OS-2019-101, Li and Tanhua

This manuscript is a companion piece to a previous study by these two authors (along with additional co-authors) to identify potential new anthropogenic transient tracers for studies of ocean ventilation on decadal timescales. While the first manuscript (Li et al., 2019) focused on reconstructing the atmospheric histories and estimating the solubilities in water and seawater for several halogenated compounds, this manuscript is focused on the feasibility of measuring these compounds in seawater, demonstrating whether they behave conservatively, and finally rating their potential as tracers. In my view, the key to the publication of this manuscript is to have demonstrated the capability to measure the seawater concentration of these tracers. As such, the authors should have focused on the methods necessary to produce high-quality "Medusa" tracer data. In the data presented from the Mediterranean Sea, only 9 of the 21 samples are evaluated as good. Another 10 are given the flag of 5. The data from the Baltic Sea may be better, but the authors spend very little time on their interpretation. The remainder of the manuscript consists of hypothetical data interpretation (e.g. Fig. 5, Figs. S1 and S2) that is not worthy of publication on their own. I cannot recommend that this manuscript be published without major revisions that may include further method development. Developing the implementation of new tracers for studying the ocean is important, and I encourage the authors to understand the issues affecting their measurements. I would suggest some laboratory measurements using seawater samples equilibrated with atmospheric gases as a potential method for future studies.

In addition to the major scientific issue, the manuscript has many other issues with both writing and with some of the interpretation. I have listed many of them below:

p. 1, L21 – "potential be tracers"
p. 1, L31 - The source function for bomb tritium is not well constrained for much of the ocean. It was the ^{3}He-^{3}H tracer pair which are used as transient tracers. On the pedantic side, ^{39}Ar is not a "transient" tracer as it's assumed to be at steady-state.

p.2, L 3 – the industrial use of CFC-12 was phased out; assign a year to the Montreal Protocol
p.2, L9 – Although restrictions on SF_6 may be implemented, the concentrations in the atmosphere will continue to rise for the foreseeable future due to its long atmospheric life-time. Note that the PFCs also have high GHG potential, yet the authors conclude that they should be considered as transient tracers.
p. 2, L15 – Note that CFC-11 is a Level 1 (required) measurement for the US GO-SHIP program (https://usgoship.ucsd.edu/level-1-data/). The measurements of tracers are complicated – fortunately, the Bullister and Wisegarver (2008) paper referenced by the authors describes an analytical system capable of measuring seawater concentrations of CFC-11, CFC-12, SF_6, and CCl_4 precisely and accurately aboard a research vessel.

p. 2, L20 – There should never be a 1-sentence paragraph
p.2, L 24 – "Well-quantified sources and sinks" instead of "known input function"
p.2, L 29 - two phrases beginning with "as" makes the sentence confusing to read
P.2, L 35 - CFC-12 is difficult to use by itself as a tracer; with some caveats, it can be used as a tracer pair with e.g. SF_6. Based upon the reported blank level (0.48 fmol/kg), the utility of SF_6 extends back about 50 y instead of 100 y.

p. 3, L8 – see https://www.ncbi.nlm.nih.gov/pubmed/12608868 for a marine natural product with a C-F bond.
p. 3, L9 – grammar: "reasoning true for PFC-116"
p.3, L 19 – replace of with to; where is "here"?
p. 3, L 24 – when I went to this ftp site, the data were available through 2004 and the only HCFC data were for HCFC-22
p. 3, L27+ - If these studies of degradation in soils have no relevance for CFC-12 in the seawater, why should they have relevance for the Medusa tracers? This contributes nothing to this manuscript.
p. 3, L. 36 – "not enough information…"

p.4, L 13 – grammar: "…rendering transient tracers to penetrate…"
p. 4, L14 – TTDs assume a time-invariant circulation. In other words, they should not be expected to produce useful information (i.e. Sec. 5.5 is just a data exercise)
p. 4, L 25 – Lobert et al., 2015 only reported HCFC-22; Beyer et al., 2014 only report CFC-1301. How are these relevant specifically to HFC-134a and HFC-125?
p.4, L 27-28 – citation for the precision of MS vs. ECD?
p. 4, L 28-29 – The last sentence in this paragraph seems to be misplaced.
p. 4, L 30 – Shortened not shorted
p. 4, L 32 – An important component of the Medusa system is the trapping system – I wouldn't say the system is "based" upon it.

p. 5, L 28 – This is the first mention of the samples from the Baltic in the manuscript. It seems like a late addition.
p. 5, L 32 – I doubt that the glass ampoule was closed with a screw.

p. 6, L 10 – Is the difference in bubble size noticeable over the depth of the ampoule? I could understand placing the purge tube near the bottom of the ampoule to increase the physical stirring of the entire water sample. In addition, the bubbles are primarily responsible for stripping the compounds out of the sample – exchange across the gas-water interface would be extremely slow.
P. 6 and Fig. 2 – The manuscript and/or figure caption need more details. For example, I assume N1, N2 and N3 are the Nafion driers mentioned in the text. If the Methods section is the major component of the paper, elaborate.
p. 6, L13 – In this method, approximately 2 L of gas is used to purge 1.3 L of sample. In comparison, the Bullister and Wisegarver CFC/SF6 system purges a 200-cc sample with approximately a liter of UHP N2. It's not surprising that it takes multiple purges.

p. 6, L 33 – What is the "stripping efficiency" for the first purge? I.e. how much of the total tracer remains in solution?

p. 7, L3 – Clarify what is meant by "noises"
p. 7, L16 – "Larger" rather than "bigger"; a single purge
p. 7 L 18 – Are the blanks additive? Since each analysis of a single purge has an error associated with it, don't those errors add up (or at least the square root of the sum of the blanks squared)?

p. 8, L 6 – monotonically
p. 8, L8 - Why does the HFC-125 source function start in the mid-1990s? In Li et al. (2019) there is a reconstructed atmospheric concentration prior to that.
p. 8, L25 – What sets the maximum useful age? Blank level? Source function?
p. 8, L31 – I'm not sure what evaluated to be transient tracers means. I would probably say demonstrated to be useful (and that includes easily measured).

p. 9 first paragraph – If both figure referred to in a paragraph are in the supplemental material, then perhaps that is where the paragraph belongs as well.
Section 5.1 – The authors stress the importance of the WML for determining the long-term average saturation, then never return to this concept. I'm not sure why this is here, since they evaluate the Medusa tracers by their surface saturation (Section 5.3)
p. 9, L 13 – Examples of density profiles – rather than exemplary
p. 9, L 23 – Takes a long time or has a long time or has been isolated from the atmosphere for a long time?
p.9, L29 – "generally comparable" is meaningless. The authors should plot CFC-12(MS) vs. CFC-12(ECD), and let the reviewer decide if they are comparable.
p. 9, L31 – Is 5.9% an absolute difference? This is much larger than the precision of the measurements. How are data flagged as "good"? Note that only 9 of 21 samples are classified as "good".
Section 5.2 – This section is entitled "Observations of Medusa Tracers in Seawater" yet the Medusa tracers are never discussed. The final sentence points to Fig. 9.

p. 10, L 5 – In Section 5.1, the authors wrote that there is too much variability in the surface saturation to be useful.
p. 10 , L7 – Do not forget bubble injection and/or vertical mixing
p. 10, L10 – If the Medusa Aqua measurements of CFC-12 are 20% too low at the surface of the Mediterranean Sea, how can the reviewer have confidence in any of the other tracer measurements? Unless the authors can provide a reasonable explanation for this discrepancy for CFC-12, the remainder of the data are suspect in my opinion. Likewise, the surface CFC-12 saturations of 122% in the Baltic Sea also seem unreasonable. Since data from only one station in the Baltic Sea is reported, I assume that the uncertainties in the surface saturation are an indication of the precision of the Medusa tracer measurements. These are also not very reassuring for the most part (except for HCFC-142b at 2%).

The manuscript would greatly benefit from measurements of the Medusa tracers in the laboratory under controlled settings. I need to be convinced that Medusa Aqua can reliably measure the tracer concentration under laboratory conditions in order to have confidence in the reported measurements from the Mediterranean and Baltic Seas.

p. 10, L 25 – grammar: easy to soluble

p. 11, L 21 – Why not take the approach of finding a IG-TTD with a delta/gamma ratio that matches CFC-12 and SF6 together (instead of assuming the ratio is 1), and then applying this IG-TTD to the Medusa tracers?

p. 12, L 31 – The discussion of solubilities should have been in Li et al. (2019)