This is an interesting paper and an enjoyable read.

on page 21, lines 24 to 29, there is some incorrect information and understanding that has led you to some incorrect conclusions, and these are likely to have had a large impact on the results of the analysis.

The authors write:

Watson et al. (2020) estimated that the sum of these two effects would shift pCO2-based estimates of the mean global CO2 flux by 0.8 to 0.9PgC yr1 (stronger25 sink).

which is correct.

However the next sentence then says:

>So far, however, it is unclear how well the water temperature at the relevant vertical positions can actually be determined (an important source of uncertainty not included in Watson et al. (2020)'s range) and how it varies in space and time.

This is sentence is partially correct but also a bit misleading and I the authors may be confusing two different issues. The depth that satellite temperature data are relevant for is well understood and well studied (eg see the information from the international Global High Resolution SST (GHRSST) team and publications eg https://www.ghrsst.org/). What is less clear is how the satellite temperature data align with the top and bottom of the mass boundary layer which is where air-sea gas exchange occurs.

The authors then continue to say:

In any case, we note that our study mainly considers the variability of the flux, for which the effect of a time-constant correction as in Watson et al. (2020) would cancel out.

The authors are confused here as well. The Watson et al work presents two corrections. the first correction focuses on the issue that surface pCO2 data (and their paired temperature data) are all collected at different depths, as sampling depth varies between ships and even within a ship track (eg as the ship changes its ballasting). Whereas the
second correction that Watson discusses is the one that the authors can ignore as the authors are interested in variability rather than absolute CO2 sink value.

So first correction in Watson et al focusses on re-analysing the SOCAT pCO2 data to common and consistent depth. These methods are published and the data are published each year and the re-analysed version of SOCATv2020 are vailable (Shutler et al 2021) (equivalent data for SOCATv2020 and SOCATv2019 are listed on the SOCAT website). This aspect will be important the work the authors present, as I suspect that some of the variability that the authors characterise in their observation-based data is likely due to the inconsistent and varying depth over which the original SOCAT pCO2 data are collected. They authors can easily check for this by repeating their analysis using the re-analysed and depth consistent SOCAT dataset using the data from the Shutler et al link below. Using the re-analysed SOCAT data may actually strengthen the conclusions in the paper.

to help, the issue of how pCO2 data collected at depth is not always representative of the surface water has been recently identified for Arctic regions by Dong et al 2021. Dong et al show that these issues can result in biased fluxes due to salinity issues. Whereas the Watson work shows that this bias due to temperauter can be more widespread. the theory is well discussed in Woolf et al 2016.

Shutler et al (2021) Reanalysed (depth and temperature consistent) surface ocean CO2 atlas (SOCAT) version 2021, https://doi.pangaea.de/10.1594/PANGAEA.939233
