2$ using CO2SYS in order to compare equivalent units (van Heuven et al. 2011). From here on, this instrument is referred to as showerhead-equilibrator system (SHS).

On the *Nuka Arctica*, the sea surface salinity was measured using an SBE 21 Seacat thermosalinograph and calibrated using in situ salinity samples according to Alory et al. (2015). The sea surface temperature was measured by a regularly calibrated thermometer (model 1524, Fluke, The Netherlands) right after the water intake, which was located at about 5 m water depth. All pipes were insulated with Styrofoam and the warming between the water intake and the equilibrator is usually in the order of 0.2–0.3°C. This warming is taken into account by correcting the fCO$_2$ using the relationship of Takahashi et al. (1993).

Crossover selection

Comparison of measurements collected by different vessels has always been a challenge for the oceanographic community due to the combination of uncertainties given by instrumental precision and dynamic heterogeneity of the sampled water masses. Crossover studies for cruises with deep stations in the open ocean define a crossover as stations within 1° ($\sim$ 100 km) of each other, and biogeochemical variables are compared only below 2000 m depth to minimize the effect of real variations (Olsen et al. 2016). For surface observations, SOCAT guidelines recommend a maximum distance of 80 km where an algorithm of both space and time is used: $[dx^2 + (30dt)^2]^{0.5} \leq 80$ km, meaning that 1 d of separation in time is equivalent (heuristically) to 30 km of separation in space (Olsen et al. 2015). In the much more dynamic coastal ocean, this recommended distance needs to be shortened when surface measurements from different platforms are compared. When defining crossover criteria for surface seawater fCO$_2$ in a shelf sea setting, Kitidis et al. (2019) used a maximum distance of 40 km by defining 1 d time difference as equivalent to a distance of 30 km.

Since both instruments investigated in this study were installed on commercial ships, the routes across the North Sea varied according to the required port calls and the weather (Fig. 1a). An area where the routes sampled by the MBS and by the SHS often overlapped was the Skagerrak Strait between Denmark and Norway. Within this region, five subregions (Fig. 1b) had the highest numbers of voyages where the two ships came in close proximity of each other. Since the size of these subregions is small (largest possible difference is 32 km), valid crossovers were identified as those when the two ships
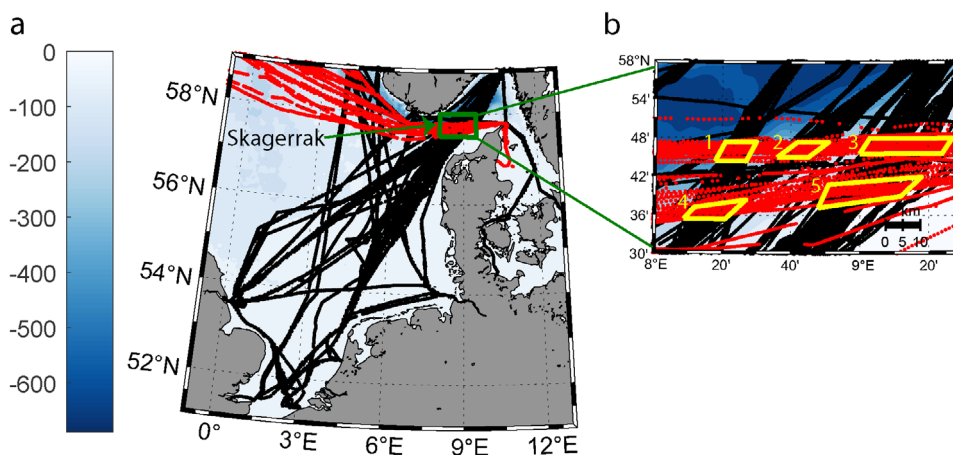

Fig 1. (**a**) Bathymetric map of the North Sea displaying the locations of the MBS pCO$_2$ measurements taken by the *Lysbris Seaways* cargo ship between 2013 and 2018 (black) and the SHS pCO$_2$ measurements taken by the *Nuka Arctica* container ship during the same time on its route toward Greenland (red). (**b**) The zoom-in on a restricted area in the Skagerrak (green box) identifies five subregions with the highest probability of valid crossovers.

passed through one of the subregions within 24 h of each other.

As an exercise in testing the typical water mass drift in this area, data were retrieved from the Drift App of the CoastMap Geoportal (www.coastmap.org) under CC BY-NC 4.0 license. The application uses marine currents, wind drag and an underlying Lagrangian transport algorithm to simulate water mass movement (Callies et al. 2017). Drifter trajectories were forward modeled for 24 h after release from the middle of the restricted Skagerrak box once per month for the 2014–2018 period. The mean distance traveled after 24 h was 12.9 km and 97% of the drifters traveled less than 31 km. Further details are provided in the Supporting Information, but we conclude that 24 h is an appropriate window of time for crossover selection in this region.

Results

Sampling in the Skagerrak

The FerryBox sampling of the North Sea covers a wide range of environments with various degrees of coastal influence and different seasonality characteristics. The deeper northern part of the North Sea experiences seasonal stratification as summertime warming of the surface establishes a strong pycnocline (Wakelin et al. 2012), while the shallower southern part is typically permanently mixed due to strong tidal forcing (Bozec et al. 2005). More recent studies have defined biome boundaries on an even finer scale depending on the stratification regime and on the freshwater influence on biogeochemical processes showcasing the high spatial variability in the North Sea (van Leeuwen et al. 2015; Kerimoglu et al. 2020). While the *Lysbris* FerryBox-MBS pCO$_2$ measurements are available over a large area of the North Sea, in order to check and validate the long-term MBS pCO$_2$ record, we focus on the small restricted box in the Skagerrak where many

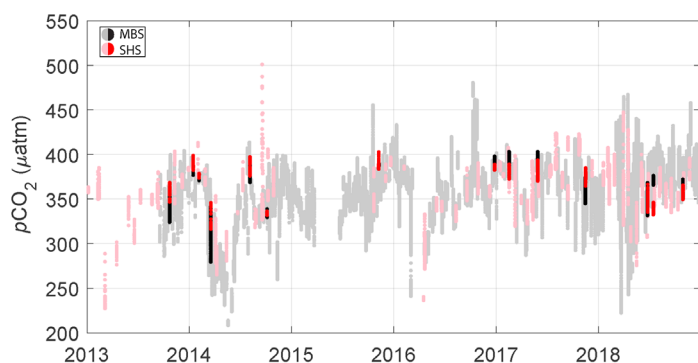**Fig 2.** Time series of the pCO$_2$ measurements taken by the two different types of sensors (MBS on *Lysbris Seaways* and SHS on *Nuka Arctica*) in the restricted box in the Skagerrak shown in the zoom-in of Fig. 1 (lighter shades) and an emphasis on the valid crossovers identified (darker shades—details below). The horizontal axis tick labels represent the start of the respective year.

crossovers with the SHS system on *Nuka Arctica* were identified.

Figure 2 displays the available pCO$_2$ data in the Skagerrak from the MBS and SHS sensors compared in this study. The temporal coverage varies during 2013–2018 according to the shipping schedules, route variations and instrument functioning status. The years of 2014, 2017, and 2018 have particularly good coverages in this region by both observing platforms. In these years, the maximum MBS pCO$_2$ values before mid-March increased from 388 μatm to 411 μatm and 438 μatm respectively, while the mean variability (taken as standard deviation of measurements for each passage through the Skagerrak box) was 7.4 μatm, 6.4 μatm, and 11.9 μatm respectively. The high sampling resolution allows observations of a distinct seasonal cycle in the evolution of surface pCO$_2$, with a typical decrease during the spring bloom, usually starting in early March, and an increase in the summer. This is a result of biological production in the spring season consuming dissolved inorganic carbon and consequently lowering the pCO$_2$, and of increasing sea surface temperature during the summer season, which increases pCO$_2$. The combined effect of biological and physical processes on the variability in surface pCO$_2$ in both midlatitude oceanic and shelf regions has been discussed before (Takahashi et al. 2002; Omar et al. 2010; Jiang et al. 2013; Macovei et al. 2020). Seasonal variability was captured by both instruments with some differences. In particular, discrepancies are observed in the second part of 2017 and 2018.

Range of measurements

The range of measurements during a single voyage through the 90 × 55 km Skagerrak box was large, so the crossover study was performed only where the distance between sampling locations was minimal. This was done to compare and validate the instruments and minimize the influence of natural variability. The small size of the five selected subregions means that each ship will pass through one subregion in less than 1 h. Even so, the ranges of recorded pCO$_2$ per passage can be large. Figure 3 shows the ranges in measurements during one passage for both the MBS and SHS instruments. Values larger than the whiskers, which represent 2.7 standard deviations from the median, were classified as outliers. A similar analysis was performed for the temperature sensors and is detailed in the Supporting Information.

Crossover differences

A total of 21 crossovers were identified using the geographical and temporal selection criteria. During these times, due to the high sampling frequency, the two ships took several measurements inside one subregion. In some cases, the range of these measurements was large (Fig. 3). In order to avoid comparisons when the ships might have been sampling multiple water masses, seven crossovers were eliminated from the analysis when at least one of the instruments reported a pCO$_2$

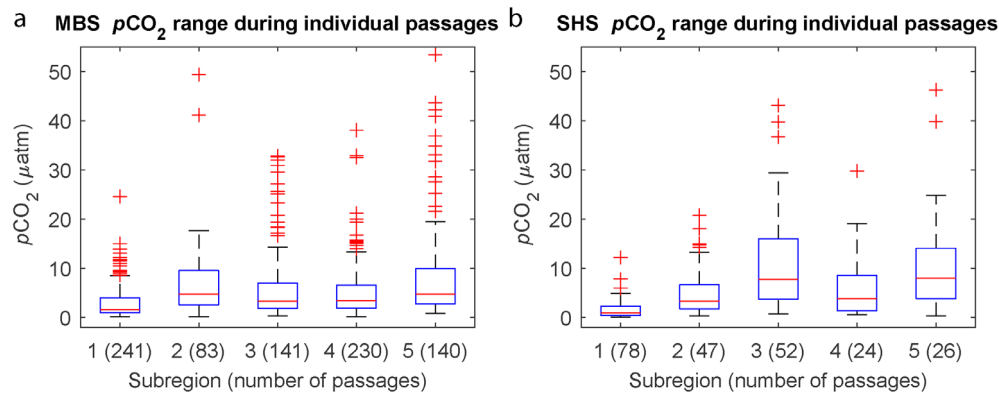**Fig 3.** Box and whisker plots showing the range of pCO_2 measured by the MBS (**a**) and SHS (**b**) during individual passages through a respective subregion. On each box, the central red mark indicates the median, and the bottom and top edges of the blue box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using red "+" symbols. The total number of passages through each subregion during the 2013–2018 period is shown in brackets. In one of the passages through subregion 5, the SHS instrument measurements range was nearly 140 μatm. This data point is not shown for better visualization.

Fig 4. Comparison between the pCO_2 measured by the MBS and the SHS (**a**) and the seawater temperature measured by the temperature sensors on the two ships (**b**) at the valid crossovers after removal of outliers. The data points are split based on the subregion they come from (1 = blue, 2 = green, 3 = black, 4 = red, 5 = cyan). In subplot (**a**), the gray shaded area denotes the uncertainty around the 1:1 line given by the respective instrument uncertainties and the thicker black line is the best fit of the linear regression model. The error bars represent 1 standard deviation of the mean of the measurements taken during the respective passage through the subregion. Only error bars larger than 1 μatm are shown for clarity. In subplot (**b**), the dashed line is the 1:1 line and the solid line is the best fit of the linear regression model with the equation shown on the plot. No error bars are shown since they are all smaller than 0.2°C.

range above the outlier threshold while passing through a subregion. The means of the pCO_2 measurements in the remaining 14 crossovers could be directly compared to identify differences between the two instruments. The geographical separation did not introduce any patterns in the crossover analysis (Fig. 4a), so henceforth, all the valid crossovers will be analyzed together, irrespective of subregion.

The MBS and SHS measurements in all the valid crossovers within a 24 h time window ($n = 14$) were linearly dependent (Pearson's coefficient of 0.91, $p < 0.01$) and the linear model

(MBS = 1.07 [± 0.14] × SHS − 25 [± 51]) had a slope not statistically different from 1 and a root mean square error of 11 μatm (Fig. 4a). The mean difference (MBS − SHS) was 1.5 ± 10.6 μatm and the range was between −16.9 and 25.0 μatm. We performed the same crossover analysis by allowing a 48 h time window and obtained similar results (details in the Supporting Information). Given the dynamic nature of the environment, we preferred to restrict our comparison to the 24 h time difference allowance. The MBS and SHS have reported uncertainties of ± 1% and ± 2 μatm, respectively. These were used to determine instrument-related uncertainty around the 1:1 line. The data points that fall within the shaded boundary in Fig. 4a also fit the high-quality crossover criterion of less than 5 μatm difference as defined in the SOCAT cookbook (Lauvset et al. 2018). The larger temperature differences and the fact that an "alternative" sensor (according to the SOCAT definitions) was used means however that the data set cannot have a quality flag better than E.

The temperature measurements from the two ships were also linearly correlated with a Pearson's coefficient of 0.97, $p < 0.01$ (Fig. 4b). The seawater temperature measured on the *Lysbris Seaways* was in most cases higher than the one on the *Nuka Arctica*. No trend was found in the temperature difference during the time series. While the temperature sensor on the *Nuka* is located immediately after the water inlet, the temperature sensor on the *Lysbris* is located at least 7 m away from the inlet, and around 0.5°C of warming may be expected (W. Petersen, pers. comm., May 2020). A past study comparing a FerryBox temperature sensor from a different ship against a fixed buoy found that the ship-based measurements were on average 0.37°C warmer than the in situ temperatures (Haller et al. 2015). While this comparison was performed for a different ship, the FerryBox installation is similar between the two ships so we can assume a similar warming inside the intake pipes of *Lysbris Seaways*. The big intercept of the linear

fit equation for the temperature comparison suggests that the difference between the sensors is greater at low temperatures, when there is a stronger gradient between seawater and ambient engine room temperature. However, no seasonal bias was identified in the pCO_2 comparison, and the difference between the pCO_2 instruments was not correlated with temperature difference between the sensors.

wAs expected, most of the valid crossovers were identified when the density of available data was highest (Fig. 5a). Of the 14 valid crossovers, 10 fall within the ± 10 μatm difference, but for 2 crossovers near the start of the time series, and for 2 others near the end, the difference was larger than 10 μatm. There is a moderate (ρ = 0.54, $p < 0.05$) correlation between the pCO_2 difference and the MBS pCO_2, but this is likely occurring because seawater pCO_2 was higher toward the end of the time series, when some of the MBS measurements at the crossovers were higher than the SHS ones. Instrument drift cannot explain the apparent trend in the difference since the sensors have been replaced multiple times. Furthermore, MBS data for the final three deployments in Fig. 5a (after December 2016) were both corrected for the zero drift and recalculated following postcalibration with both pre- and postcalibration coefficients. The deployments when crossover differences were larger than 10 μatm also included smaller crossover differences (< 10 μatm), so any apparent trend in the evolution of these differences is coincidental. There is also no evidence that the

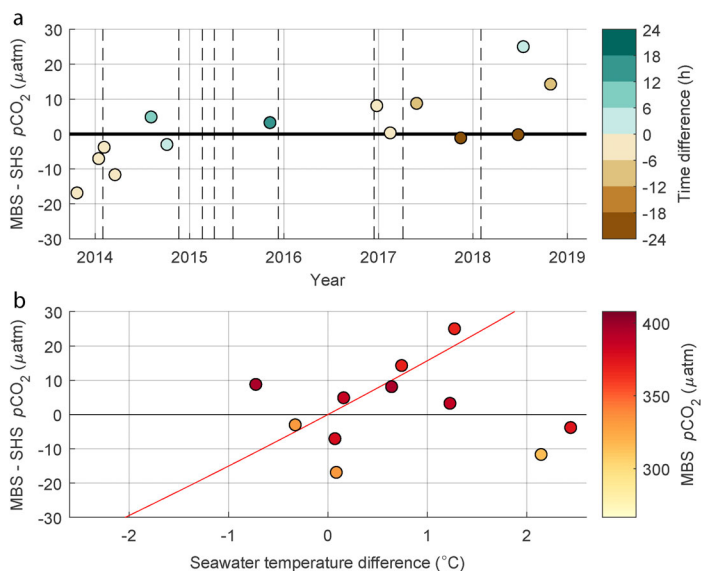time differences between the measurements influence the pCO_2 differences. Most of the valid crossovers happen at night and none occurred at the two extremes of the diel cycle.

Figure 5b shows the pCO_2 difference between the two platforms vs. the sea surface temperature difference. There are fewer data points shown in this subplot since, in three cases, the temperature range did not pass the valid temperature crossover criteria. Seawater pCO_2 changes between two hypothetical temperatures states of the water mass ($T_{initial}$ and T_{final}) according to the (pCO_2 at T_{final}) = (pCO_2 at $T_{initial}$) × exp [0.0423 × ($T_{final} - T_{initial}$)] relationship (Takahashi et al. 2002). For example, at typical seawater pCO_2 values in our study area, the pCO_2 measured in a ≈ 1°C warmer water mass would be ≈ 16 μatm higher without any change in the chemical composition. An idealized line of no-difference between the two instruments following the temperature dependence is also shown. Any points that fall on this line would yield a 0 μatm crossover difference if either pCO_2 measurement would be corrected to the in situ temperature measured by the other ship. While some crossover differences appear to fall on this line, there was no conclusive evidence that a temperature adjustment to a common value would decrease the crossover pCO_2 difference. The ship intake valves and piping, and the location of the temperature sensors are different, so correcting one set of measurements to the temperature of the other risks introducing artificial errors. In addition, the SHS pCO_2 was corrected for the difference between the in situ and equilibration temperature, while this was not done for the MBS data since the seawater temperature was not measured immediately after the intake location. Attempts were made to correct the MBS pCO_2 data to satellite-derived in situ temperature. While this improved the crossover differences in 2018, it increased the differences in earlier crossovers. We also calculated a hypothetical pCO_2 correction range assuming a warming between the intake and the MBS from 0.2 to a more drastic 0.8°C. This translated to a 3.1–12.2 μatm decrease of the MBS pCO_2 values, which similarly improves positive crossover differences and worsens negative ones. The lack of directly measured in situ temperature remains a limitation of the current setup.

Discussion

Advantages of higher frequency sampling

The shipping schedules mean that *Lysbris Seaways* samples the Skagerrak at a higher frequency than *Nuka Arctica*. In the fall of 2016, the direct comparison of the two instruments in the restricted Skagerrak box revealed that some measurements from the MBS were substantially higher. These are highlighted in blue in Fig. 6a. Measurements from two crossings of *Nuka Arctica* in this period match closely with the *Lysbris Seaways* measurements from the nearest (with respect to time) crossings and with most other measurements in-between. Investigating the location of the anomalously high measurements revealed that these were taken when the ship was using a

Fig 5. Difference between pCO_2 measured by MBS and SHS during the valid crossovers vs. (**a**) date of measurement and (**b**) difference between the temperature sensors on the two ships (*Lysbris – Nuka*). The colormap in subplot (**a**) shows the time difference between the passages of the two ships through a subregion in 6 h intervals. The colormap in subplot (**b**) represents MBS pCO_2. Vertical dashed lines in subplot (**a**) show the dates when newly calibrated MBS were installed on the *Lysbris*. Subplot (**b**) also shows the idealized temperature-dependent line of no difference in red using a reference pCO_2 of 363 μatm – the mean of the MBS crossover values.

different route (Fig. 6b). Crossing the shallower, south-eastern part of the Skagerrak, the ship was transiting a region with a different water mass type. It has been shown that pCO$_2$ values and total alkalinity concentrations and their seasonal cycles can vary tremendously depending on the proximity to the coast (Voynova et al. 2019). FerryBox salinity measurements in the Skagerrak were not significantly different during this period, irrespective of sampling location. Therefore, the biogeochemical rather than physical properties drive the differences in pCO$_2$ in this case. These measurements were taken at the end of the March–October "biologically active" period (Andersson and Rydberg 1993). Some of the dissolved inorganic carbon will have been consumed and thus lower pCO$_2$ levels are expected. While this was true for the western locations, the eastern regions had high pCO$_2$ values. Furthermore, investigating additional FerryBox measurements confirmed that, compared to the western side, the water mass in the eastern side of the Skagerrak in the fall of 2016 was distinctly different, with lower dissolved oxygen (Fig. 6c) and lower pH (Fig. 6d).

In this example, the high frequency FerryBox sampling captured both the spatial and temporal variability of pCO$_2$ in the Skagerrak. The water mass in the eastern Skagerrak had higher pCO$_2$ levels, but this was a short-lived event, since at the end of October, when *Nuka* crossed the region again, the measurements were back to levels from early October. The

exact reason for this event is beyond the scope of this manuscript, but the finding does show that larger and more frequent observational footprints are necessary for accurate biogeochemical characterizations, particularly in coastal regions where seasonal and spatial variations can be large compared to the open ocean. The Skagerrak is an important region since 70% of North Sea water passes through here before being exported to the North Atlantic (Danielssen et al. 1996). It is also a location where mixing between water masses with different biogeochemical characteristics takes place, so it is important to resolve for inclusion in models, which are used to create larger scale budgets or basin-wide gas exchange estimates. The lack of observations in nearshore and estuarine regions is the largest factor contributing to uncertainties in air–sea CO$_2$ fluxes (Legge et al. 2020). While shelf seas as a whole are considered to be carbon sinks, the coastal input of carbon to the water column makes the very nearshore waters a likely source of CO$_2$ to the atmosphere (Borges et al. 2005; Chen and Borges 2009). The atmospheric dry air mole fraction of carbon dioxide in October 2016 was 404 ppm at Mace Head, Ireland (World Data Centre for Greenhouse Gases 2020). Using atmospheric pressure and dew point temperature from the ERA-Interim reanalysis product in the Skagerrak (Dee et al. 2011), we calculated the water vapor pressure (Alduchov and Eskridge 1996; Lawrence 2005) and subsequently an atmospheric partial pressure of carbon dioxide of 406 μatm. Had
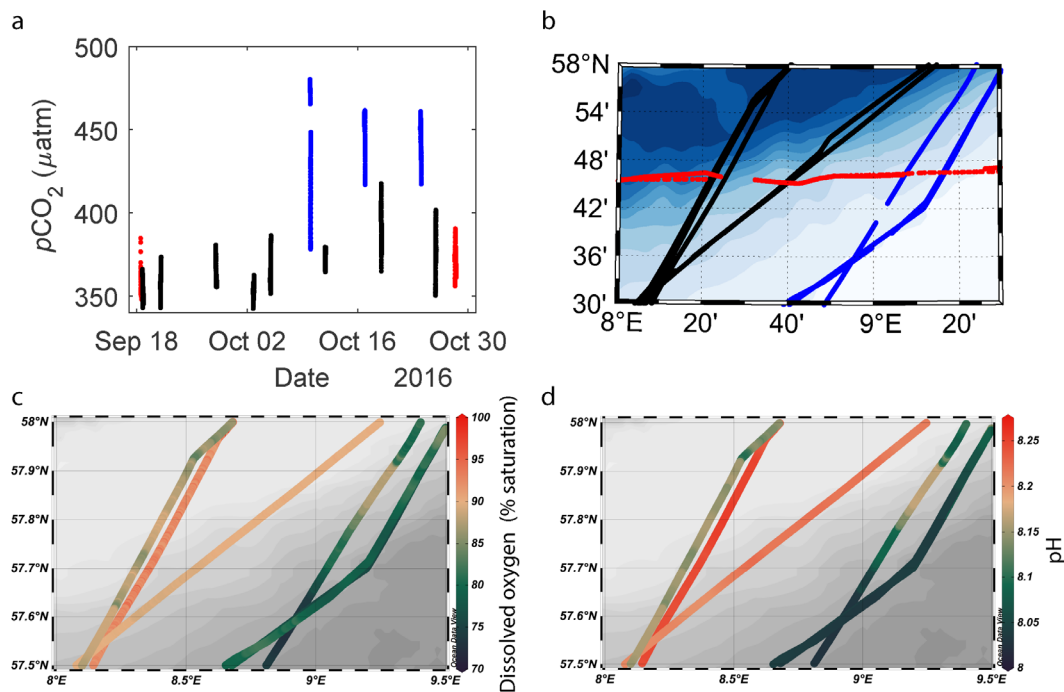

Fig 6. Continuously measured pCO$_2$ (**a**) during the fall 2016 passages through the restricted Skagerrak region using the SHS (red) and the MBS (black/ blue) instrument. Blue was used to identify the crossings with the highest pCO$_2$ values. Map of the restricted Skagerrak box (**b**) with the locations of the measurements shown in the respective color. Dissolved oxygen saturation (**c**) and NBS scale pH (**d**) measured on the *Lysbris* and shown in relation to the ship routes during fall 2016. Subplots (**c**) and (**d**) were produced with the ocean Data View software (Schlitzer 2020).

only the available measurements (SHS and MBS) on the usual route through the middle of the Skagerrak been considered (Fig. 6, red and black colors), one could conclude that the entire region was acting as a sink for atmospheric carbon in the fall of 2016 (seawater pCO_2 < atmospheric pCO_2). The high MBS pCO_2 measurements taken on the route closer to the Danish coast (Fig. 6, blue color) revealed that this nearshore region was in fact a carbon source during the second part of October 2016. This highlights the issue that observation-based estimates of shelf-wide CO_2 uptake are likely to be overestimated (Legge et al. 2020).

Considerations for present and past intercomparisons

The comparison between the MBS and SHS instruments in this study indicated that these two measurement methods were both able to capture the seasonal variability of the surface waters in the Skagerrak and, when instrumental precision and the rapidly changing environment are considered, showed similar results. The comparison yielded a small average difference of 1.5 μatm considering the dynamic environment investigated, but a large range between the biggest positive and negative differences. The MBS data were not corrected to in situ temperature since this measurement was not available. As seen from Fig. 5b, correcting the MBS data to the in situ temperature of *Nuka* improves some of the comparisons and worsens others. However, this correction does not resolve the difference between the diurnal cycles of temperature and pCO_2 and it assumes the temperature measured inside the FerryBox is the same as the in situ temperature. A further attempt was made to correct the MBS data to satellite derived OSTIA sea surface temperature (Donlon et al. 2012). This correction improves the 2018 comparisons but worsens the overall difference to −12.2 ± 22.4 μatm. The satellite product has a grid resolution of approximately 28 km which might be too big for such a dynamic region. Finally, the MBS data were recalculated assuming a 0.2–0.8°C range of warming between the intake and sensor. The corresponding range of pCO_2 decrease is between 3.1 and 12.2 μatm so, as before, this correction is inconclusive in improving the comparison and we do not use it to avoid introducing artificial errors.

Diurnal variations of biologically mediated parameters such as pCO_2 are large in coastal surface waters. Rapid drawdowns of more than 50 μatm were observed after dawn in a high latitude coastal environment during the productive season (Tortell et al. 2014). Daily variations ranging from 10 μatm in an oligotrophic setting to over 60 μatm in a coral reef system were seen in a low latitude coastal environment (Dai et al. 2009). Photosynthetically active radiation, temperature, and biological metabolism variability dominate the pCO_2 variability in coastal settings. We expect that diurnal variability may have an influence on our comparison, since the MBS and SHS instruments discussed here were installed on different ships and their measurements were not synchronous. While the agreement between the instruments is better when a 24 h

time allowance is used compared to a 48 h one, further restricting this does not improve the comparison and limits the number of available valid crossovers.

In addition, the Skagerrak Strait in particular is a challenging location for a crossover study, due to mixing of multiple water masses (Albretsen et al. 2011; Kristiansen and Aas 2015). This is true especially for biologically mediated parameters. A comparison study between FerryBox data and conventional research vessel observations in the nearby Kattegat Strait and Baltic Sea found a very good match for temperature and salinity, but a weaker match for oxygen and Chl- *a* (Karlson et al. 2016). Similarly, pCO_2 is controlled by biological processes and can vary significantly over short spatial and temporal scales.

The Skagerrak Strait is characterized by a counter-clockwise circulation (Gröger et al. 2019) and is a mixture of water masses, including Atlantic, Central North Sea, Coastal North Sea, and Baltic (Albretsen et al. 2011). The slow flushing times of the Baltic Sea means that the ratio of water entering the Skagerrak from the east compared to the west is about 1:10 (Rodhe 1996). In addition, the Baltic Sea water outflow flows next to the Norwegian coast, away from our restricted box. What instead makes a difference in the restricted Skagerrak box is the Coastal North Sea water, which is shown to influence the eastern region (Holt and Proctor 2008; Kristiansen and Aas 2015). This water mass contains the products of remineralized organic matter and mostly influences the sites of the "blue" sampling points in Fig. 6. Higher dissolved inorganic carbon concentrations in this coastal water could be causing the unusual observations.

While ICOS/SOCAT remain the reference networks with the highest quality data and long distance routes, the quality of membrane sensors can be verified against them. Through the benefit of a high temporal resolution given by short voyage times, the two data sets can complement each other to fill-in temporal and spatial gaps. Crossover studies such as the present one should be repeated regularly to detect any potential issues with the membrane sensors and add confidence to the results. A summary of such past comparisons is shown in Table 1.

In a previous attempt to compare MBS data from *Lysbris Seaways* to a SHS reference instrument, Kitidis et al. (2019) found a regression slope statistically indistinguishable from unity and a residual of 16.7 μatm for their 91 identified crossovers. The comparison was done with a previous version of the MBS data, which have since been reprocessed. In the study of Jiang et al. (2014), the difference between MBS and SHS data on two different ships was −0.3 ± 3.9 μatm, but the comparison was done for one single crossover with a time difference of less than 12 h when the two ships were in proximity in the equatorial ocean. In addition, a temperature correction was applied. The study of Arruda et al. (2020) also compared instruments installed on the same ship. Their mean differences between the MBS instruments and the SHS one were

Table 1. The mean residual between the results of membrane-based and equilibrator-based pCO$_2$ instruments for previous crossover and intercalibration studies on moving platforms.

Reference	MBS instrument	Mean residual (μatm)	Comments and study duration
Fietzek et al. (2013)	Contros HydroC FT	-3.1 ± 2.9 1.8 ± 3.4 -0.7 ± 2.8	• Two deployments (the first one with two MBS instruments) on the same ship • Two 1-month cruises
Jiang et al. (2014)	ProOceanus CO2-Pro CV	-0.3 ± 3.9	• One crossover on different ships • 1 d time window
Kitidis et al. (2019)	Contros HydroC FT[*]	16.7	• 91 crossovers on different ships • 1 yr
Arruda et al. (2020)	SubCTech OceanPack2[†]	-4.7 ± 2.9 -12.6 ± 2.0	• Two deployments on the same ship • A 1-week and a 1-month cruise
Arruda et al. (2020)	ProOceanus CO2-Pro CV	-5.7 ± 4.0 -8.7 ± 3.9[‡] -26.0 ± 6.8[§]	• Two deployments on the same ship • A 1-week and a 1-month cruise
This study	Contros HydroC FT	1.5 ± 10.6	• 14 crossovers on different ships • 5 yr

[*]Comparison done with a previous version of the data.
[†]Calibrated with standard gases during deployment.
[‡]Only using data before a severe storm.
[§]Only using data after a severe storm.

-5.7 ± 4.0 and -4.7 ± 2.9 μatm during their first deployment but as high as -26.0 ± 6.8 μatm during a second deployment affected by biofouling and a storm event.

The mean residual in this study including its uncertainty interval is larger than the usual goals of weather (2.5% relative—so 10 μatm for a 400 μatm measurement) and climate (0.5% relative—so 2 μatm for a 400 μatm measurement) accuracy in marine carbonate system observations (Newton et al. 2015). However, the dynamic nature of the coastal ocean where the crossovers are located makes the difference between the data sets relatively small. Compared to previous intercomparison studies, ours is done over a period of over 5 yr, on separate ships and with different instrument types. Even when two similar instruments (both equilibrator type sensors) installed on the same ship were compared by Ribas-Ribas et al. (2014), the range in the differences varied by up to 41 μatm in spite of having fewer factors that can introduce uncertainty compared to our study. Despite having a lower reported accuracy than SHE, MBS pCO$_2$ instruments can adequately handle the variability in the coastal ocean.

Conclusion and recommendations

Despite the lack of reference gases for MBS checks, we demonstrated that the MBS measurements are comparable to SHS pCO$_2$ data over a long period in a dynamic shelf sea environment. This was accomplished by carefully selecting crossovers in the extremely dynamic Skagerrak region. Integrating the instruments on the high-frequency FerryBox sampler allowed capturing the high spatial and temporal variability in this area.

Public availability of data from coastal observation research infrastructures is a key deliverable in many international projects (DANUBIUS, JERICO, etc.) and necessary for a better understanding of these highly dynamic environments. Following this crossover analysis, we will submit FerryBox pCO$_2$ data from *Lysbris Seaways* for inclusion in the widely used SOCAT database. This will increase the temporal coverage in regions such as the Skagerrak and the spatial coverage in the Central and Southern North Sea.

We recommend that such crossover studies continue to be performed in the future, as membrane-based instruments can provide valuable pCO$_2$ data sets when compared to gas standard-calibrated equilibrator-style instrument data. Future technological developments and participation in field and laboratory intercalibration exercises will likely increase even further the agreement between the two instrument types.

Following the lessons learned through this study, we compile a list of recommendations for the future improvement of the installation on *Lysbris Seaways* as well as new sensor installations. It is important to measure the temperature at the water intake to correct for any potential warming before the equilibration takes place. If possible, antifouling piping material should be used. It is important to test the performance of the sensors by comparing their results to pCO$_2$ calculated from bottles samples that have had two other carbonate system parameters analyzed in a quality-controlled manner. If postcalibration of a MBS is not possible, the deployment length should be limited to minimize the span drift. The user can have more control on the quality of the data if the processing is done starting from the raw instrument values. Finally, intercomparison studies between different ships, such as the

current one, while undoubtedly useful, should be done together with and not as a replacement for side-by-side comparisons of multiple sensors.

Alongside the *Lysbris Seaways*, MBS-equipped FerryBoxes exist on other SOOP-CO₂ that have been transiting the North Sea on various routes. There are now nearly a decade of data covering different regions of the seasonally dynamic and rapidly changing North Sea. The successful comparison with SHE data combined with the higher sampling frequency and spatial coverage in North Sea can allow for a better-characterized marine carbonate system, especially in terms of short-lived anomalous events.

References

Albretsen, J., J. Aure, R. Sætre, and D. S. Danielssen. 2011. Climatic variability in the Skagerrak and coastal waters of Norway. ICES J. Mar. Sci. **69**: 758–763. doi:10.1093/icesjms/fsr187

Alduchov, O. A., and R. E. Eskridge. 1996. Improved magnus form approximation of saturation vapor pressure. J. Appl. Meteorol. **35**: 601–609. doi:10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2.

Alory, G., and others. 2015. The French contribution to the voluntary observing ships network of sea surface salinity. Deep-Sea Res. Part I Oceanogr. Res. Pap. **105**: 1–18. doi:10.1016/j.dsr.2015.08.005

Andersson, L., and L. Rydberg. 1993. Exchange of water and nutrients between the Skagerrak and the Kattegat. Estuar. Coast. Shelf Sci. **36**: 159–181. doi:10.1006/ecss.1993.1011

Arruda, R., D. Atamanchuk, M. Cronin, T. Steinhoff, and D. W. R. Wallace. 2020. At-sea intercomparison of three underway pCO₂ systems. Limnol. Oceanogr.: Methods **18**: 63–76. doi:10.1002/lom3.10346.

Bakker, D. C. E., and others. 2018. Surface ocean CO₂ atlas (SOCAT) V6. PANGAEA.

Borges, A. V., B. Delille, and M. Frankignoulle. 2005. Budgeting sinks and sources of CO₂ in the coastal ocean: Diversity of ecosystems counts. Geophys. Res. Lett. **32**: L14601. doi:10.1029/2005GL023053.

Borges, A. V., L. S. Schiettecatte, G. Abril, B. Delille, and F. Gazeau. 2006. Carbon dioxide in European coastal waters. Estuar. Coast. Shelf Sci. **70**: 375–387. doi:10.1016/j.ecss.2006.05.046

Bozec, Y., H. Thomas, K. Elkalay, and H. J. W. de Baar. 2005. The continental shelf pump for CO₂ in the North Sea - evidence from summer observation. Mar. Chem. **93**: 131–147. doi:10.1016/j.marchem.2004.07.006

Byrne, R. H. 2014. Measuring ocean acidification: New technology for a new era of ocean chemistry. Environ. Sci. Technol. **48**: 5352–5360. doi:10.1021/es405819p

Cai, W.-J., M. Dai, and Y. Wang. 2006. Air-sea exchange of carbon dioxide in ocean margins: A province-based synthesis. Geophys. Res. Lett. **33**: L12603. doi:10.1029/2006GL026219.

Callies, U., and others. 2017. Surface drifters in the German Bight: Model validation considering windage and stokes drift. Ocean Sci. **13**: 799–827. doi:10.5194/os-13-799-2017

Chen, C.-T. A., and A. V. Borges. 2009. Reconciling opposing views on carbon cycling in the coastal ocean: Continental shelves as sinks and near-shore ecosystems as sources of atmospheric CO₂. Deep-Sea Res. Part II Top. Stud. Oceanogr. **56**: 578–590. doi:10.1016/j.dsr2.2009.01.001

Chierici, M., A. Fransson, and Y. Nojiri. 2006. Biogeochemical processes as drivers of surface *f*CO₂ in contrasting provinces in the subarctic North Pacific Ocean. Global Biogeochem. Cycles **20**(1): 1009. doi:10.1029/2004GB002356.

Clargo, N. M., L. A. Salt, H. Thomas, and H. J. W. de Saar. 2015. Rapid increase of observed DIC and pCO(2) in the surface waters of the North Sea in the 2001-2011 decade ascribed to climate change superimposed by biological processes. Mar. Chem. **177**: 566–581. doi:10.1016/j.marchem.2015.08.010

Cooper, D. J., A. J. Watson, and R. D. Ling. 1998. Variation of pCO₂ along a North Atlantic shipping route (U.K. to the Caribbean): A year of automated observations. Mar. Chem. **60**: 147–164. doi:10.1016/S0304-4203(97)00082-0

Dai, M. H., and others. 2009. Diurnal variations of surface seawater pCO₂ in contrasting coastal environments. Limnol. Oceanogr. **54**: 735–745. doi:10.4319/lo.2009.54.3.0735

Danielssen, D. S., E. Svendsen, and M. Ostrowski. 1996. Long-term hydrographic variation in the Skagerrak based on the section Torungen–Hirtshals. ICES J. Mar. Sci. **53**: 917–925. doi:10.1006/jmsc.1996.0113

Davidson, F., and others. 2019. Synergies in operational oceanography: The intrinsic need for sustained ocean observations. Front. Mar. Sci. **6**: 450. doi:10.3389/fmars.2019.00450.

Dee, D. P., and others. 2011. The ERA-interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. **137**: 553–597. doi:10.1002/qj.828

Dickson, A., C. L. Sabine, and J. R. Christian. 2007. Guide to best practices for ocean CO₂ measurments, p. 191. PICES Special Publication.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer. 2012. The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sens. Environ. **116**: 140–158. doi:10.1016/j.rse.2010.10.017

Dumousseaud, C., and others. 2010. Contrasting effects of temperature and winter mixing on the seasonal and interannual variability of the carbonate system in the Northeast Atlantic Ocean. Biogeosciences **7**: 1481–1492. doi:10.5194/bg-7-1481-2010

Fiedler, B., P. Fietzek, N. Vieira, P. Silva, H. C. Bittig, and A. Körtzinger. 2013. In situ CO₂ and O₂ measurements on a profiling float. J. Atmos. Ocean. Technol. **30**: 112–126. doi:10.1175/JTECH-D-12-00043.1

Fietzek, P., B. Fiedler, T. Steinhoff, and A. Körtzinger. 2013. In situ quality assessment of a novel underwater pCO₂ sensor based on membrane equilibration and NDIR spectrometry.

J. Atmos. Ocean. Technol. **31**: 181–196. doi:10.1175/JTECH-D-13-00083.1.

Friedlingstein, P., and others. 2019. Global Carbon Budget 2019. Earth Syst. Sci. Data **11**: 1783–1838. doi:10.5194/essd-11-1783-2019

Fröb, F., and others. 2019. Wintertime fCO₂ variability in the subpolar North Atlantic since 2004. Geophys. Res. Lett. **46**: 1580–1590. doi:10.1029/2018GL080554

Gattuso, J. P., M. Frankignoulle, and R. Wollast. 1998. Carbon and carbonate metabolism in coastal aquatic ecosystems. Annu. Rev. Ecol. Syst. **29**: 405–434. doi:10.1146/annurev.ecolsys.29.1.405

González-Dávila, M., S. K. Lauvset, L. Coppola, and U. Schuster. 2016. Deliverable D11.3 cross-over analysis of parameters between platforms. Work Package 11 Optimisation of ocean observing capability.

Gröger, M., L. Arneborg, C. Dieterich, A. Höglund, and H. E. M. Meier. 2019. Summer hydrographic changes in the Baltic Sea, Kattegat and Skagerrak projected in an ensemble of climate scenarios downscaled with a coupled regional ocean–sea ice–atmosphere model. Clim. Dyn. **53**: 5945–5966. doi:10.1007/s00382-019-04908-9

Haller, M., F. Janssen, J. Siddorn, W. Petersen, and S. Dick. 2015. Evaluation of numerical models by FerryBox and fixed platform in situ data in the southern North Sea. Ocean Sci. **11**: 879–896. doi:10.5194/os-11-879-2015

Holt, J., and R. Proctor. 2008. The seasonal circulation and volume transport on the northwest European continental shelf: A fine-resolution model study. J. Geophys. Res. Oceans **113**: C06021. doi:10.1029/2006JC004034.

Jiang, Z.-P., and others. 2013. Key controls on the seasonal and interannual variations of the carbonate system and air-sea CO₂ flux in the Northeast Atlantic (Bay of Biscay). J. Geophys. Res. Oceans **118**: 785–800. doi:10.1002/jgrc.20087

Jiang, Z.-P., and others. 2014. Application and assessment of a membrane-based pCO₂ sensor under field and laboratory conditions. Limnol. Oceanogr.: Methods **12**: 264–280. doi:10.4319/lom.2014.12.264.

Jiang, Z.-P., J. Yuan, S. E. Hartman, and W. Fan. 2019. Enhancing the observing capacity for the surface ocean by the use of volunteer observing ship. Acta Oceanol. Sin. **38**: 114–120. doi:10.1007/s13131-019-1463-3

Juranek, L. W., R. C. Hamme, J. Kaiser, R. Wanninkhof, and P. D. Quay. 2010. Evidence of O₂ consumption in underway seawater lines: Implications for air-sea O₂ and CO₂ fluxes. Geophys. Res. Lett. **37**: L01601. doi:10.1029/2009GL040423.

Karlson, B., and others. 2016. A comparison of FerryBox data vs. monitoring data from research vessels for near surface waters of the Baltic Sea and the Kattegat. J. Mar. Syst. **162**: 98–111. doi:10.1016/j.jmarsys.2016.05.002

Kerimoglu, O., and others. 2020. Interactive impacts of meteorological and hydrological conditions on the physical and

biogeochemical structure of a coastal system. Biogeosciences **17**: 5097–5127. doi:10.5194/bg-17-5097-2020

Kitidis, V., and others. 2019. Winter weather controls net influx of atmospheric CO₂ on the north-west European shelf. Sci. Rep. **9**: 20153.

Körtzinger, A., H. Thomas, B. Schneider, N. Gronau, L. Mintrop, and J. C. Duinker. 1996. At-sea intercomparison of two newly designed underway pCO₂ systems — encouraging results. Mar. Chem. **52**: 133–145. doi:10.1016/0304-4203(95)00083-6

Kristiansen, T., and E. Aas. 2015. Water type quantification in the Skagerrak, the Kattegat and off the Jutland west coast. Oceanologia **57**: 177–195. doi:10.1016/j.oceano.2014.11.002

Laakso, L., and others. 2019. JERICO-NEXT TNA: Intercomparison of instruments for carbonate system measurements (INTERCARBO). SEANOE.

Laruelle, G. G., H. H. Dürr, C. P. Slomp, and A. V. Borges. 2010. Evaluation of sinks and sources of CO₂ in the global coastal ocean using a spatially-explicit typology of estuaries and continental shelves. Geophys. Res. Lett. **37**: L15607. doi:10.1029/2010GL043691

Laruelle, G. G., W.-J. Cai, X. Hu, N. Gruber, F. T. Mackenzie, and P. Regnier. 2018. Continental shelves as a variable but increasing global sink for atmospheric carbon dioxide. Nat. Commun. **9**: 454. doi:10.1038/s41467-017-02738-z.

Lauvset, S. K., and others. 2018. SOCAT Quality Control Cookbook -For SOCAT version 7. Available from https://www.socat.info/wp-content/uploads/2019/01/2018_SOCAT_QC_Cookbook_for_SOCAT_Version_7.pdf. Accessed 20 July 2020.

Lawrence, M. G. 2005. The relationship between relative humidity and the dewpoint temperature in moist air - a simple conversion and applications. Bull. Am. Meteorol. Soc. **86**: 225. doi:10.1175/BAMS-86-2-225

Le Quéré, C., and others. 2018. Global Carbon Budget 2017. Earth Syst. Sci. Data **10**: 405–448. doi:10.5194/essd-10-405-2018

Legge, O., and others. 2020. Carbon on the northwest European shelf: Contemporary budget and future influences. Front. Mar. Sci. **7**: 143. doi:10.3389/fmars.2020.00143.

Lorenzoni, L., M. Telszewski, H. Benway, and A. Palacz. 2017. A user's guide for selected autonomous biogeochemical sensors. *In* An outcome from the 1st IOCCP International Sensors Summer Course, June 22–July 1, 2015, Kristineberg, Sweden. IOCCP.

Lüger, H., D. W. R. Wallace, A. Körtzinger, and Y. Nojiri. 2004. The pCO₂ variability in the midlatitude North Atlantic Ocean during a full annual cycle. Global Biogeochem. Cycles **18**: 1–16. doi:10.1029/2003GB002200.

Macovei, V. A., S. E. Hartman, U. Schuster, S. Torres-Valdés, C. M. Moore, and R. J. Sanders. 2020. Impact of physical and biological processes on temporal variations of the ocean carbon sink in the mid-latitude North Atlantic

(2002–2016). Prog. Oceanogr. **180**: 102223. doi:10.1016/j. pocean.2019.102223

Newton, J. A., R. A. Feely, E. B. Jewett, P. Williamson, and J. Mathis. 2015. Global ocean acidification observing network: Requirements and governance plan. GOA-ON. Available from http://www.goa-on.org/documents/general/ GOA-ON_2nd_edition_final.pdf

Olsen, A., K. R. Brown, M. Chierici, T. Johannessen, and C. Neill. 2008. Sea-surface CO$_2$ fugacity in the subpolar North Atlantic. Biogeosciences **5**: 535–547. doi:10.5194/bg-5-535-2008

Olsen, A., N. Metzl, D. C. E. Bakker, and K. O'Brien. 2015. SOCAT quality control cookbook – for SOCAT version 3. Available from https://www.socat.info/wp-content/ uploads/2017/04/2015_SOCAT_QC_Cookbook_v3.pdf

Olsen, A., and others. 2016. The global ocean data analysis project version 2 (GLODAPv2) - an internally consistent data product for the world ocean. Earth Syst. Sci. Data **8**: 297–323. doi:10.5194/essd-8-297-2016

Omar, A. M., A. Olsen, T. Johannessen, M. Hoppema, H. Thomas, and A. V. Borges. 2010. Spatiotemporal variations of fCO$_2$ in the North Sea. Ocean Sci. **6**: 77–89. doi:10.5194/ os-6-77-2010

Omar, A. M., H. Thomas, A. Olsen, M. Becker, I. Skjelvan, and G. Reverdin. 2019. Trends of ocean acidification and pCO$_2$ in the northern North Sea, 2003–2015. J. Geophys. Res. Biogeosci. **124**: 3088–3103. doi:10.1029/2018JG004992.

Otto, L., J. T. F. Zimmerman, G. K. Furnes, M. Mork, R. Saetre, and G. Becker. 1990. Review of the physical oceanography of the North Sea. Neth. J. Sea Res. **26**: 161–238. doi:10. 1016/0077-7579(90)90091-T

Petersen, W. 2014. FerryBox systems: State-of-the-art in Europe and future development. J. Mar. Syst. **140**: 4–12. doi:10.1016/j.jmarsys.2014.07.003

Petersen, W., S. Reinke, G. Breitbach, M. Petschatnikov, H. Wehde, and H. Thomas. 2018. FerryBox data in the North Sea from 2002 to 2005. Earth Syst. Sci. Data **10**: 1729–1734. doi:10.5194/essd-10-1729-2018

Pierrot, D., and others. 2009. Recommendations for autonomous underway pCO$_2$ measuring systems and data-reduction routines. Deep-Sea Res. Part II Top. Stud. Oceanogr. **56**: 512–522. doi:10.1016/j.dsr2.2008.12.005

Ribas-Ribas, M., and others. 2014. Intercomparison of carbonate chemistry measurements on a cruise in northwestern European shelf seas. Biogeosciences **11**: 4339–4355. doi:10. 5194/bg-11-4339-2014

Rodhe, J. 1996. On the dynamics of the large-scale circulation of the skagerrak. J. Sea Res. **35**: 9–21. doi:10.1016/S1385-1101(96)90731-5.

Roobaert, A., G. G. Laruelle, P. Landschutzer, N. Gruber, L. Chou, and P. Regnier. 2019. The spatiotemporal dynamics of the sources and sinks of CO$_2$ in the global coastal ocean. Global Biogeochem. Cycles **33**: 1693–1714. doi:10.1029/ 2019GB006239

Salt, L. A., H. Thomas, A. E. F. Prowe, A. V. Borges, Y. Bozec, and H. J. W. de Baar. 2013. Variability of North Sea pH and CO$_2$ in response to North Atlantic oscillation forcing. J. Geophys. Res. Biogeosci. **118**: 1584–1592. doi:10.1002/ 2013JG002306

Schlitzer, R. 2020. Ocean data view. Available from http://odv. awi.de

Schuster, U., and A. J. Watson. 2007. A variable and decreasing sink for atmospheric CO$_2$ in the North Atlantic. J. Geophys. Res. Oceans **112**: C11006. doi:10.1029/ 2006JC003941.

Takahashi, T., J. Olafsson, J. G. Goddard, D. W. Chipman, and S. C. Sutherland. 1993. Seasonal variation of CO$_2$ and nutrients in the high-latitude surface oceans: A comparative study. Global Biogeochem. Cycles **7**: 843–878. doi:10.1029/ 93GB02263

Takahashi, T., and others. 2002. Global sea–air CO$_2$ flux based on climatological surface ocean pCO$_2$, and seasonal biological and temperature effects. Deep-Sea Res. Part II Top. Stud. Oceanogr. **49**: 1601–1622. doi:10.1016/S0967-0645(02) 00003-6

Takahashi, T., and others. 2009. Climatological mean and decadal change in surface ocean pCO$_2$, and net sea–air CO$_2$ flux over the global oceans. Deep-Sea Res. Part II Top. Stud. Oceanogr. **56**: 554–577. doi:10.1016/j.dsr2.2008.12.009

Tamburri, M. N., and others. 2011. Alliance for coastal technologies: Advancing moored pCO$_2$ instruments in coastal waters. Mar. Technol. Soc. J. **45**: 43–51. doi:10.4031/MTSJ. 45.1.4

Thomas, H., Y. Bozec, K. Elkalay, and H. J. W. de Baar. 2004. Enhanced open ocean storage of CO$_2$ from shelf sea pumping. Science **304**: 1005–1008. doi:10.1126/science. 1095491

Thomas, H., and others. 2005. The carbon budget of the North Sea. Biogeosciences **2**: 87–96. doi:10.5194/bg-2-87-2005

Tortell, P. D., and others. 2014. Metabolic balance of coastal Antarctic waters revealed by autonomous pCO$_2$ and ΔO$_2$/ Ar measurements. Geophys. Res. Lett. **41**: 6803–6810. doi: 10.1002/2014GL061266

Tsunogai, S., S. Watanabe, and T. Sato. 1999. Is there a "continental shelf pump" for the absorption of atmospheric CO$_2$? Tellus B **51**: 701–712. doi:10.1034/j.1600-0889.1999.t01-2-00010.x.

van Heuven, S., D. Pierrot, J. W. B. Rae, E. Lewis, and D. W. R. Wallace. 2011. CO2SYS v 1.1 MATLAB program developed for CO$_2$ system calculations. ORNL/CDIAC-105b. Oak Ridge National Laboratory.

van Leeuwen, S., P. Tett, D. Mills, and J. van der Molen. 2015. Stratified and nonstratified areas in the North Sea: Long-term variability and biological and policy implications. J. Geophys. Res. Oceans **120**: 4670–4686. doi:10.1002/ 2014JC010485

Voynova, Y. G., W. Petersen, M. Gehrung, S. Aßmann, and A. L. King. 2019. Intertidal regions changing coastal

alkalinity: The Wadden Sea-North Sea tidally coupled biore-actor. Limnol. Oceanogr. **64**: 1135–1149. doi:10.1002/lno.11103.

Wakelin, S. L., J. T. Holt, J. C. Blackford, J. I. Allen, M. Butenschön, and Y. Artioli. 2012. Modeling the carbon fluxes of the northwest European continental shelf: Validation and budgets. J. Geophys. Res. Oceans **117**: C05020. doi:10.1029/2011JC007402.

Wanninkhof, R., J. Triñanes, G.-H. Park, D. Gledhill, and A. Olsen. 2019. Large decadal changes in air-sea CO_2 fluxes in the Caribbean Sea. J. Geophys. Res. Oceans **124**: 6960–6982. doi:10.1029/2019JC015366

World Data Centre for Greenhouse Gases. 2020. Mace head xCO_2. Available from https://gaw.kishou.go.jp/. Accessed 20 July 2020.

Conflict of Interest

None declared.