We would like to thank Anonymous Referee #1 for the thoughtful review and the constructive feedback provided. We have carefully considered all comments, as detailed in the attached document.

We appreciate the helpful comments from the reviewer #1 that help to improve the readability of the manuscript. The reviewer comments are listed below with the italicized responses following. Changes to the manuscript and their associated line number are highlighted in blue text.

RC1: ['Comment on essd-2025-759'](), Anonymous Referee #1, 19 Jan 2026

This is a very useful study showing the aggregated data of the MedSea. The data is harmonized and quality controlled using state of the art methods, strongly increasing their availability and usefulness for future studies. The data set and product include data that were almost lost and now saved for the future.

Thank you for the compliment. We hope that our CARIMED data product will be useful to the oceanographic and modeling community working in the Mediterranean Sea, helping to improve current risk assessments and inform future measures for global change adaptation and mitigation.

<u>Some general comments:</u>

I would be interested to know how many data points (and their percentage) were discarded for the different variables. This would give an idea of the quality of the data entering the data product. This is important information enabling the assessment of the reliability of the data product.

> *The original aggregated data product includes all collected data along with the corresponding quality flags, including data of poor quality and/or outliers (flag 3), as described in Section 3.3. In contrast, the final bias-adjusted data product contains only data considered acceptable (flag 2) and interpolated data (flag 0). Interpolated data correspond to ancillary variables (such as dissolved oxygen or inorganic nutrients) that originally had flag 3 but are linked to CO_2 system variables (see Section 6.2). Variables from the CO_2 system—pH, TA, and DIC—with an original flag 3 are discarded from the final bias-corrected data product.*
>
> *Here is a quantification of the flags for the different biogeochemical variables in the two data products. This table can be found as Table S5 in the Supplementary Information and references to this new table can be found in the main manuscript. Data discarded would be those with flag 3 in the aggregated original cruise data product. In the CARIMED Data synthesis product, some of these inorganic nutrient and oxygen data are interpolated if possible and assigned a flag 0.*

Table S5. Number of data points with corresponding data flags in the two data products released within CARIMED. The data synthesis product with adjusted data contains no suspicious or wrong data (flag 3) while it contains adjusted data according to the secondary quality control (2QC) procedure.

	Aggregated original cruise data product					Bias adjusted data product			
	Total data	Flag 2	Flag 3	Flag 0	Flag 9	Flag 2	Flag 0	Flag 9	2QC
	With values	Originally measured and good	Originally measured and probably bad	Not measured but interpolated from similar variable	Not measured	Originally measured and good	Not measured but interpolated (similar variable or hermitian)	Not measured or eliminated	Adjusted data
CTDSAL	27590	27158	0	432	0	27158	432	0	2137
SALINITY	26176	746	0	25430	1414	--	--	--	--
CTDOXY	24133	15901	6	8226	3443	--	--	--	--
OXYGEN	24152	13267	25	10860	3438	13267	10882	3441	3880
NITRATE	18794	18628	166	0	8796	18628	1057	7905	4806
SILICATE	18351	18263	88	0	9239	18263	1042	8285	4783
PHOSPHATE	18548	18206	342	0	9042	18206	1118	8266	6972
DIC	4888	4762	126	0	22702	4762	0	22828	1323
TA	9187	9074	113	0	18403	9074	0	18516	1783
pH	6778	6726	50	0	20812	6726	0	20862	0

The use of NUT throughout the manuscript just for nutrients is not meaningful and strongly reduces the readability.

We have substituted the acronym with the full name in the whole manuscript.

Eastern and western only get capitals when they are part of a standing name. It is thus the eastern Med basin and western Med basin, or eastern Med Sea and western Med Sea. I identified some below, but please correct throughout the manuscript.

Thanks for the suggestion. We have corrected the corresponding capitals.

I suggest not to use RM as abbreviation, since it diminishes the readability and its use is not necessary to reduce the number of words or length of the manuscript.

Thanks for the suggestion. We have corrected the corresponding abbreviation.

Some minor comments and corrections in the following:

L61-63 Please provide some more info on these abbreviations. The reader is now just presented with a list of letters without knowing what they intend. *Thanks for the suggestion. Corrected.*

L66-67 GLODAP double definition. *Thanks for the suggestion. Corrected.*

L72 … efforts include … (add -s, delete comma). *Thanks for the suggestion. Corrected.*

L72 nutrients (add -s). *Thanks for the suggestion. Corrected.*

L74 and further … (add and). *Thanks for the suggestion. Corrected.*

L97 Why 'carbon-relevant variables' within parentheses?. *Thanks for the suggestion. Corrected.*

L99 parameters or variables?. *"Variables" thanks for the suggestion. Corrected.*

L111 and 118 This is the same information, one of these can be deleted. *Thanks for the suggestion. Deleted.*

L113 Figure 1: Please explain NCEI-OCADS and IEO-CSIC here. *Thanks for the suggestion. Explained.*

L194-195 “Nitrite and nitrate were summed to obtain total oxidised nitrogen (NO2- + NO3-), and listed as nitrate” I do not understand this. If nitrate occurs in the data set, why would you take nitrite, add it to nitrate and call the sum nitrate? This is not logical at all, because you change existing (and probably correct) nitrate values. If there is some logic behind this, please explain in the text. *Thanks for the suggestion. The new text now reads:* Nitrite and nitrate are measured together as total oxidised nitrogen ($NO_2^- + NO_3^-$), given the low nitrite concentration in open ocean and specially in the MedSea, total oxidised nitrogen is listed as nitrate.

L202 Why are the units not mentioned, as this is done for all other variables?. *There is a suit of transient tracers, each one with the corresponding standard units (Table 3). The text now reads as:* Transient tracers: Expressed in their standard units (see Table 3) (as retrieved, see Table S1).

L220 lower instead of reduced. *Thanks for the suggestion. Corrected.*

L224-225 Please provide some examples of such other issues. *The text now reads*: Other issues (wrong latitude and longitude coordinates, non-existent stations and/or depths, weird or badly reported units…) identified during 1QC were, wherever possible, discussed with the respective PIs.

L258 research vessels instead of RV’s. *Thanks for the suggestion. Corrected.*

L262 north, eastern. *Thanks for the suggestion. Corrected.*

L305, 306 helium, tritium (no capitals). *Thanks for the suggestion. Corrected.*

L318 eastern and western (no capitals). *Thanks for the suggestion. Corrected.*

L329-332 METEOR_51_2 was also included in GLODAPv2. Were the adjustments applied taken from that project, or were new adjustments developed? If the second case, how do these adjustments compare?. *Thanks for the suggestion. In section 4.3.3 we added this phrase:* Please note that CARIMED cruises also contained in GLODAPv2 (METEOR_51_2 and 84_3, see next section) were retrieved and treated using original uncorrected data. *In section 5.3,, in the Carbonate System paragraph we added this phrase:* The METEOR_51_2 TA was corrected both in CARIMED (+5 µmol kg-1) and GLODAPv2 (+9 µmol kg-1) while no corrections were applied for biogeochemical variables in the METEOR_84_3 cruise in both data products.

L363 characterises (with s as you are using British spelling). *Thanks for the suggestion. Corrected.*

L378-379 … GLODAP updates, integrating ship-based biogeochemical observations (Olsen et al., 2016; Lauvset et al., 2024), have established a benchmark for 2QC procedures. (insert commas after updates and after 2024 for better readability). *Thanks for the suggestion. Corrected.*

L398-399 “MedSea biogeochemical variables are distinct in terms of their cycles and drivers (Álvarez et al., 2023) and challenging in terms of analytical SOPs and RMs” How can BGC variables be challenging as to SOPs and RM? Such SOPs were especially produced for just such measurements. If you just want to state that measuring is difficult, keep it simple and rephrase this sentence. *Thanks for the suggestion. We have just added that they are analytically challenging.*

L400 In contrast instead of Conversely. *Thanks for the suggestion. Corrected.*

L405 western (no capital). *Thanks for the suggestion. Corrected.*

L406 use instead of propose (you have used it, right? Not just proposed it). *Thanks for the suggestion. Corrected.*

L422-425 I would be helpful to show a map with the locations of the basins and the basin boundaries. The readers do not have any info on that. Also geographical names of seas that occur in the manuscript should be shown there. Also the info in the first paragraph of §5.2 cannot be appreciated without a chart. *We agree that visual representation is helpful. A map of the sub-basins is provided in Figure S3, and we have clarified this reference in the main text.*

L436 The 250 km radius for cross-overs was taken from in the GLODAP effort, right? Because of the elevated variability in the Med Sea, the criteria for the vertical range were modified, which I agree with. However, I would also expect that the horizontal criteria for cross-overs be modified, because of the elevated variability which also occurs on horizontal spatial scales. Please justify why you did not do it. Please show with examples what the differences between cross-overs over 250 km and smaller radii could be. *The 250 km criterion was retained to ensure a sufficient number of crossover pairs, given the limited spatial and temporal overlap of cruises. Reducing the radius significantly decreases the number of valid crossovers, particularly in the eastern basin.*

L457 sub-basins (+s). *Thanks for the suggestion. Corrected.*

L472-474 "the CARIMED synthesis comprises historical and recent cruises collected using evolving methodologies, SOPs, and quality assurance procedures. Many early cruises were conducted before the availability of RMs for biogeochemical variables" This is obviously your argument for less stringent limits of adjustments. However, exactly the same holds for GLODAP with recent and historical cruises, so your argument is not valid. *Thanks for the suggestion. Corrected. The text now reads as:* To guide the required adjustments, minimum adjustments limits were prescribed for each variable. Consistent with 2QC procedures, these limits represent the minimum systematic bias that can be confidently identified relative to the measurement uncertainty, typically approximated by measurement precision. The CARIMED synthesis comprises historical and recent cruises collected using evolving methodologies, SOPs, and quality assurance procedures. Many early cruises were conducted before the availability of reference materials for biogeochemical variables, and on top the long-term temporal variability in the MedSea is evident. Therefore, the adjustment limits are the same order of magnitude as GLODAP but less restrictive (Table 2).

L513-515 "All corrections were rigorously inspected to ensure they did not remove true temporal trends or natural variability before being applied." Which trends did you detect and use? Please explain and possibly make a list of such trends, including references. *Long term trends in temperature and salinity are evident even in the low variability layers in the different sub-basins (Please have a look at Figures S9 y S12). We did not explicitly quantify trends. Shift changes are mainly detected in deep and bottom waters which are not used as reference layers. We added some references and further explanations in Section 5.1:…*. Ideally, final adjustments would be determined via the weighted least-squares inversion procedure proposed by Johnson et al. (2001), as utilised in GLODAP, which minimises offsets for a given cruise and variable across all crossover pairs. However, strict application of this procedure in CARIMED would inadvertently remove genuine temporal trends effecting the whole water column (Béthoux et al., 1998; Ozer et al., 2017; Schroeder et al., 2017; Kubin et al., 2023) and regime shifts (Roether et al., 1996; Rubino and Hainbucher, 2007; Schroeder et al., 2016) in the deep MedSea water masses.

L521-524 Please skip "units" after the numbers as practical salinity does not have a unit. *Thanks for the suggestion. Corrected.*

L537-538 "sometimes exceeding the adjustment limit by an order of magnitude" Is it still valid to use adjustments for DIC and TA of this magnitude? I do not think so. If the data is so far off the scale, there are probably also other problems with the measurements or with data processing. There are additional good arguments needed to justify the inclusion in the data product of such data. *The OTRANTO-5 DIC and TA data were presented in Krasokopoulo et al. (DSR 2011), their vertical distribution looks adequate but completely lacking accuracy, although they used reference materials. We have clarified it in the new text:* Carbonate system: For the seawater CO2 system variables, no adjustments were applied to pH. Despite the change in methodologies, transitioning from potentiometric to spectrophotometric measurements, the overall agreement for pH was comparatively better, which may be related to a better agreement of pH at the higher MedSea pH values (Álvarez et al., 2025). However, adjustments for TA and DIC were particularly high, sometimes exceeding the adjustment limit by an order of magnitude, even for cruises that reported using CO2 reference materials. A clear example is the OTRANTO_5 cruise where DIC and TA profiles have a coherent vertical distribution, a good precision, but bad accuracy. The METEOR_51_2 TA was corrected both in CARIMED (+5 µmol kg-1) and GLODAPv2 (+9 µmol kg-1) while no corrections were applied for biogeochemical variables in the METEOR_84_3 cruise in both data products. Preliminary findings from an inter-laboratory comparison exercise based on high DIC, TA, and pH mid-depth water from the Levantine Basin (Ibello et al., 2026) will also help to devise a strategy to improve the consistency of these climate-relevant biogeochemical variables.

L538-540 Findings of an intercomparison are mentioned. Were results of that study already included in the judgement on the CARIMED data? If not, this should be rephrased and made clearer. As it is now, this does not add any more confidence to the data. *The results from the interlaboratory comparison exercise are not published yet but were presented in the Ocean Science Meeting 2026 (https://x.com/oa_medhub/status/2026724021211935111). They have not been used to judge CARIMED data, but they will be used to design and implement adapted to the MedSea SOPs and reference materials. The text now reads as the underlined phrase in the paragraph above.*

Figures in the Supplementary Material are partly, for example S6 and S7, of lesser quality and hard to read completely. *Thanks for pointing this issue, they have been improved.*

Citation: https://doi.org/10.5194/essd-2025-759-RC1

We would like to thank Anonymous Referee #2 for the thoughtful review and the constructive feedback provided. We have carefully considered all comments, as detailed in the attached document.

We appreciate the helpful comments from the reviewer #2 that help to improve the readability and coherence of the manuscript. The reviewer comments are listed below with the italicized responses following. Changes to the manuscript and their associated line number are highlighted in blue text.

RC2: ['Comment on essd-2025-759'](), Anonymous Referee #2, 24 Jan 2026

Álvarez et al. present a highly valuable data compilation product focused on the Mediterranean Sea. To my knowledge, this is the first discrete bottle based, full water column data product for this region, representing a major advance in regional data synthesis efforts. The data product is of high quality, and the manuscript merits timely publication. Below are some suggested revisions that I hope the authors will consider:

> *We appreciate the thoughtful comments from Referee #2. We hope that our CARIMED data product will be useful for the oceanographic and modeling community working in the Mediterranean Sea, helping to improve current risk assessments and support future measures for global change adaptation and mitigation. Below, we provide our responses to the comments.*

Major comments:

1. While the authors' efforts to conduct the crossover analysis are commendable, I have concerns regarding the underlying basis for this approach. In Lines 421–426, the basins in which the 2QC was conducted are listed. Please provide figures, either in the main text or in the Supplementary Material, demonstrating that seasonal and interannual natural variability in these basins is sufficiently small over time periods of decades to justify such adjustments. Without this evidence, there is a risk that real natural variability could be inadvertently removed.

> *In Section 5.1, we explain that deep and bottom waters in the Mediterranean sub-basins cannot be used as a reference layer, unlike in GLODAP for other open-ocean deep basins. The temporal evolution and changes in these deep and bottom water masses in the Mediterranean Sea are clearly evidenced by transient tracers, as shown in Li and Tanhua (2020, https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2020.00594/full). The magnitude of these interannual changes exceeds that of seasonal variability.*
>
> *We identified the Tracer Minimum Zone (TMZ) layer in Figure S2, which presents the mean vertical profiles of CFC-11, CFC-12, and SF_6 across different sub-basins. In addition, Figures S15 and S16 show the temporal evolution of these tracers within the corresponding TMZ of each area. Clearly, even within the TMZ, a temporal evolution is observed in the Alboran, Algerian, Tyrrhenian, and Ionian sub-basins. Therefore, interannual variability, although small, is still significant. This justifies the use of both crossover analysis and the assessment of the temporal evolution of each property within the TMZ layer across different sub-basins or areas (Figures S9 to S14). Only clearly identifiable biases were corrected. As a result, the CARIMED data product is more likely to be slightly under-corrected rather than overcorrected. Seasonal variability is expected to be confined to the upper layers, well above 600 dbar, which corresponds to the minimum depth of the TMZ. Therefore, we consider seasonal variability within the TMZ to be minimal.*

2. The international research community has moved away from the outdated WHP format and has adopted a new standard (https://doi.org/10.3389/fmars.2021.705638). The WHP format was originally shaped by technical constraints of its time, resulting in non-intuitive abbreviations such as ALKALI and TCARBN. Notably, such abbreviations are not used consistently by the authors themselves in the manuscript, where terms such as DIC and TA are used instead. Another example is SAMPNO (Niskin Bottle #), which can be easily confused with Sample_ID, the unique identifier for a data row.

We are aware of the proposed updated vocabulary and column headers for WHP files stored in OCADS (Table 1 in the Frontiers paper), and we agree that they are more consistent and intuitive for certain properties. However, we followed the guidelines provided by A. Kozyr, who was our OCADS point of contact. Implementing these changes would require substantial effort, as it would involve transforming the original files as well as the data products that have already been corrected, formatted, and uploaded. Additionally, such modifications would introduce inconsistencies with the GLODAP-style files currently stored in OCADS. Therefore, we prefer to wait for a broader, coordinated decision among the National Data Centers, particularly NCEI-OCADS, regarding the adoption of standardized vocabulary and header abbreviations in archived files.

3. The provided NetCDF file does not seem to follow a specific template, with numerous empty strings as attributes, impeding readability. The CMIP community provides well-established examples of standardized NetCDF file structures, and I recommend following their templates. Please pay close attention to variable class types. For example, flags, station identifiers, cast numbers, and similar fields can be stored as integers (int32 or int64) rather than doubles to save disk space. In addition, please adopt a single convention for missing values (either NaN or −999, but not both); I recommend using −999. Finally, where possible, please maintain the same variable ordering as in the Excel file to improve consistency.

The NetCDF file has been completely regenerated. The file now uses the CMIP/GO-SHIP/CCHDO profile template structure, following the standards used by in CCHDO library, the conventions = "CF-1.10 ACDD-1.3". The file uses the incomplete multidimensional array representation (CF-1.10 9.3.2) with dimensions N_PROF=2066 (profiles) × N_LEVELS=35 (max bottle samples per profile), with proper coordinate axes and full ACDD-1.3 global attributes. Furthermore, following "CF-1.10 ACDD-1.3", every profile has now defined with featureType = "trajectoryProfile", and spatial axes were included: latitude(N_PROF) (axis=Y), longitude(N_PROF) (axis=X), pressure(N_PROF, N_LEVELS) (axis=Z, positive=down). This eases readability with programs like panoply.

4. When the same variable is measured using multiple methods (e.g., CTDSAL vs. Salinity, CTDOXY vs. Oxygen), both values should be reported. Although a merged column will serve the needs of most users, providing both measurements allows additional quality control by users who wish to perform independent assessments.

We agree with the referee. CARIMED outcomes (Figure 1) solved this issued as following:

- *Individual original cruise data and metadata: they contain the original sensor and discrete data for salinity and oxygen. Only oxygen data could be properly checked, as CTDOXY and OXYGEN were both measured in several cruises (see Section 3.3 and Figure S1). This is not the case for salinity it was mostly measured as either only CTDSAL or only discrete salinometer data (MILLERO_76, GEOSECS_Leg3, OTRANTO_5), just the MEDIPROD_IV, PROSOPE & CRELEV cruises had CTDSAL and SALINITY data, only 315 samples out of 27590 compiled samples.*
- *The Aggregated original cruise contains only CTDSAL data, discrete salinometer data were assigned to the CTDSAL variable with a flag 0, this happened in the MILLERO_76, GEOSECS_Leg3 and OTRANTO_5 cruises (432 samples vs a total of 27590, Table S5). This product contains both CTDOXY values calibrated with OXYGEN (wherever possible) and OXYGEN. Missing OXYGEN were substituted with calibrated CTDOXY and viceversa and assigned a flag 0 (see Section 6.1 and Table S5)*
- *The Bias-adjusted data product: contains only CTDSAL as the previous one, and only OXYGEN. Discrete Winkler data are less abundant but more reliable, missing OXYGEN data was substituted with calibrated CTDOXY data (10860 samples) and original OXYGEN data with flag 3 from 22 samples were interpolated with the Hermitian procedure. Our aim is to provide the best data in a friendly format following the GLODAP procedures, so only single variables for oxygen and salinity were provided in the bias adjusted product.*
- *Curious users can download the two products (samples are ordered equally) or even each original cruise file and check the consistency of our approach.*

More clarifications about this issue have been added in Section 6.1 and 6.2, and Table 3 in the main manuscript.

5. Checking the internal consistency between calculated and measured pH, as well as between calculated and measured carbonate ion content, is a powerful approach for quality control of ocean carbon system variables. If this has not already been done, please consider incorporating this step into the QC procedure.

We respectfully disagree with the referee. There is substantial evidence in the recent literature cautioning against using the mean difference between calculated pH (from DIC and TA) and measured spectrophotometric pH to adjust measured pH values. Detailed discussions can be found in the review papers by Álvarez et al. (2020, https://pubs.acs.org/doi/10.1021/acs.est.9b06932) and Carter et al. (2024, https://aslopubs.onlinelibrary.wiley.com/doi/10.1002/lno.12674). This issue is further complicated when combining spectrophotometric pH measured with purified and unpurified meta-cresol purple, as was done during the CARIMED cruises (see examples in Álvarez et al., 2025, https://aslopubs.onlinelibrary.wiley.com/doi/10.1002/lom3.70023).

Regarding carbonate ion measurements, Guallart et al. (2022, https://pubs.acs.org/doi/10.1021/acs.est.1c06083) clearly demonstrate discrepancies between measured and calculated carbonate ion concentrations (from pH and TA, DIC and TA, or pH and DIC), particularly for samples with high carbonate content, such as Mediterranean Sea waters. Several Mediterranean cruises were included in this review (TALPRO 2016, HOTMIX 2014, MSM72 2018, MEDWAVES 2016), all of which support the conclusion that internal consistency should not be used to identify biases in measured data.

6. pH was reported at 25oC. Consider adding an additional column showing the pH values adjusted to the in-situ conditions.

We have added this variable in Table 3 regarding the bias adjusted data product. Correspondingly, this new variable is described in Section 6.2.

7. SECT_ID values in the NetCDF file are empty, all entries are set to −999, and should instead be stored as strings.

The column has been renamed from SECT_ID to ALIAS throughout all data products (see also Minor comment 2 below). In the new NetCDF file, ALIAS is stored as a character-array (string) variable. Values are taken directly from the source Excel file (column ALIAS), stripped of trailing whitespace, and written without any numeric encoding. Representative values such as MILLERO_76, MEDIPROD_IV, or GEOSECS_LEG3 are now fully readable in the NetCDF file.

8. Time values are not stored correctly in either the Excel or NetCDF files, all values are zero, appearing as "0000" in Excel and 0 in the NetCDF file.

Times are now written as int32 in the NetCDF file and as four-digit zero-padded strings in the WHP-Exchange CSV. Of the 27,590 bottle records, 19,554 now carry a non-zero sampling time. The remaining 8,029 records correspond to older cruises for which sampling time was not originally recorded and retain a value of 0000 (or −999 where explicitly flagged as missing).

9. The date information is currently stored as numerical values in a non-standard format. I recommend splitting the date into three separate columns (Year, Month, and Day). Alternatively, if a single column is preferred, the date should follow the ISO-8601 format (YYYY-MM-DD), which is widely supported by standard software libraries.

The date information should be YYYYMMDD (text) as indicated by NCEI-OCADS rules. Please see Section 3.2 paragraph (i).

10. All the variables should be rounded to an appropriate number of decimal places. Otherwise, the current presentation may give a misleading impression of the measurement uncertainty.

All numerical variables are now rounded to a precision consistent with measurement uncertainty and standard oceanographic practice before any output is produced. The rounding scheme applied is:

Group of variables	*Decimal places*
LATITUDE, LONGITUDE	*4*
CTDPRS, DEPTH, MAXSAMPDEPTH, OXYGEN, AOU, ALKALI, TCARBN	*1*
NITRAT, NITRIT, SILCAT	*1–2*
PHSPHT	*3*
CTDTMP, CTDSAL, THETA, SIGMA0–4, GAMMA, HELIUM, NEON, TRITUM, pH	*4*
CFC-11, CFC-12, CFC-113, CCL4, SF6, partial pressures, DELHE3	*3*
Integer fields (flags, station/cast/date/time)	*0 (stored as int32)*

11. Consider creating two separate folders: one containing the unadjusted values and another containing the bias-adjusted values. That will facilitate data use, as many programs can capture the columns automatically.

We do not clearly understand the suggestion. The final CARIMED data product contains two columns for each biogeochemical variable, the original one (NAME) and the adjusted one (NAME_2QC) as showed in Table 3 of the main manuscript. We wanted, as the referee comments, to facilitate the comparison between original and adjusted data.

12. Consider adding a table to the supplementary information showing all the QC flag changes during the QC process.

We are not sure about this question. Maybe the referee is asking about the 1QC process, how many samples have a 1QC 2, 3 and 0. We guess it relates with the first comment of Referee #1, asking about how many data was discarded from the original cruise data.

Here is a quantification of the flags for the different biogeochemical variables in the two data products. This table can be found as Table S5 in the Supplementary Information and references to this new table can be found in the main manuscript. Data discarded would be those with flag 3 in the aggregated original cruise data product. In the CARIMED Data synthesis product, some of these inorganic nutrient and oxygen data are interpolated if possible and assigned a flag 0.

Table S5. Number of data points with corresponding data flags in the two data products released within CARIMED. The data synthesis product with adjusted data contains no suspicious or wrong data (flag 3) while it contains adjusted data according to the secondary quality control (2QC) procedure.

	Aggregated original cruise data product					Bias adjusted data product			
	Total data	Flag 2	Flag 3	Flag 0	Flag 9	Flag 2	Flag 0	Flag 9	2QC
	With values	Originally measured and good	Originally measured and probably bad	Not measured but interpolated from similar variable	Not measured	Originally measured and good	Not measured but interpolated (similar variable or hermitian)	Not measured or eliminated	Adjusted data
CTDSAL	27590	27158	0	432	0	27158	432	0	2137
SALINITY	26176	746	0	25430	1414	--	--	--	--
CTDOXY	24133	15901	6	8226	3443	--	--	--	--
OXYGEN	24152	13267	25	10860	3438	13267	10882	3441	3880
NITRATE	18794	18628	166	0	8796	18628	1057	7905	4806
SILICATE	18351	18263	88	0	9239	18263	1042	8285	4783

PHOSPHATE	18548	18206	342	0	9042	18206	1118	8266	6972
DIC	4888	4762	126	0	22702	4762	0	22828	1323
TA	9187	9074	113	0	18403	9074	0	18516	1783
pH	6778	6726	50	0	20812	6726	0	20862	0

13. Consider to adopt the concept of a Sampling ID, as it uniquely identifies each data row. This is particularly important when merging measurements from multiple sources into a single file, as done in this study. See the paper mentioned above for additional details about how to derive Sample IDs.

The referee is right that creating a unique Sampling ID such as the Bedford numbers (the PI of the Meteor 84_3 cruise in 2011 Toste Tanhua insisted in using them to identify each sample from the same Station_cast_Niskin) is useful. However, as commented in Section 3.2 different situations were encountered and we could not easily set a Sampling ID using Eq (1) in Jiang et al. (Frontiers, 2022).

Sample_ID=Station_ID×10000+Cast_number×100+Niskin

Therefore, we preferred not reporting the Sampling ID we used for each cruise as it was differently defined depending on the treated cruise.

14. Access to individual cruise data files could be improved by adding, in the abstract, a link to the CARIMED page hosted on NCEI's server: https://www.ncei.noaa.gov/access/ocean-carbon-acidification-data-system/oceans/CARIMED/. This would substantially facilitate access to the underlying cruise-level datasets. In addition, please consider including the dataset DOIs in Table 1. If the table becomes overly wide, some columns could be removed (e.g., Year, which is already contained in the date field).

The referee is right. We have added the link to NCEI-OCADS-CARIMED in the abstract and in the Data availability section. CARIMED will be also accessible through the GLODAP web page (no yet released). A link to the individual cruise data is already provided in Table S2 in the Supplementary Material.

15. Strictly speaking, the use of the term "concentration" is incorrect when the variables are reported using per mass based units. According to the IUPAC Gold Book, the term "content" should be used when referring to amounts expressed per unit mass of seawater, whereas "concentration" refers to the amount of solute per unit volume of solution.

The referee is right and we are happy to correct the term concentration. We have worked with metrologist and are used to use the term "content". We have changed concentration wherever it appear in the manuscript.

<u>Minor comments:</u>

1. Table 1. Align the ordering with the table served through NCEI. Whatever you prefer, but it will be nice to be consistent with each other. *With the permission of the referee we prefer to keep Table 1 with the cruises ordered with time, we think is a nice way of showing the evolution of the measured variables. While in Table S2 and S4 the cruises are presented as in OCADS.*

2. Page 8, Line 171: Cruise ID is not equivalent to Section ID. The latter is a leg or transect that is frequently visited for research purposes, e.g., P16. The former refers to a specific cruise visiting that leg, e.g., P16_2025. *The column is now labelled ALIAS in all three output files (WHP-Exchange CSV, Parquet, NetCDF).*

3. Page 9, line 190. Please specify the equations or software used for this unit conversion and provide the appropriate citations. When calculating density for the conversion, please also indicate which temperature was used.

For oxygen measurements, it is essential to use the temperature at which the Winkler samples were fixed, rather than the room or water-bath temperature that may be used for other variables (e.g., DIC). *The referee is right. But it is impossible to retrieve the "fixing" temperature for winkler samples therefore we used as described in Section 3.2 paragraph (ii): "* Dissolved oxygen (O_2): Expressed in µmol kg^{-1}; data originally in mL L^{-1} were converted using the factor 44.66 mL $µmol^{-1}$ and seawater density derived from salinity and potential temperature."

4. Page 22, Table 3. Hour (HHMM). Shouldn't "hour" be time here?. *The time-of-day column is named TIME in all output products. Its long_name attribute in the NetCDF file is set to "Time (HHMM, UTC)". No field named "hour" appears now in any of the outputs.*

5. In the Excel file, Temperature has no units, only scale. *The units attribute for both CTDTMP and THETA are now set to DegreeCelsius in the NetCDF file and in the units row of the WHP-Exchange CSV. This applies to both the in-situ temperature (CTDTMP) and the potential temperature (THETA). The scale descriptor ITS-90 is retained only as a note in the long_name of CTDTMP ("CTD temperature (in situ)"), while the units field correctly reflects the physical unit.*

Citation: https://doi.org/10.5194/essd-2025-759-RC2

We would like to thank Anonymous Referee #3 for the thoughtful review and the constructive feedback provided. We have carefully considered all comments, as detailed in the attached document.

We appreciate the helpful comments from the reviewer #3 that help to improve the readability and coherence of the manuscript. The reviewer comments are listed below with the italicized responses following. Changes to the manuscript and their associated line number are highlighted in blue text.

RC3: ['Comment on essd-2025-759'](), Anonymous Referee #3, 31 Jan 2026

The current manuscript describes CARIMED and the workflow and general approach behind it. CARIMED comes as two datasets that collect research cruise data of inorganic carbon and/or transient tracers in the Mediterranean Sea from 1976 to 2018, reformatted to a common, typically used format in the field (WHP) and made internally consistent. The procedures underlying this effort of data harmonization, quality control, and bias identification and adjustment (1QC, 2QC) are described in detail and follow state of the art techniques, e.g., those developed in GLODAPv2. Where necessary, they were expanded and adapted to the regional MedSea peculiarities, e.g., by focussing on specific water layers of low variability. The outcome is a benchmark and go-to dataset on MedSea carbon and ventilation dynamics, that will help re-define understanding of the interior Mediterranean carbon dynamics and beyond (e.g., as reference for adjustment of sensor-based observations through improved regressions, e.g., like CANYON-MED). The manuscript is well-written and it is easy to follow the line of thoughts and arguments.

Behind this data set description paper stands an enourmous work of tracing, finding, and recovering historic data as well as harmonizing data across 4 decades. The authors are to be commended for this work as well as for their comprehensive description (incl. historical background, e.g., section 4.3, which is extremely important to allow informed use of the data).

> *We thank Referee #3 for their positive comments and for recognizing our efforts. We hope that the CARIMED data product will be a valuable resource for the oceanographic and modelling communities working in the Mediterranean Sea, contributing to improved risk assessments and supporting future adaptation and mitigation strategies in the context of global change. Our detailed responses to the comments are provided below.*

<u>Questions and suggestions:</u>

This work focuses on carbon and transient tracer cruises and data, with nutrients as ancillary parameters. While it mentions regionally-focused data products for nutrients in the Western MedSea, it would be nice to get an impression of how many more cruises there would be (than the current 46), if the data collection would relax its carbon/transient tracer requirement and include all MedSea cruises with high quality nutrient data? I guess this can only be done as an order of magnitude estimate for the entire Mediterranean at this point. But it may be helpful, to illustrate the potential and to kindle motivation of the same/other researchers to next target such an expanded, stronger nutrient-focused FAIR MedSea dataset?

> *Several initiatives mentioned in the Introduction have developed biogeochemical data products for the Mediterranean Sea. These products not only compile and disseminate biogeochemical data in accordance with FAIR principles—such as dissolved inorganic nutrients (Belgacem et al., 2020, , https://essd.copernicus.org/articles/12/1985/2020/essd-12-1985-2020.html) and dissolved oxygen from CTD sensors (Belgacem et al., 2025, https://essd.copernicus.org/articles/17/5315/2025/essd-17-5315-2025.html)—but, importantly, they also make a substantial effort toward data harmonization and quality control. This constitutes a key added value of CARIMED, in line with other data products in the Mediterranean Sea. Additional examples focused on the CO_2 system are provided by Jiang et al. (2026, https://essd.copernicus.org/articles/18/1405/2026/).*

Regarding the question of how many additional cruises with biogeochemical data could contribute to such data products, we note the existence of the EMODnet Chemistry aggregated datasets for the Mediterranean Sea, particularly those addressing eutrophication. These datasets include 206,650 samples from 2,036 cruises spanning the period from August 1911 to July 2024 (HCMR/HNODC and OGS, 2025).

> *Hellenic Centre for Marine Research, Hellenic National Oceanographic Data Centre (HCMR/HNODC), National Institute of Oceanography and Applied Geophysics - OGS, Division of Oceanography. Mediterranean Sea - Eutrophication and Acidity aggregated datasets 1911/2024 v2025. https://doi.org/10.13120/13d2a552-ee18-49a0-abdf-ac7936bc070f*

A major challenge, however, lies in assessing the level of harmonization and overall consistency of these aggregated datasets. EMODnet provides detailed documentation of its data management workflow and quality control procedures, aimed at producing standardized, harmonized, and validated regional data collections for eutrophication (nutrients, chlorophyll, and oxygen) and ocean acidity (pH) across European sea basins (Lipizer et al., 2026). These procedures are primarily based on statistical evaluations of nutrient concentration ranges within the regions defined by Manca et al. (2004), with out-of-range values being flagged.

> *Lipizer M., Molina Jack M. E., Wesslander K., Fyrberg L., Tsompanou M., Iona A., Buga L., Sarbu G., Gatti J., Larsen M. M., Østrem A. K., Giorgetti A. 2026 Guidelines for eutrophication data aggregation, harmonisation and Quality Control. Version 3. Pp. 45, doi: 10.13120/8XM0-5M67*

Extending such approaches to ensure full consistency across datasets spanning more than a century (1911–2024) would represent a substantial effort, particularly if more advanced secondary quality control (2QC) procedures were to be applied.

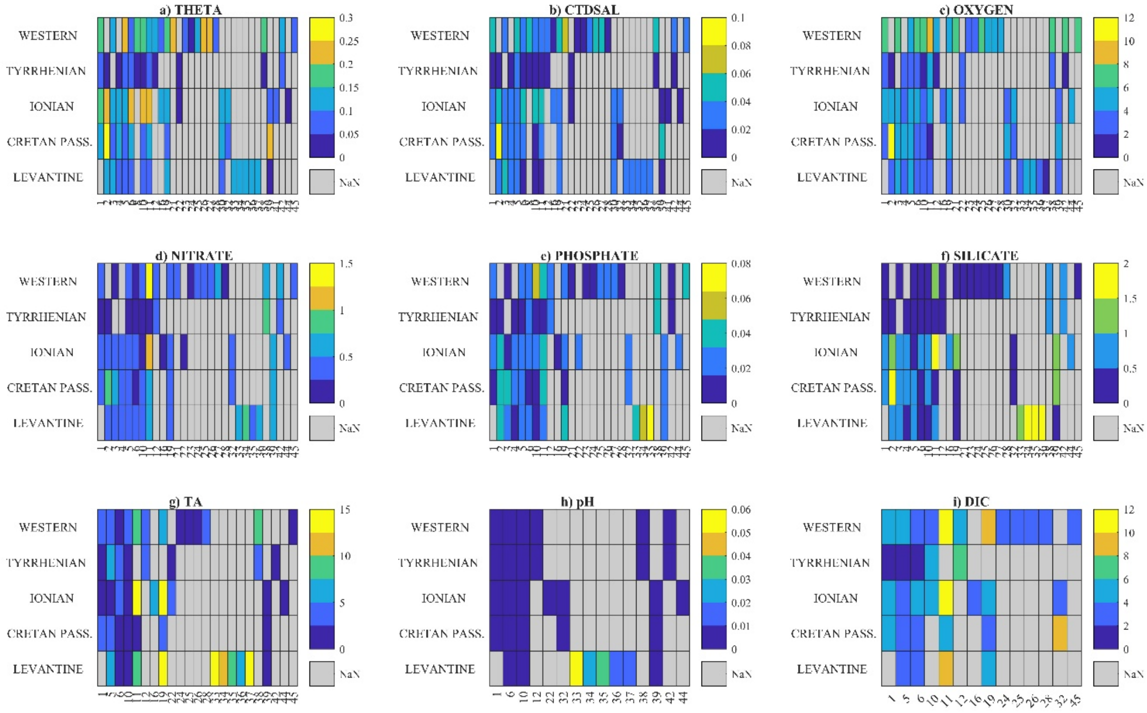
We have added a phrase about EMODnet Mediterranean Sea Eutrophication and acidity aggregated dataset in the introduction section. The text now reads as: Complementary regional efforts include the aggregated data regarding eutrophication (inorganic nutrients, dissolved oxygen, chlorophyll and acidity accomplished by EMODnet Chemistry (2025) with specific quality control procedures tailored for the Mediterranean Sea (Liziper et al., 2026) but no specific bias detection and correction procedures. In contrast the data products DIN-WMED for dissolved inorganic nutrients (Belgacem et al., 2019, 2021) and CTDO2-WMED for dissolved oxygen in the western Mediterranean Sea (Belgacem et al., 2025a, b), developed procedures to detect and finally apply adjustments to the aggregated data, and further highlight the value of basin-specific synthesis products tailored to regions with strong temporal and spatial variability.

I feel that a few key figures have been placed in the supplemental materials. While one can easily get overwhelmed by the number of figures that come out of such an analysis, I suggest the authors go through the supplemental material and check whether they (visually) support key aspects of the work and approach, and should rather move some of them to the main manuscript instead. Views may differ on which of them are "key", but, e.g., I would have thought a version of Fig. S8 on the intra-cruise homogeneity across depth/time as one such key figure, relevant to get readers of the main manuscript an idea of the reliability of the data products across depth/time.

> *We appreciate the referee's comment and have carefully reviewed the manuscript and the supplementary tables and figures. We prefer to maintain the current distribution of figures to preserve a simple and clear readability of the main manuscript. Introducing Fig. S8 into the main text, which shows the standard deviation of the original physical and biogeochemical data for each cruise in the tracer minimum zone (layer with the lowest temporal variability), would also require presenting the results after bias adjustment for the sake of coherence. Including both would make the main manuscript overly complex. For this reason, we have decided to keep Fig. S8 in the supplementary material, retain Table 2 in the main manuscript, and present Fig. 3 showing the overall CARIMED homogeneity before and after the bias adjustments.*

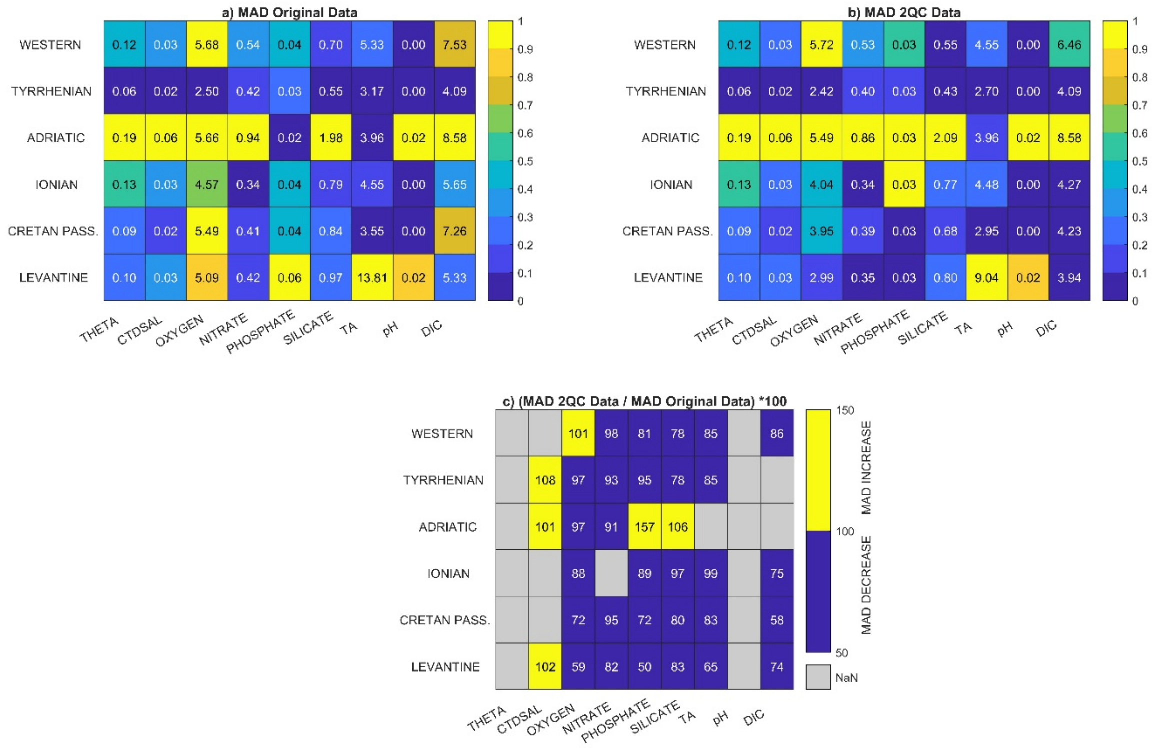
For the main manuscript, Fig. S8 should be re-worked by (1) using a colormap with a number of discrete rather than continuous color shading, which typically helps to better make out the data magnitude, and (2) using a less intense color for NaN (e.g., light grey instead of black) to focus the eye on the data rather than on the gaps.

Thanks for the suggestion. We hope the figure now looks clearer.



The same holds for Figure 3.

Thanks for the suggestion. We hope the figure now looks clearer.



Aspect to improve:

- Fig. 1 / early section 2: In most of the text you refer to two synthesis products whereas in figure 1 you talk about three complementary outputs of your work. Maybe this can be harmonized a little more (whether you created two or three data outcomes?).

Thanks for the suggestion. We have now changed Figure 1 and instead of products we use the word outcomes, indeed we release two data products (aggregated original and synthesis bias-adjusted products), and in addition we think the individual original cruise data and metadata in CARIMED NCEI-OCADS within the CARIMED Cruise Summary Table and Original Cruise data is a useful outcome.

Minor items:

- l.151: check sentence. *Thanks for the suggestion. The phrase now reads:* When available, additional information on analytical methods and measurement QC was collected, with comprehensive cruise reports proving particularly valuable information.

- l.268: check sentence. *Thanks for the suggestion. The phrase now reads:* More recently, the Med-SHIP program has promoted repetition every 5–6 years of North–South hydrographic lines in each MedSea sub-basin, with TALPRO and CRELEV cruises in 2016 (Schroeder et al., 2024) serving as reference occupations.

- l.425: check sentence. *Thanks for the suggestion. The phrase now reads:* For the western Mediterranean Basin, crossovers were performed for waters at depths between 600 m and 2,000 m.

- l.637-640: paragraph can be removed, because repeated at least the fourth time here. *Thanks for the suggestion. Indeed, it is repeated in the DATA AVAILABILITY section.*

- l.676: LINK. *Thanks for the suggestion. Added.*
https://www.ncei.noaa.gov/data/oceans/ncei/ocads/metadata/0309255.html

Citation: https://doi.org/10.5194/essd-2025-759-RC3