Anonymous Referee #2, 14 Nov 2025

General Comments

I really enjoyed reading this paper. It is well written, thoughtfully laid out, and has incorporated many rigorous experiments explained in significant detail. I liked the background information provided for the A25 section and felt it gave good context to the remainder of the manuscript. This large pHT dataset will be a great resource for researchers. I particularly commend the authors for including the absorbance values in their pHT dataset, which the authors correctly point out will be useful in the future if pHT characterizations are updated. I also appreciate that the authors put these pHT revisions in the context of the larger scale OA trends. I recommend acceptance of the manuscript after some minor revisions.

We would like to thank Anonymous Referee #2 for the thoughtful review and the constructive feedback provided. We are pleased that the referee found our work valuable and engaging. All comments have been carefully considered, as detailed in the responses below.

Anonymous Reviewer #2's evaluation is reproduced in black, *the author's responses appear in green and italics,* *the original manuscript text appears in black and italics,* *and the changes introduced in the manuscript are indicated in blue and italics.*

Specific Comments

Lines 74-75: pHT calculated from AT and CT is known to have potential issues, such as the pH-dependent pH offset and larger uncertainty. I would argue that this sentence should be rephrased to convey that both (1) pHT reported with no additional internal consistency corrections AND (2) pHT calculated from AT and CT could both affect the reliability of OA analyses. The key is having directly measured pHT that can be "trusted".

Thank you for the appropriate suggestion, we agree with your comment. We have rephrased the text as follows: "A_T and C_T measurements are generally considered more reliable due to the availability of standardized reference materials, consensually accepted methods, and quality control procedures. In contrast, although pH_T measurements are technically precise, easy to perform, and cost-effective, their intercomparability is more challenging, arising from methodological inconsistencies across various research initiatives (Dickson et al., 2015; Ma et al., 2019; Álvarez et al., 2020; Capitaine et al., 2023). The lack of traceability to a common reference, preferentially the International System of units for spectrophotometric pH measurements (Dickson et al., 2015), and the unavailability of pH reference materials within the seawater pH range (Capitaine et al., 2023), together with the documented issues affecting pH_T calculated from A_T and C_T (including pH_T-dependent offsets and larger propagated uncertainties; Álvarez et al., 2020; Carter et al., 2024b), mean that neither unadjusted direct observations nor calculated values currently provide a fully trusted global reference. Both limitations may therefore affect the reliability of pH data for climate-quality OA assessments. ~~*These factors contribute to the relative scarcity of pH_T measurements in GLODAPv2, and lead to the future GLODAPv3 providing pH_T data*~~

~~as originally reported with no additional internal consistency corrections, a limitation that may potentially affect the reliability of OA analysis from directly measured pH$_T$ data instead of those calculated from A$_T$ and C$_T$ (Álvarez et al., 2020; Carter et al., 2024b)~~."

Lines 78-79: The pH method was detailed previous to the 1990s (Robert-Baldo et al. 1985; Byrne and Breland, 1989), using mCP as the indicator dye began in the 1990s.

Thank you. We have rephrased the text as follows: "Briefly, the spectrophotometric pH method is a straightforward technique that involves adding an acid-base indicator dye~~, usually meta-cresol purple (mCP),~~ to the seawater sample. The method was initially defined in the 1980s (Robert-Baldo et al., 1985; Byrne and Breland 1989), and the use of meta-cresol purple (mCP) as the indicator dye began ~~The method was defined~~ in the 1990s (Clayton & Byrne, 1993). The technique has been updated since then, ~~and~~ but it still lacks metrological traceability and reference materials (Ma et al., 2019; Carter et al., 2024a)."

Line 100: Was this the same lot of dye used for all 11 cruises? Or were different batches of dye made over time, just with repeat purchases from the same brand?

Only a few different batches of indicator dye from the same brand were used over the 11 cruises. For example, the same lot of indicator dye was used for the last four cruises. We have included it in Sect. 2.2.3 as follows: "Throughout the OVIDE-BOCATS program, the mCP dye used was from Sigma-Aldrich (Cat. No. 11,436-7 in the basic form; $C_{21}H_{17}NaO_5S$; molecular weight 404.41 g), with only a few different batches from this brand used over the 11 cruises."

Eq. 1: Need to define what pK2 is for this equation. You should also make it clear that this pK2 is different from the pK2 defined in Clayton and Byrne (1993). They list their K1 and K2 as formation constants, whereas the K2 in this manuscript is a dissociation constant. Their K2 describes the formation of H2I, whereas the K2 in this manuscript describes the dissociation to form H+ and I2-. **Lines 171-172:** Again, make clear that the pK2 shown here is actually labeled as log K1 (so – pK1) in the CB'93 paper. I understand why you are showing yours as dissociation constants, but since you are referencing CB'93, you need to make it clear that they are not the same.

Thank you for pointing this out. We agree, and have clarified it in the manuscript as follows: "The first three terms of Eq. (2) represent the second dissociation constant of mCP dye (pK$_2$, or -pK$_1$ in CB'93). The -pK$_1$ ~~pK$_2$~~ reported by CB'93 is based on the TRIS (...)".

Lines 171-173: This sentence is almost identical to a sentence in Lee et al. (2000) and should be properly cited, not just citing the last sentence of the paragraph.

Done, thank you.

Line 178: Would be useful to mention here that the mCP in CB'93 was also made in DI water. Which is different than what is done in this study (and described in detail later).

We agree, and we have included the clarification as follows: "The CB'93 parameterization was developed using Kodak mCP dye, prepared in deionized water, which contained impurities contributing~~presenting~~ significant absorbance at 434 nm (referred to as $_{434}A_{imp}$)."

Line 185: I think this value should be 0.018 from L'11.

Corrected, thank you.

Line 205: Since you're referencing eq. (11) of DB'17, you should note that your "Runpur" is noted as "Robs" in their paper.

We agree, thank you. We have revised the text as follows: "Both the purified R values and their associated $_{434}A$ values were then used in Eq. (11) of DB'17 to compute the adjusted R values (R_{unpur}; referred to as R_{obs} in DB'17) that reflect the contribution of mCP dye impurities ($_{434}A_{imp} \neq 0$) as follows: (...)".

Fig. 2: It needs to be made clear that these 434Aimp values are specific to one lot from each of these vendors, which are listed in the CB'93 paper. These 434Aimp values aren't generic for all lots of impure mCP with the same vendor name.

Yes, thank you. We have clarified this in the Figure 2 caption as follows: "(...) Each mCP dye is represented by a different color, with its corresponding $_{434}A_{imp}$ value (in units of 10^{-3} absorbance) specific to the lot used in DB'17, listed after the indicator dye name. Bold indicates mCP dye brands discussed in this work. All $_{434}A_{imp}$ values are taken from DB'17 and are specific to the lot used, except for Sigma-Aldrich, whose value was determined in this study."

Fig. 2: What is the purified line showing? Is this using the same Runpur and comparing the CB'93 and L'11 parameterizations? It's difficult to follow the text in lines 220-222 describing this.

Thank you for pointing out this confusion. We apologize, there was a typo in those lines. The 'purified' line shows the pH difference resulting from applying the two different parametrizations (CB'93 and L'11) to a purified mCP dye. No, it does not use the same Runpur; rather, the same Rpur is applied. In contrast, the other lines represent the difference between the pH computed using Runpur with CB'93 and Rpur with L'11. We have corrected the text as follows: "When the CB'93 parameterization is applied to $R_{\cancel{un}pur}$ ($_{434}A_{imp} = 0$) at $S = 35$ and 25ºC, the largest theoretical pH_T differences (> 0.015 pH_T units) are observed relative to pH_T values obtained by applying the L'11 parametrization to the corresponding R_{pur} under the same conditions (see turquoise line in Fig. 2)"

Line 219: List the concentration for the mCP dyes determined in this study.

We have included the final indicator dye concentration in the sample cell in the first and third experiment sections (not in the second one, since it is the same as in the first) to facilitate comparison with the concentration used by DB'17. A detailed response is provided to the comment below regarding "line 425".

Line 240: Clarify that the blank measurement were recorded at the three target wavelengths (434, 578, 730 nm) like you do for the dye-addition measurements (line 246).

We have clarified the text as follows: "For each sample, a blank measurement was performed after drying and cleaning both faces of the optical cell and placing it in the spectrophotometer's cell holder. Following blanking at the three target wavelengths (434 nm, 578 nm, and 730 nm) with sampled seawater, 75 µL of 2 mM mCP dye solution were added to each 28 mL sample cell using an adjustable repeater pipette (SOCOREX), resulting in a final mCP dye concentration of 5.36 µM in the cell. Dispenser syringes were wrapped in aluminum foil to prevent photodegradation of the mCP dye (Fontela et al., 2023). After the mCP dye addition, the cell was thoroughly shaken and placed back in the holder in~~maintaining~~ the same orientation as for the blanking, and triplicate absorbance readings were carried out at the same target wavelengths as for the blank.~~Triplicate absorbance measurements were carried out at the three target wavelengths (434 nm, 578 nm, and 730 nm).~~ All absorbance readings were carried out in the spectrophotometer's thermostatted cell compartment, maintained at 25.0 ± 0.2ºC."

Lines 249-252: It should be clarified that measurements of R are insensitive to slight temperature changes (which is what Byrne and Breland 1989 noted). Also, Byrne and Breland were using cresol red, not mCP. However, pH is quite sensitive to temperature so accurate T measurements are necessary.

Thank you for the clarification. We have rewritten the text as follows: "Byrne & Breland (1989) demonstrated that ~~pH_T~~ R ~~determinations~~ measurements are largely insensitive to small temperature variations, for cresol red dye.~~;specifically,~~ The same general behaviour applies to mCP dye, in which R values between 0.~~7~~5 and 2.8 exhibit pH_T errors of less than 0.001 pH_T units per 0.5~~1~~°C change when using the ~~CB'93~~ L'11 parameterization. This insensitivity arises because, for mCP, the temperature dependence of the indicator dye's pK_2 ~~of the mCP dye~~ closely parallels ~~matches~~ the temperature dependence~~effect~~ of~~n~~ seawater pH. The CB'93 pH_T parametrization shows a slightly greater temperature sensitivity, such that temperature deviations must be kept within approximately ±0.5°C to limit the pH_T error to ≤0.0012 pH_T units. Therefore,~~ Nevertheless~~, it is recommended to ensure that temperature deviations remain within ±0.5°C of the reference temperature (25°C). In our procedure, sample temperature was monitored every five measurements to verify~~confirm~~ that it remained within this tolerance ~~range.~~"

Line 264: Specify that an R of 1 corresponds to a pHT of 7.67 at what salinity and temperature?

We have specified this as follows: "(…) R values remained close to 1, corresponding to a pH_T of approximately 7.67 (Li et al., 2020; at $S = 35$ and $T = 25$ºC)."

Line 279: Can authors justify why their double-dye additions used 50 uL increments instead of 75 uL like their experiments? The authors account for the difference correctly in their equation 6, but I think readers will wonder the reasoning.

The 75 µL addition corresponds to the center or intermediate range of the deltaR adjustment. Using 50 µL increments allowed us to examine the indicator dye effect both "behind" and "in front" of the 75 µL reference point, providing a clearer understanding of the response while still correctly accounting for the difference in Eq. 6. We have justified this in the manuscript as follows: "Following blanking, an initial addition of 50 µL of mCP dye solution was made to each sample, and absorbance was measured as described in Section 2.2.2. A second addition of 50 µL of mCP dye solution was then made (resulting in a total of 100 µL of mCP dye solution added), and absorbance measurements were repeated. These double-addition experiments enabled the determination of linear regressions of the change in pH_T ($\Delta pH_T = pH_{T,2} - pH_{T,1}$; where subscripts 2 and 1 refer to 100 µL and 50 µL of mCP dye solution added, respectively) or R ($\Delta R = R_2 - R_1$) as a function of the initial $pH_{T,1}$ or R_1, respectively (Supporting Information Fig. 1). This two-step 50 µL addition bracketed the typical 75 µL reference volume added to the sample, allowing us to evaluate the dye effect on the ΔR or ΔpH_T both below and above this reference. The corresponding relationship was expressed as: (...)".

Line 297: Might be useful to add a clarifying sentence here to explain that at pHT y=0, the R (pH) of the original sample and the R (pH) of the indicator are the same, which is why no change is observed.

Thank you. We have clarified the text as follows: "At $pH_T^{y=0}$, the R (pH_T) of the original sample and the R (pH_T) of the indicator dye are the same, so no change is observed. If $pH_m > pH_T^{y=0}$, then (...)".

Line 307: For a couple of the cruises, normalizing the delta R with the isosbestic point reduced the R2. Can the authors explain why that would be?

We apologize for the confusion. Upon checking the computations, we identified a few typos when transferring the data for the 2018 cruise, which have now been corrected. Fortunately, these do not affect the equations applied to the data (last column). In contrast, the 2014 cruise equations are correct, and the slightly lower R^2 for some cases arises from an anomalous average of the indicator dye effect perturbation at high pH values.

Line 339: R=1 (or pHT=7.65) at what temperature and salinity?

We have clarified this as follows: "*In contrast, deviations from R = 1 (or $pH_T = 7.65$; at S = 35 and T = 25°C) enhance this (...)".*

Fig. 3: Also need to specify which temperature condition.

Done, thank you.

Lines 368-369: How did the researchers determine when the volume deviated by more than 20%?

We determined deviations by monitoring the absorbance at 488 nm (measured or computed). For each cruise, the mean A488 was computed across all samples measured in the field layer *(approximately 2,000 per cruise), providing a robust reference value. Any sample whose A488 deviated by more than 20% from this mean was considered to exceed the 20% threshold. This procedure was repeated for all 11 OVIDE-BOCATS cruises. We have clarified in the text as follows: "During the OVIDE-BOCATS cruises, mCP dye was manually added to samples using an adjustable repeater pipette (see Section 2.2.2). For each cruise, volume deviations associated with manual addition were assessed by comparing each sample to the cruise-specific mean $_{488}A$, computed from all field-layer samples (~2000 per cruise). Manual addition ~~introduced~~resulted in volume deviations exceeding 20% in approximately 3% of the samples (~ 706 cases), potentially affecting ΔR and ΔpH_T determinations. To address variability in the volume of mCP dye solution added, (...) ".*

Lines 395-396: Authors should note that some uncertainty is introduced by collecting duplicates on 2 different Niskin bottles, rather than 2 samples from the same Niskin. (i.e., small leaking in one bottle, biological activity, delay in closing between Niskins, etc.)

We have clarified this as follows: "Throughout the 11 OVIDE-BOCATS cruises, a total of 502 duplicate samples were collected to evaluate the reproducibility of pH_T measurements using an unpurified ~~impure~~ mCP dye. At selected stations, two Niskin bottles were closed at the same depth to obtain replicates. Any uncertainty introduced by collecting duplicates on two different Niskin bottles (e.g., small leaks, biological activity, or delay in closing) was neglected. Figure 4 displays the absolute pH_T differences between replicates for each cruise. The overall mean and standard deviation of these differences is 0.0014 ± 0.0015 pH_T units (N = 502)."

Fig. 4: The grey shading is useful but hard to see. Can the authors make it slightly darker? Min and max values are also very small and hard to read.

Done, thank you; and we applied the same recommendations to Fig. 8.

Line 425: Would be useful to include the final dye concentration in the sample cell for the 2 dye types for a direct comparison, since the experiments used different volumes and stock indicator concentrations.

Thank you for the suggestion. We agree and have included the final indicator dye concentration in the sample cell in the first and third experiment sections (not in the second one, as it is the same as in the first) as follows:

"3.1 TRIS buffer validation

(...) Multiple bottles from both batches were measured during the BOCATS2-2021 and BOCATS2-2023 cruises using two mCP dye solutions: (i) unpurified mCP (75 µL of 2 mM solution; Sigma-Aldrich; 5.36 µM final mCP dye concentration in the sample cell), and (ii) purified mCP (10 µL of 11 mM solution; provided by Prof. R. Byrne's laboratory; 3.93 µM final mCP dye concentration in the sample cell). (...)

3.3 Duplicate measurements of modified seawater

(...) For each batch, four to six samples were analyzed for pH_T using the standard OVIDE-BOCATS procedure (CB'93+DVD'98 and 75 µL of unpurified Sigma-Aldrich mCP dye [2 mM] solution added to the cell; 5.36 µM final mCP dye concentration in the sample cell), and an equivalent number was measured using purified mCP dye (75 µL of mCP dye [2.5 mM] solution added to the cell; 6.70 µM final mCP dye concentration in the sample cell; FB5-2017 from Dr. Byrne's laboratory and the L'11 parametrization). (...) ".

Fig. 5: Might be useful to reorganize the figure so that the purified dye column is on the right-most side. The unpurified dye columns could then all be shown together as increasingly improving the results to most closely match the purified dye results.

Done, thank you.

Fig. 6: The "N=176" within the figure seems unnecessary, since it's listed in the figure caption. Its placement is confusing since it makes it appear that the "N=176" is specifically referring to the middle column of the figure, rather than all of the data.

Removed from the figure, thank you.

Lines 514-516: Mention this finding is reasonable since you are working with two different lots of dye. The phrasing now makes it seem like you are recommending a change to the DB'17 value, rather than that you have a different 434Aimp value because you have a different lot of dye.

Thank you, we agree. We have rephrased the phrase as follows: "DB'17 proposed a value of $_{434}A_{imp}$ = 0.004413 absorbance units for a Sigma-Aldrich mCP dye at a final concentration of 3.3 µM in the sample cell ($_{488}A$ = 0.225; Takeshita et al., 2021). ~~However, b~~Based on our experimental results, we found that a better fit under our conditions corresponded to a $_{434}A_{imp}$ value of 0.0034 absorbance units, which is reasonable given that we are working with a different lot of mCP dye."

Line 543: No pH values were flagged as 4? If you're going to highlight the number of flag 3, you should include the number of other flags as well.

Thank you for noticing this. We have identified a mistake in the generation of the files in different formats for upload to the repository, in which some pH values were inadvertently removed. This occurred when one or some marine carbonate system parameters (pH, total alkalinity, dissolved inorganic nutrients, etc.) had a flag 4, which prevented our conversion routines from resolving the seawater CO_2 system and transforming pH values to in-situ conditions. Only complete pairs with pHT at 25ºC and in-situ were previously included, resulting in the removal of some data. We have fixed this issue to allow the conversion for all values, and re-uploaded the files in the repository. Moreover, from the new total of 23,843 samples, 23,773 were flagged as good (2), 19 as questionable (3) and 51 as bad (4). We have corrected this number in the manuscript, and this text as follows: " The dataset incorporates a quality flagging scheme consistent with GLODAPv2 recommendations (Key et al., 2015; Olsen et al., 2016), where flag 2 denotes good data (23,773 *samples), 3 (19 samples) and 4 (51 samples) indicate questionable and bad data, respectively, and 9 denotes not measured.* ~~*Specifically, 71 pH$_T$ values were flagged as questionable (flag = 3).*~~*"*

In addition, we have updated the DOI to the new corrected version in Sect. 6 as follows: " The complete OVIDE-BOCATS pH$_T$ dataset presented in this study is made available at ~~*https://doi.org/10.5281/zenodo.17141184*~~ *https://doi.org/10.5281/zenodo.17141183 (Pérez et al., 2025) in multiple formats to ensure broad accessibility and compatibility with different research workflows."*

Lines 545-547: So, were the pHT data from 2002-2018 not corrected using the DB'17 adjustment and L'11 equation? Or this was what was in GLODAPv2 before and will be updated using DB'17 and L'11 for v3? It needs to be very clear which data is in each version. Same for the 2021 and 2023 cruises – which procedure was used and what is the final product? Spell it out here for readers who skim, since this is a very important point.

We fully agree on the importance of clarifying this point and thank Reviewer #2 for highlighting it. We have revised the text as follows: "Spectrophotometric pH$_T$ data collected between 2002 and 2018, available in the GLODAPv2.2023 release (Lauvset et al., 2024), were computed using the CB'93 parameterization with the DVD'98 correction (+0.0047 pH$_T$ units)~~*and were already available in the GLODAPv2.2023 release (Lauvset et al., 2024)*~~*. The newly compiled pH$_T$ dataset presented here updates and corrects these data following the procedure presented in Sect. 3.4, applying the DB'17 adjustment and L'11 equation. In addition, the dataset significantly extends the temporal coverage by including pH$_T$ measurements from the 2021 and 2023 cruises, which were not previously available; these measurements are also corrected using the same procedure, resulting in a consistent final product across all cruises."*

Line 603: What's the uncertainty in these saturation depth changes? I would be curious what is the range in the depth that saturation = 1 based on the uncertainty in pH measurements themselves?

The uncertainty in the saturation depth change due to the uncertainty in pH measurements (given the reproducibility of the doubloons and the STD of the average pH$_T$ in the NEADW layer over 11 cruises) is 17 meters, and for Ω_{Ar} 0.0033 units. We have included in the manuscript as follows: "This reevaluation reveals a more pronounced reduction in aragonite saturation at the surface (from -0.040 to -0.065), relative to pre-industrial conditions, which progressively diminishes with depth, reaching changes of -0.016 near the seafloor. Although changes at depth appear small in absolute terms, the weak vertical gradient in aragonite saturation in deeper layers translates into a significant vertical shift in the saturation horizon—rising by approximately 120 m to 200 m. Based on the reproducibility of the doubloons (Sect. 2.2.5) and the standard deviation of the mean pH$_T$ in the NEADW layer over 11 cruises (Sect. 4.2), the uncertainty in the saturation-depth change is estimated at 17 meters, while the uncertainty in the aragonite saturation state is 0.0033 units. For instance, in the NA subpolar gyre, where the aragonite saturation horizon currently resides (...) "

Fig. 9: Difficult to read pH scale numbers and numbers in each panel, also difficult to read the x and y axes labels and numbers.

Corrected, thank you.

Lines 649-651: What are the uncertainties on these pH rate changes per year?

We have addressed this by adding a new figure to the Supporting Information (Supporting Information Fig. 7) that shows the standard deviation of pH_T across the 11 cruises (panel A) and the standard error of the pH_T rate (panel B). The text has been updated to guide the reader to this figure as follows: "These upper layers also exhibit high interannual pH_T variability (~~Figure 10~~ Supporting Information Fig. 7), which correlates negatively with AOU (Supporting Information Fig. 6b)."

Technical Corrections

Line 14: Provide the acronym meanings for GO-SHIP and OVIDE-BOCATS. *Done*

Line 97: Provide acronym meaning for CLIVAR. *Done*

Line 163: Should be CB'93. *Done*