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.

Some general comments:

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