

Supplementary Material

Metrological checks for stability assessment of probe and sensors

In the following, a description of the control procedures applied at MZS to each employed probe is presented.

SeaFET - The assessment of SeaFET, in terms of both accuracy and stability, is a well-known issue that has given rise to the development of various metrological methods over the past few years [79,80]. During the present project, an *ad-hoc* procedure was developed to check the stability of pH_r_EXT sensor during the one-year deployment period of the SeaFET. This procedure, based on pH comparative measurements with a different pH probe, used as stability-reference standard (a Mettler Toledo, SG2 potentiometric model glass electrode sensor, measuring pH in NBS (National Bureau of Standards, now National Institute of Standards and Technology (NIST)) scale [81]), can be summarized in the following steps:

1. pre-deployment operations (period 07th–20th of November 2018): a total number of 40 pH measurements on various 200 μm -filtered natural seawater samples, at atmospheric pressure and mean temperature t_{lab} of 21.8 °C (s.d. of mean: ± 0.4 (°C)), were performed on bench in the MZS laboratory. The pH measurements were carried out in cascade between both the pH-meters, SeaFET and SG2 respectively (after at least a 24-hour period of stabilization for SeaFET measuring a new sample). Regarding SG2 sensor, (for which an error limit equal to ± 0.01 pH is declared by the manufacturer), prior the use of seawater, a three-points daily calibration was performed through a proper set of NBS certified calibration buffers for pH values equal to 4.01, 7.00 and 9.21 (at 25 (°C), with an accuracy declared equal to ± 0.02 pH). Values of pH measured by the SeaFET EXT sensor were corrected for salinity measured by an Hach-HQ-30d probe on each sample;
1. post-deployment operations (28th of November 2019): a total number of 20 pH simultaneous measurements were carried out by SeaFET and SG2, both immersed in a tank filled with seawater continuously pumped directly from the sea, at atmospheric pressure and mean temperature of -0.858 °C (s.d. of mean: ± 0.002 (°C)). SeaFET EXT sensor measures were corrected for salinity measured by the CTD used in conjunction with SeaFET. Values of pH measured at both scales (Total and NBS) were then reported to t_{lab} by means of CO2Calc software (v. 4.0.9, U.S. Geological Survey [42]), set according to the following calculation preferences:
 - CO₂-system dissociation constants (K1 and K2): in analogy with literature [31,82], constants estimated by Roy and coauthors [43] were used, having turned out as the most suitable for cold and surface waters with low salinity;
 - Total Boron: reference to Uppstrom [44];
 - dissociation constant for potassium sulfate (KHSO₄): reference to Dickson [45];

- total alkalinity (TA): calculated by applying well-known empirical relationship between CTD-measured sea surface temperature and practical salinity, overall valid for the Southern Ocean [36]. By applying a more localized empirical formula [31], more suitable for Ross Sea, no significant variations of pH values reported to t_{lab} was noticed.

The aim of the adopted procedure was to verify the possible drift over time of the mean difference pH between the values of pH_T (measured by SeaFET) and pH_{NBS} (measured by SG2), actually not directly comparable each other [83]: being SG2 daily calibrated with NBS calibration buffers during pre- and post-deployment comparative tests, and so used as a stable reference, any changes found between the values of the two scales would have been attributable to the drift of the SeaFET itself. This method, clearly, didn't allow an absolute calibration of the SeaFET probe (for which the manufacturer calibration was used), but allowed its control in terms of stability, using the SG2 output as a reference for each seawater sample analyzed at the same pressure and temperature conditions (atmospheric pressure and t_{lab} , respectively).

Comparative tests led to the following mean value of $\Delta pH = pH_{T,SeaFET} - pH_{NBS,SG2}$ (where s.d. of mean is indicated, together with the sample numerosity n):

- pre-deployment: $\Delta pH = 0.062 \pm 0.025$ ($n = 40$)
- post-deployment: $\Delta pH = 0.037 \pm 0.001$ ($n = 20$).

The post deployment tests were performed on a smaller number of simultaneous pH measurements due to the greater stability of the environmental conditions guaranteed by the pumped tank, comparable to a thermostatic bath, that is evidenced by the lower dispersion of values.

Diagram in Figure S1 shows how the mean post-deployment ΔpH value resulted reasonably comparable with the pre-deployment one, taking also into account that both values were contained into the band drift declared by the manufacturer (whose estimated value after 12 months of use is equal to about ± 0.06 pH_T unit). As a further proof of the compatibility between mean ΔpH values a year later, two sample t-Test [84] was performed on the two samples of data under study (pre- and post-deployment); with the significance level set at 0.05, the null hypothesis ($\Delta pH_{post} = \Delta pH_{pre}$) is accepted, i.e. the difference of the populations means is not significantly different. Finally, the hourly drift estimates of the pH values measured by the SeaFET was given by:

- $drift(pH_T) = (\Delta pH)_{post} - (\Delta pH)_{pre} \approx -3E-6$ pH_T h^{-1}

Therefore, each hourly pH_T measure made by the SeaFET during the deployment was corrected for this drift (although at a first sight it seems negligible, such a correction after about 362 days of deployment resulted in a positive variation of about 0.02 pH_T unit). The obtained drift value for the SeaFET was extremely limited thanks to two main reasons:

1. the almost total absence of bio-fouling, typical of the environment in which the experimental station was deployed, as it can be seen in Figure S2;
2. SeaFET actively flushed (i.e., pumped, as in this case) showed “far greater stability than passively flushed packages, with the former remaining stable in a wide variety of environments on timescales approaching one year...” [79]. Besides, antifoulant devices were used, directly mounted in the intake and out-take of the hydraulic circuit.

For what concerns the pressure effect suffered by SeaFET [79], a test was performed during its recovery from the bottom: SeaFET and CTD were left deployed at a depth of 4 m for 20 hours and mean values of all quantities of interest were compared with mean values at 25 m depth (acquired immediately before, on the same time period). Results are shown in Figure S3. As expected, decreasing pressure results in an increasing pH signal: being the difference considerably restrained (about 0.01 pH unit), and considering that at the same time the seawater itself showed reasonably different values of both temperature and density, it was not possible to deduce and apply any type of correction for the pressurizing effect of the SeaFET EXT sensor.

SBE37-SMP-ODO - Regarding the dissolved oxygen measurements, the factory specifications (manufacturer calibration document of January-February 2018) were followed. The stability of the CTD probe (in terms of p , t and C measurements) was instead verified through the following steps:

1. at the end of the deployment period (in the period 8–12 November 2019), before recovering the probe, 10 seawater samples (1 l each) were collected using a Go-Flo bottle from the hole in the ice-pack near the submerged structure, at the same depth: date and time of the seawater sample collection were marked in order to compare the subsequent laboratory measurements with the data acquired by the submersed probe;
2. in a short time (less than 1 h) each seawater sample was brought to the MZS laboratory and subjected to the following measurements: p_{NBS} and t_{situ} (by SG2 probe), C by the Hach-HQ-30d probe;
3. starting from these measurements, and using the SBE SeaCalc III software (v. 3.0), the value of σ_t of each sample was calculated (imposing $p = 0$ (dbar));
4. after the recovery of both CTD and SeaFET probes, t and S_p (from which TA was calculated in accordance with [41]), p and p_{Hr} data were downloaded in correspondence of date and time of each seawater samples;
5. from this dataset, and by means of CO2Calc software, the density reported at t_{situ} (and $p = 0$ (dbar)) was recalculated for each seawater sample: in this way the two values of σ_t (measured respectively in the laboratory and at the submersed station) were reported under the same environmental conditions and as such could be compared.

By applying the procedure described above to the ten seawater samples, a relative mean difference equal to about 0.08 % was calculated between the two density measurements. This result allowed to conclude that, during the one-year deployment, no significant drifts in the triad of parameters measured by the CTD probe could be found.

HOBO - The stability of these sensors over time was not verified. It should be emphasized that light intensity values were not used as absolute measures but rather at a qualitative level (e.g. as indicators of absence, presence or fragmentation of the ice-pack).

Figures and figure captions

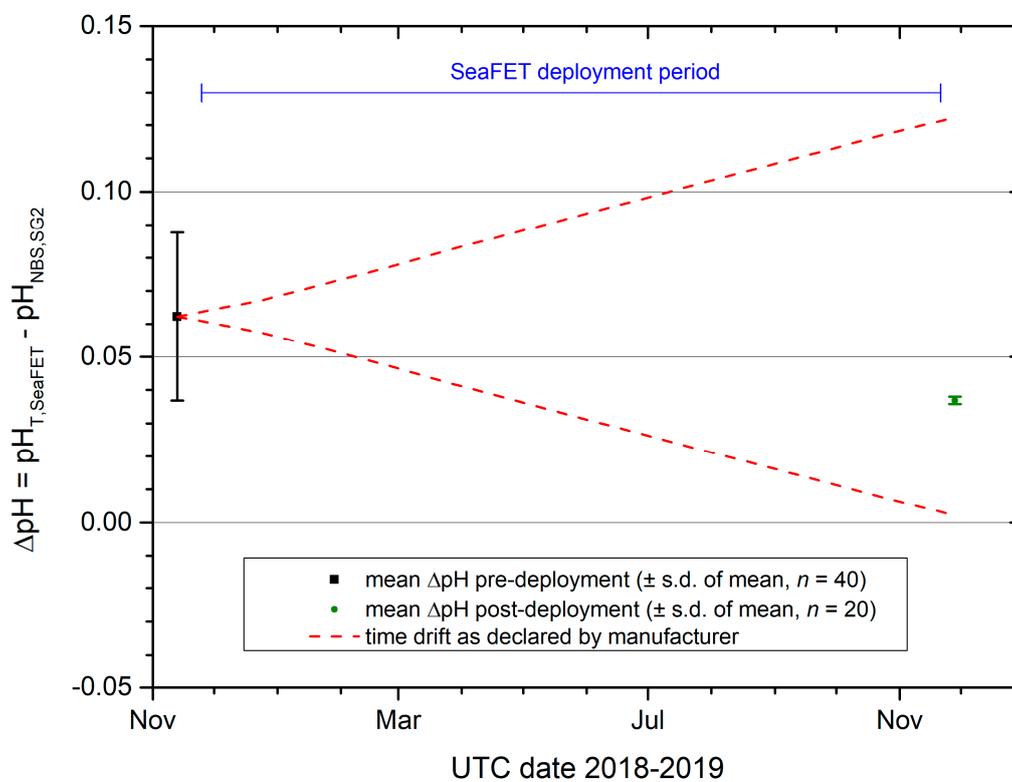


Figure S1. Evaluation of SeaFET annual drift: comparison between mean differences of pH measures performed vs the reference pH-meter (SG2) pre- and post-deployment, respectively.

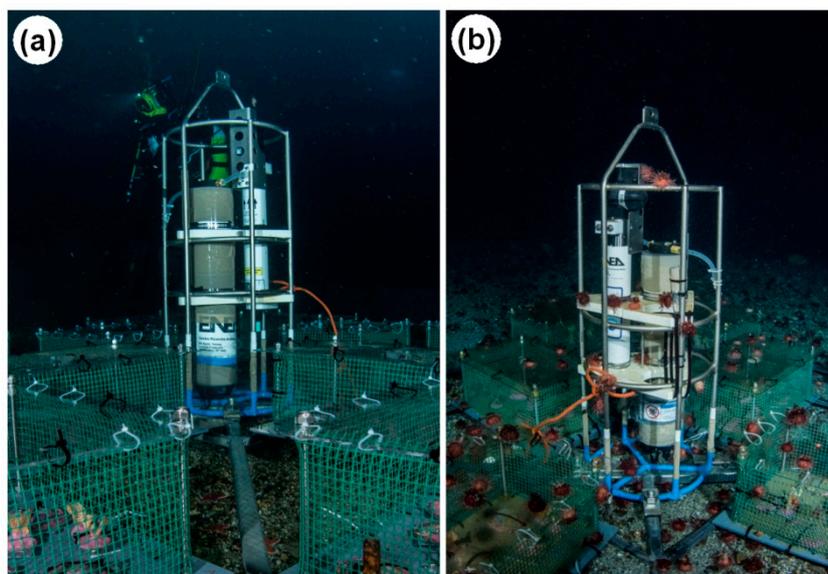


Figure S2. Images of the pre- (a) and post-deployment (b) multi-parametric probes (SeaFET coupled to CTD): no significant evidence of fouling after one year at 25 m depth.

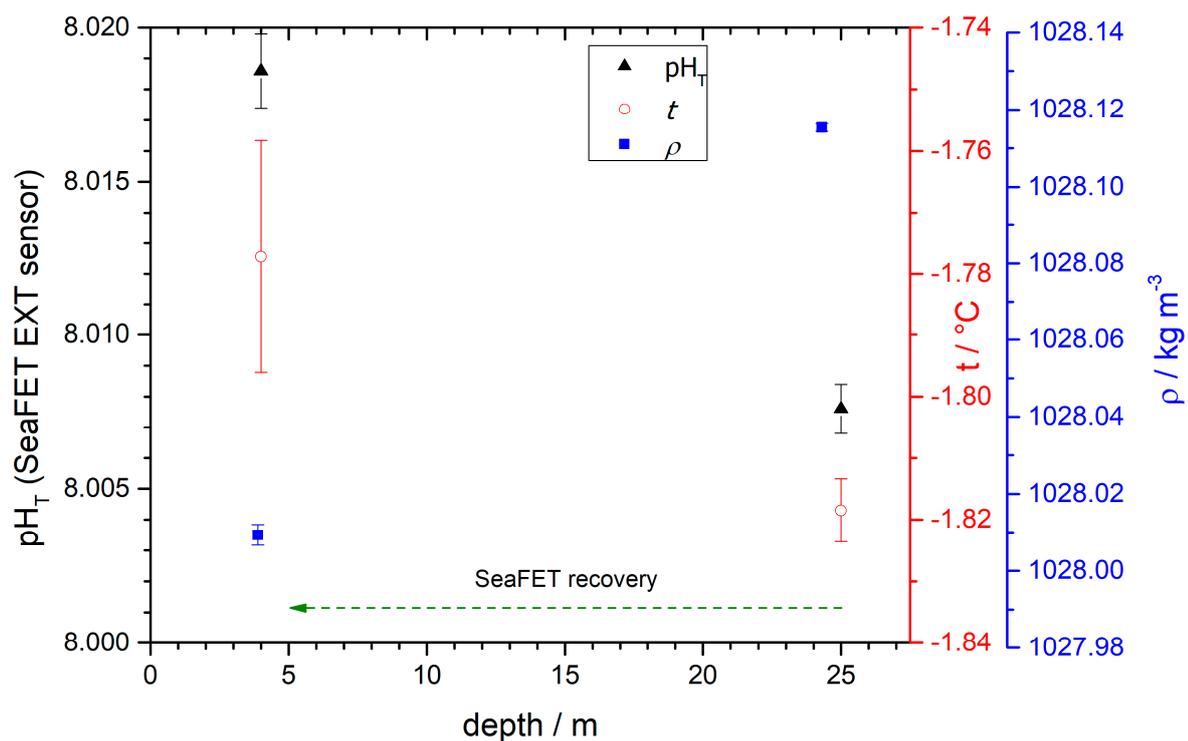


Figure S3. SeaFET (pH_T) and CTD (t , ρ) mean measures pre- and post a pressure decrease, corresponding to a change in depth of about 21 m. Mean measures ($n = 20$) are reported, together with their s.d.

References:

31. Rivaro, P.; Messa, R.; Ianni, C.; Magi, E.; Budillon, G. Distribution of total alkalinity and pH in the Ross Sea (Antarctica) waters during austral summer. *Polar Res.* **2014**, *33*, 20403, doi:10.3402/polar.v33.20403.

36. Umami, S.F.; Monti, M.; Bergamasco, A.; Cabrini, M.; De Vittor, C.; Burba, N.; Del Negro, P. Plankton community structure and dynamics versus physical structure from Terra Nova Bay to Ross Ice Shelf (Antarctica). *J. Mar. Syst.* **2005**, *55*, 31–46, doi:10.1016/j.jmarsys.2004.05.030.
42. Robbins, L.L.; Hansen, M.E.; Kleypas, J.A.; Meylan, S.C. CO2calc—A User-friendly Seawater Carbon Calculator for Windows, Max OS X, and iOS (iPhone). U.S. Geological Survey: St. Petersburg, FL, USA, 2010, 2010–1280.
43. Roy, R.N.; Roy, L.N.; Vogel, K.M.; Porter-Moore, C.; Pearson, T.; E Good, C.; Millero, F.J.; Campbell, D.M. The dissociation constants of carbonic acid in seawater at salinities 5 to 45 and temperatures 0 to 45°C. *Mar. Chem.* **1993**, *44*, 249–267, doi:10.1016/0304-4203(93)90207-5.
44. Uppström, L.R. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. *Deep. Sea Res. Oceanogr. Abstr.* **1974**, *21*, 161–162, doi:10.1016/0011-7471(74)90074-6.
79. Bresnahan, P.J.; Martz, T.R.; Takeshita, Y.; Johnson, K.S.; LaShomb, M. Best practices for autonomous measurement of seawater pH with the Honeywell Durafet. *Methods Oceanogr.* **2014**, *9*, 44–60, doi:10.1016/j.mio.2014.08.003.
80. Miller, C.A.; Pocock, K.; Evans, W.; Kelley, A.L. An evaluation of the performance of Sea-Bird Scientific's SeaFET™ auto-nomous pH sensor: Considerations for the broader oceanographic community. *Ocean Sci.* **2018**, *14*, 751–768.
81. Pérez, F.F.; Fraga, F. The pH measurements in seawater on the NBS scale. *Mar. Chem.* **1987**, *21*, 315–327, doi:10.1016/0304-4203(87)90054-5.
82. Chierici, M.; Fransson, A. Calcium carbonate saturation in the surface water of the Arctic Ocean: Undersaturation in freshwater influenced shelves. *Biogeosciences* **2009**, *6*, 2421–2431, doi:10.5194/bg-6-2421-2009.
83. Luchetta, A.; Cantoni C.; Catalano, G. New observations of CO₂-induced acidification in the northern Adriatic Sea over the last 696 quarter century. *Chem Ecol*, **2010**, *26*, 1-17.
84. Hoel, P.G. Introduction to Mathematical Statistics. 5th ed. New York: Wiley, 1984.