

Review of Metzl et al: Anthropogenic CO₂, air-sea CO₂ fluxes and acidification in the Southern Ocean.

The paper describes a comprehensive data set of in situ observations spanning multiple decades that is backed up by a gap-filled surface ocean pCO₂-product in the Indian sector of the Southern Ocean. It provides critical numbers for anthropogenic CO₂ uptake and accumulation, acidification rates etc. It also uses an older data set from the 1960s to put the data into perspective and uses the relationship between atmospheric CO₂ and anthropogenic carbon to extrapolate the results into the future based on atmospheric CO₂ of two emission scenarios. This is a very valuable scientific contribution and I fully support its publication. Thank you for collecting these data and making them available. It is particularly interesting to see the decadal variability in the carbonate system variables and how an increase in biological production can compensate the anthropogenic CO₂ uptake in recent years to lead to a relatively stable carbonate system. I have a couple of comments and clarifying questions that I would like the authors to address. The list is lengthy, but many are detailed comments on sentences being not clear.

General:

- 1) The paper is rather long and very detailed. It is certainly a strength of the paper to provide all numbers (convincing evidence) and this is appropriate for a publication in Ocean Sciences. I would nevertheless recommend the authors to see if the text can be somewhat shortened and stream-lined, for example by moving some of the numbers into a table.
 - a. In particular, I would ask the authors to shorten the summary section, which does not need to present all numbers again.
 - b. Also, I was a bit confused about the sections 3.3.1 “surface pH trend” and “3.3.2 Temporal changes in the water column”. The “surface” section 3.3.1 also comments on the depth profiles, which should be covered in the following section.
- 2) The figures are not of highest quality, probably plotted in Excel? Maybe consider whether there is a way to use a higher-quality plotting software for future work. I will list specific comments below, on font size, colors, missing legend entries where applicable. Font sizes are also generally too small in the supplementary figures.
- 3) Given the availability of a decadal time-series, could you comment on the discussion of the variability of the Southern Ocean carbon sink? E.g. the stagnation of the Southern Ocean carbon sink in the 1990s and/or the reinvigoration in 2015 (Le Quere et al., 2007, doi: 10.1126/science.1136188 Landschützer et al., 2015, doi:10.1126/science.aab2620). You touch on SAM, but I would appreciate a clear statement on decadal variability and its drivers in the summary/concluding section.

Specific comments:

- 1) There are quite a few language issues, I will mention some in the technical corrections, but the list is not exhaustive. Also, quite a few times, the sentences weren't super clear. See comments below. In general, please double-check whether it is always crystal-clear whether you talk about numbers/trends from in situ data or from FFNN.

- 2) Line 21: “At the surface during austral winter the oceanic fCO₂ increased at a rate close or slightly lower than in the atmosphere. “I’d appreciate if you added what that means for the ocean sink.
- 3) Line 28 “desperately”. I suggest changing to “is expected to increase” or such. How atmospheric CO₂ evolves depends on the decisions made by human societies.
- 4) Line 44 and elsewhere: why do you use parentheses around uncertainties? They belong to the number, so I would just report them in plain text (no parentheses)
- 5) Line 45: could update to Friedlingstein et al., 2023, 10.5194/essd-15-5301-2023
- 6) Line 54: “ongoing debate” add “on...” (explain the debate a little bit).
- 7) Lines 55-64: please mention that there is some doubt about the previously reported magnitude of decadal variability in the Southern Ocean carbon sink based on tests with the mapping products (Gloege et al., 2021, 10.1029/2020GB006788, Hauck et al., 2023, 10.1098/rsta.2022.0063)
- 8) Line 57 here and elsewhere you cite Hauck et al. 2023: “Sparse observations...”, doi: 10.1098/rsta.2022.0063 though you probably meant to refer to the RECCAP chapter (Hauck, J.; Gregor, L.; Nissen, C.; Patara, L.; Hague, M.; Mongwe, P.; Bushinsky, S.; Doney, S. C.; Gruber, N.; Le Quéré, C.; Manizza, M.; Mazloff, M.; Monteiro, P. M. S.; Terhaar, J. The Southern Ocean Carbon Cycle 1985–2018: Mean, Seasonal Cycle, Trends, and Storage. *Global Biogeochemical Cycles* **2023**, 37 (11), e2023GB007848. <https://doi.org/10.1029/2023GB007848>.)
- 9) Line 72: upwelling → enhanced upwelling
- 10) ~Line 74: I would personally add a note on the suggested secondary saturation horizon that is expected as the anthropogenic carbon uptake is strongest at the surface (Hauck et al., 2010; Negrete-Garcia et al., 2019). You seem not to confirm this with your results, so might be worth introducing this here in the first place.
- 11) Line 90: “modest source of CO₂”: Combined with the next sentence it reads as if this is a model flaw. A large sink certainly is, but a modest source is possible I would say. We have very few preindustrial observations ;) and so don’t know the number well.
- 12) Line 91-98: I would tone down the comment on the models a bit, because, yes there are systematic flaws, but also, often they fall into the observational range and/or we don’t really have the necessary process understanding that would be needed to cure those uncertainties/biases.
- 13) Line 175: can you comment on whether the Leseurre et al paper is based on exactly the same paper?
- 14) Line 190: maybe add Hauck & Völker, 2015, 10.1002/2015GL063070
- 15) Figure 1: please increase fontsize of all text on axes and within figure. Yellow text “Kerguelen” is a bit hard to read
- 16) Line 289: please give a quick summary of accuracy of 1990-1995 data
- 17) Please comment on accuracy of 1960s data
- 18) Line 298: CRMs were used for all data since 1998 and for none before that?
- 19) Line 322: can you please add the section number where this is shown.
- 20) Section 2.2.3: please comment on 1960s data as well.
- 21) Line 359: suggest to delete the sentence on pH here, as this is the topic of the next paragraph and alkalinity has not yet been introduced.
- 22) Line 419 “although”: add that there also wasn’t much winter data for training?
- 23) Line 420-423: I’m confused about this sentence. Why is it here and what is the message? This is the methods section, not results.

- 24) Table 1 and Figure 2 provides averages over certain time-periods, but these are not defined/explained until the end of section 3.1. Please explain this earlier (methods, even if referring to results). I am also puzzled over the chosen *five* seasons, some of them being single months. Why not stick to DJF, MAM, JJA, OND? This needs justification. Also for the time-periods considered, it would be useful to also report numbers for “standard decades” 1990-1999, 2000-2009 etc even if in supplement. The chose time periods also use the “edge” years twice, e.g. 2001-2010 and 2010-2020 both have 2010 included. Is that on purpose?
- 25) Line 430: “sampling locations were mainly reoccupied in austral summer”: unclear, where do the data for different seasons then come from? Do you simply mean: “...different seasons, though most of them stem from summer”?
- 26) Figure 2 and 4: please increase font size of all text (axes labels, legend). The inserted map is far too small. Colors of the symbols in the legend are not readable. Suggest changing the purple to grey or such, red symbols on top of purple line are hard to read. Still puzzled about the chosen seasons (that don’t seem to be interpreted much) and decadal periods.
- 27) Line 487: looks to me the positive SAM started rather in 2008 than in 2010. There were also positive phases earlier.
- 28) Line 506: “part of”: how much?
- 29) Figure 3: increase font sizes.
- 30) Figure 5: increase all font sizes. Change decimal comma “,” to “.” Orange dashed line missing in legend.
- 31) Line 590: “stability” this needs explanation before figure is shown, or delete in the caption.
- 32) Line 600: “albeit...” does this refer to October? Unclear.
- 33) Line 614-615: how does this increasing chl fit with no trend observed in nutrients? Everything used up immediately? Hauck et al., 2013 also found in a modelling study that increasing NPP can counteract the outgassing during a positive SAM phase. Would be nice to at some point in the discussion comment on the system response to SAM (summarize your findings).
- 34) Line 627: “higher”: higher than what? It is still well below atm CO2.
- 35) Line 637: this information is needed much earlier. I would also welcome a more detailed justification of those periods, it seems a bit arbitrary here.
- 36) Line 649: add depth range (100-300 m?)
- 37) Figure 6: Increase font sizes of axes and color bar labelsI know Cant detection methods can result in negative values, but it would be good to comment on them (artefacts or real?) What is the lowest value that you find?
- 38) Line 674: “top layers (0-400m, Figure 6b): Figure 6b only starts at 200 m...
- 39) Line 679-681: suggest ordering from surface to depth, so change order of these two sentences.
- 40) Line 693: “natural variability”. I’m curious whether more details on processes at work can be given?
- 41) Line 695: explain links between CT, O2, temp and Cant here.
- 42) Figure 7: moderate increase in font sizes
- 43) Line 729: I’m curious why none of these data was used by Gruber et al? Is it not in Glodap?

- 44) Line 744-746: "CT trend faster in summer <-> negative CT anomaly" doesn't make sense to me, more CT or less?
- 45) Line 749: "twice the rate ... that could be explained": how much can be explained and how was that estimated?
- 46) Line 753: "processes at the surface": which ones?
- 47) Line 756/57: I'm a bit lost which information comes from in situ obs and which from FFNN.
- 48) Line 759: "the temporal change" → absence of temporal change
- 49) Figure 8: subscripts for CO₂ in the legend would be nice.
- 50) Line 811ff: I am not sure how useful it is to list values from different sources which are all from different periods. This is acknowledged in the next sentence, but maybe this would a good place to shorten and refer to the numbers in the table without repeating them in the text.
- 51) Figure 9: see comments on Figure 2. I guess TS stands for total scale. Please spell out in the caption.
- 52) Line 863: the first number is from in situ obs? It would help me and maybe also other readers to specify.
- 53) Line 894: "no trend is observed for pH-PI": then why do you give a trend number in the figure?
- 54) Figure 11: increase font sizes. It would be easier to read if a box or background shading would be used to indicate the data for depth < 500 m (instead of the arrow). Otherwise, the arrow should be labelled directly. Also mark and label the data for depth > 500 m.
- 55) Line 935: "no change of Cant". Well, the figure shows between 0 and 30 μmol/kg Cant, this is not 0.
- 56) Line 939: "any appreciable": how defined?
- 57) Line 940: saturation state → saturation horizon!!
- 58) Line 947: saturation → saturation horizon
- 59) Line 960ff: the percentage numbers, do they really refer to pH or to H⁺. H⁺ would make more sense to me given that pH is a logarithmic scale.
- 60) Line 965: carbonate properties → carbonate system properties
- 61) Line 969: seasonality → seasonal amplitude
- 62) Line 972: I do not understand what this sentence is meant to say.
- 63) Line 977-981: I am confused about this part, I guess you mean to say that WW layer data can be compared to surface data, but please simplify the sentences.
- 64) Figure 12: increase font sizes. I found it very confusing to have the FFNN data with time on the y-axis plotted over the profiles of in situ data. After a while I understood, but it would be much appreciated if this could be simplified. An alternative could be to make a second panel for the FFNN data where color of dots is used to mark the years. This could be a smaller panel with only CT on the x-axis, no y-axis and would avoid the impression of a depth profile. Caption refers to "Jan 2020 obs", but only Jan 2021 obs are in legend.
- 65) Line 1039: does this calculation use salinity from 1962 or a climatology?
- 66) Figure 13: increase all font sizes.
- 67) Line 1071: 2014: I can't see SST in 2014 sticking out in Figure 13a. (2012?)
- 68) Line 1096: did you calculate the effect of the Delta T on pCO₂?

- 69) Line 1110: “extrapolation of trends”: do I understand correctly that this is using the trend/year and multiplying with number of years? i.e. independent of atm CO₂ assumptions? Please specify.
- 70) Line 1140/41: the aragonite saturation → the surface aragonite undersaturation
- 71) Line 1145: “correction”: this is not a correction, but a sensitivity test. Maybe simply say “warming”? Also note that this offline approximation of temperature will overestimate the temperature effect as it neglects circulation and mixing (warming will be limited by how much not-yet-warmed water is brought into the surface mixed layer)
- 72) Table 2: I am a bit lost here. First: why is SSP2-4.5 (the more realistic emission scenario) only shown for winter, and not also for summer (and the test cases)? Another idea might be to list the test cases in another table so that the design of these can be grasped quicker. Short for SSP2-4.5 is usually ssp245 and for SSP5-8.5 it is ssp585 (ssp85 → ssp585). I also overlooked the ssp45 in the table entirely for a while. Could you add a white space between the lines with the high and low emission scenario?
- 73) Figure 14 and 15: increase font sizes, use subscripts in legend, use complete name for scenarios.
- 74) Line 1260: the → anthropogenic
- 75) Line 1264: at minimum → at its minimum.
- 76) Line 1269: detectable increase: of what? Concentration or trend?
- 77) Line 1283: interesting, and it would be nice to comment on the complete effect of positive SAM index on the CO₂ system. Upwelling and outgassing limited to further south? At this location mostly nutrient effects?
- 78) Line 1290: any change of giving some numbers for source and sink from Rödenbeck et al?
- 79) Line 1303: “desperately” see comment 3 above.
- 80) Line 1313: “coupling of .. not well represented...” : I think this is a bit too general. The coupling between CT, AT, fCO₂ and pH is actually well represented with carbonate chemistry routines. It is more some of physical drivers and most of the biological drivers that cannot be represented with sufficient detail/process understanding.

Technical corrections (not complete):

- Line 19: In subsurface → In the subsurface
- Line 41: grammar, “taking up a large part ... since decades”
- Line 49: grammar
- Line 65: is → in
- Line 87: reach 0.7 → reach up to 0.7
- Line 89: ESM → ESMs
- Line 123: are in bracket → are given in brackets
- Line 309: be not → not be
- Line 321: latitude → latitudes
- Line 343: excepted → except
- Line 565: would occurred → would have occurred
- Line 742: “count” → contribute to XX
- Line 899: grammar

Line 990: the all → all the

Line 1066: corrected to → corrected for

Line 1158/59: grammar

Line 1312: BG → BGC