os-2019-101-EC1

Topic editor: The authors have responded to the comments of the reviewers and made a large number of changes to the manuscript. As a result, the paper is somewhat shorter and easier to read than previously and more of the details have been placed in the supplementary material. I believe that the paper now needs only a few relatively minor changes to be suitable for publication, but would like to give reviewers #2 and #3 a chance to comment on the revised version.

Dear editor and referees,

We thank the editor and two referees for your constructive suggestions and comments. Below, we address all the comments and describe our responses to them one by one. In the revised manuscript, all changes from the original text are marked.

To quantify the reproducibility of the analytical set-up and make our results more convincing, we took seawater from the tropical Atlantic Ocean, and let it equilibrate with the atmosphere in the laboratory. The water was then sampled from Niskin bottles in the same way as during a cruise and flame sealed in ampoules, although for this experiment we used 300 mL samples. The reproducibility for CFC-12, HCFC-22, HCFC-141b, and HCFC-142b measurements from four duplicate samples are 3.7%, 2.0%, 3.5%, and 3.4%, respectively. The added results validate the feasibility of the analytical method of ampoule sampling combined with the Aqua-Medusa system and make our results more credible.

With this in mind, we hope that the paper is now suitable for publication in Ocean Science.

Some comments I have are:

1. The abstract does not mention the samples from the Baltic; this seems odd given that these samples are used to justify some of the conclusions in the paper.

Response: The underlined part has been added to the sentence "*HCFC-22, HCFC-141b, HCFC-142b, HFC-134a, and HFC-125 have been measured in depth-profiles in the Mediterranean Sea for the first time, and reproducibility in the Baltic Sea*".

2. Nowhere in the paper do the authors state specifically that at present none of the suggested compounds can be considered to be suitable transient tracers. The last paragraph of the authors' overall response to reviewer #2 is a very clear statement of this (starting "The main result of this study…" and finishing "to guide further, more targeted experiments."). This could be added to the conclusion.

Response: Thanks for the suggestion. The mentioned paragraph has been added as the third paragraph in the Conclusion section.

3. In section 5.2 the authors discuss variability in their CFC-12 measurements. Reviewer #2 queried this, and the authors produced Table A1 in response. While these numbers can presumably be calculated from Table S6, a bit more explanation in the text would help understanding.

Response: This part has been moved up to Sect. 3.5. The underlined part has been added to the sentence "*For such samples, the averaged difference of CFC-12 concentrations measured by the two different instruments is 5.9 ± 4.6 % at roughly the same depth by ignoring their distance differences*.".

4. Like reviewer #3, I feel the paper seems to suggest liabilities of the present Medusa system for seagoing use as much as the problems of using these particular tracers. The authors have made some suggestions for future improvement; these could perhaps be expanded.

Response: The expanded part has been added in the third and fourth paragraphs in the Conclusion section, see the revised manuscript.

os-2019-101-referee-report-1

Presentation is greatly improved, however it is still difficult to determine what exactly the authors are trying to say.

This is a worthwhile study, and I think it should see the light of day (i.e., be published) after some work on presentation.

This draft is greatly improved over the prior. Unfortunately, the ms. is not currently ready. Although the text has been streamlined considerably (and I applaud the first author on the effort) the text is still confusing to follow.

The discussion of TTDs, for example, has been omitted from Section IV and moved to the supplement. Please, at least introduce, in the main manuscript, the idea (somewhere in section 4) of using D/G = 0.2 and D/G = 1.8 as a validity range (Waugh et al., 2003, or one of Hall's or Stoven's or Sonnerup's papers. Oh, and incidentally, this "valid range" is based on what is visible using the CFC / SF6 pair and does not reflect what may actually be happening in the ocean.) Without this introduction, the TTD-based curves in Fig. 10, and the discussion of TTDs on p. 9 (line 32) and p. 10 (line 8) have little explanation and appear to come out of nowhere.

Response: Many thanks for your nice reminder. We have added a paragraph (shortly described the TTD and mean age) to the Sect. 4.2 and changed the subtitle "*4.2 Tracer age*" to "*4.2 Tracer age and the Transit Time Distribution (TTD)*". For the "valid range", we do make a point in the manuscript that due to the time-variant ventilation, the validity range are not hard boundaries, and deviations can be conceived, see the first two sentences in the second paragraph in Sect. 5.4. In addition, we revised some sentences in the Sect. 5.4, see the revised manuscript.

The summary section (and Fig. 10) often disagrees with the final sentence of section 5.3, in which are listed sea surface saturations from the Baltic station:

CFC-12 122 +- 8 %

HCFC-22 77 +- 8 %

HCFC-141b 74 +- 12 %

HCFC-142b 114 +- 2 %

HFC – 134a 125 +- 23 %

HFC – 125 252 +- 35 %

One would ordinarily conclude, from this list, that HCFC-142b, HFC-134a have solubility functions that are approximated reasonably, and that they are stable in seawater.

Response: We agree with the reviewer's statement and have revised words in the "**HCFC-142b**" and "**HFC-134a**" paragraphs in Sect. 6, see the revised manuscript.

Jumping ahead to page 11:

For HCFC-141b one should note that the low saturation measured in this paper points to a problem with the solubility function, or degradation in seawater, or both.

Response: We have added a sentence in the "**HCFC-141b**" paragraph in Sect. 6. "*The low surface saturation measured in the Baltic Sea suggests that there is likely an issue with the solubility function*."

For HCFC-142b – are the Baltic data omitted from Fig. 10? I interpret the Baltic results as promising and wonder if some of the analytical issues identified in the Med. are the reason for the high (they are profoundly high, not 'slightly', as suggested) values in Fig. 10?

I would recommend that the Baltic measurements be included in the 'Synopsis' for each tracer.

Response: The results from cruise AL516 in the Baltic Sea have been added to Fig. 10 as black dots. We removed the word "slightly" from section 6.

I started the process of identifying individual sentences that are confusing or could use some clarification, but concluded that the list would be too exhaustive. This is the authors' responsibility.

In summary, while the ms. is greatly improved, I would appreciate it if the authors would take three steps: 1) put some more effort into clarifying the presentation, 2) have peers/ colleagues go over it once or twice, and 3) let it sit for a few days, then go over it again before resubmitting. I stand by my original review statement that this is a worthwhile effort that should be published eventually.

Response: We did the language check again following these suggestions and hope that it has been improved.

os-2019-101-referee-report-2

I have concerns with the accuracy of the measurements.

Review of Li & Tanhua (OS-2019-101, revised)

This version of the manuscript is much improved in focus. Many of my previous concerns have been addressed. However, my major issue is with the quality of the reported measurements from the Aqua-Medusa system. As a proof of concept that this instrument can detect the proposed tracers in the water column, I think the manuscript is fine. However, I have low confidence in the quality of the data - primarily based upon the inability to accurately measure CFC-12. The decision on which data are good (quality flag = 2) seems to be based on an undefined cutoff for consistency with the CFC-12 concentrations measured using the standard PT-ECD instrumentation. The assumption follows that the accuracy issue is due to the sampling or extraction, such that good CFC-12 concentrations are a necessary and sufficient condition for the other measured Medusa concentrations to also be good. I would prefer an explanation, supported by data, for the "bad" CFC-12 data. As Tanhua is aware, it is possible to measure one compound accurately but still have issues with other compounds using a PT-ECD system. Without confidence in the accuracy of the tracer data, I find the interpretation of the data to be interesting but not convincing. In my opinion, laboratory experiments conducted using seawater equilibrated with ambient air, sampled into glass ampoules and analyzed in the same manner as the samples from the Mediterranean and Baltic Sea, would greatly improve my confidence in their water column measurements. If I return to their list of five requirements for a transient tracer, I am not convinced of #4 and cannot therefore evaluate #5.

Response: To quantify the reproducibility of the analytical set-up and make our results more convincing, we took seawater from the tropical Atlantic Ocean, and let it equilibrate with the atmosphere in the laboratory. The water was then sampled from Niskin bottles in the same way as during a cruise and flame sealed in ampoules, although for this experiment we used 300 mL samples. The reproducibility for CFC-12, HCFC-22, HCFC-141b, and HCFC-142b measurements from four duplicate samples are 3.7%, 2.0%, 3.5%, and 3.4%, respectively. The added results validate the feasibility of the analytical method of ampoule sampling combined with the Aqua-Medusa system and make our results more credible. The samples show saturations very close to 100% for CFC-12, and over 100% for HCFC-22, HCFC-141b and HCFC-142b. For HFC-134a, we observed very high concentrations, undoubtedly from contamination by nearby cold-labs. Therefore, we do not use the saturations for this experiment to evaluate the tracers in this manuscript. But these measurements do increase our confidence in the analytical procedure, although it does not really explain why we had issues in the upper water column of the Mediterranean Sea.

We are currently setting up an experiment to reliable (we hope) formulate the solubility for those tracers. This will hopefully be conducted during 2021, but we are experiencing some difficulties to make a reliable standard as some of these compounds are hard to come by.

While we agree that it is indeed possible, and not uncommon, to be able to reliably measure one compound correctly and others not, we prefer to be on the safe side in this manuscript. That means that we used the criteria of "large" deviations from expected CFC-12 concentration to flag ALL data for those samples. Note that we still show samples flagged “5” in the figures, although with a different symbol.

Other specific comments – both editorial and scientific – are listed below.

p. 1, Line 29 - Ventilation is usually defined as a process, not a time (e.g. the process that transports water and climatically important trace gases from the surface mixed layer into the ocean interior).

Response: “*Ventilation is defined as the time elapsed since the water parcel has left the mixed layer and been transported to the ocean interior. Ocean ventilation and mixing processes play significant roles in climate as they are important processes to propagate perturbations on the ocean surface to the interior.”* has been changed to “*Ocean ventilation is defined as the process that transports perturbations such as the concentrations of tracer gases from the surface mixed layer into the ocean interior. Ocean ventilation and mixing processes play significant roles in climate.*”.

p. 2, line 12 – add the word “respectively” for clarity

Response: “*, respectively*” has been added.

p.2, line 16+– grammar (“have been” or “might be” imposed); will continue to rise?; “have readily measured”

Response: The target sentence “*Due to its very high global warming potential, some local restrictions on the production and use of SF_6 implemented. However, the concentrations of SF_6 in the atmosphere continue to rise due to its long atmospheric lifetime.*” has been changed to “*Due to its very high global warming potential, some local restrictions on the production and use of SF_6 has been imposed since 2006. However, the concentrations of SF_6 in the atmosphere is still increasing, partly due to its long atmospheric lifetime.*”

p. 2, line 19 – None of the US tracer groups have stopped measuring CFC-11. It is clear that the Tanhua group has stopped measuring it. I’d just leave out the discussion entirely. In addition, Lee & Bullister found evidence for the degradation of CFC-11 in permanently anoxic waters (e.g. the Black Sea), but not in the waters of the ODZs in the open ocean.

Response: The target sentences have been removed and the related sentences have been changed to “*CFC-12, SF_6, and CFC-11 can readily be measured onboard research vessels at a reasonable rate from one seawater sample. 1,1,2-trichloro-1,2,2-trifluoroethane (CFC-113) and carbon tetrachloride (CCl_4) have previously been used as oceanic transient tracers, but…*”.

p. 22, line 27 – citation for the need for multiple tracers?

Response: "*(Stöven and Tanhua, 2014; Holzer et al., 2018)*" has been added.

p. 3, line 8 – stable and stability are repetitive. Be clearer about what is meant by stable chemical structure. Long atmospheric lifetimes are indicative that a compound is probably stable to processes such as UV degradation of bonds even in the stratosphere.

Response: "*One indication of the stability is the stable chemical structure*" has been changed to "*One indication of the stability is the chemical structure and the atmospheric lifetime*".

p.3, line 20 – "with" not "to"

Response: "*associated to their estimates*" has been removed.

p 3, line 23+ - As the authors point out, a comparison of the surface saturations with CCl4 only suggest that the medusa tracers are perhaps more conservative than CCl4. That does not mean they are conservative. A better comparison would be to CFC-12 which is known to be conservative.

Response: We compared the average surface saturation of HCFCs and CFC-12. As seen in the newly added **Table S2 in the Supplement**, the average surface saturation of HCFCs tends to be higher than those of CFC-12. This may suggest that HCFCs are stable enough to be suited as tracers in the ocean. The corresponding paragraph has been revised: "*Another route is to compare surface saturations of a tracer with unknown stability to those of a compound that is known to be stable in seawater, e.g., CFC-12. The average surface saturation of HCFCs tends to be higher than those of CFC-12. This may suggest that HCFCs are stable enough to be suited as tracers in the ocean.*".

Table S2. Average surface saturation (%) of CFC-12, HCFC-22, and HCFC-142b from cruises BLAST III, GasEx98, PHASE1, and GOMECC

Cruise name	Sampling year	CFC-12	HCFC-22	HCFC-142b
BLAST III [a]	1996	-3.0 ± 10.4	21.6 ± 74.2	
GasEx98 [a]	1998	0.3 ± 5.8	7.6 ± 21.0	
PHASE1 [a]	2004	-22.8 ± 114.0	1.5 ± 6.7	
GOMECC [b]	2007	-0.4 ± 8.6	6.8 ± 108.7	6.0 ± 13.2

[a] National Oceanic and Atmospheric Administration (NOAA) cruises in 1992-2004 (ftp://ftp.cmdl.noaa.gov/hats/ocean/, last access: 20 January 2020); [b] Gulf of Mexico and East Coast Carbon Cruise (GOMECC) in 2007 (https://seabass.gsfc.nasa.gov/cruise/gomecc-1, last access: 10 June 2020).

p.4, line 26 – Clarify that Medusa was developed and utilized to measure volatile gases in the atmosphere.

Response: "*volatile trace gases has been …*" has been changed to "*volatile trace gases* *in the atmosphere* *has been …*".

p. 5, line 16 – "empirically" implies that improvements were made without any logical thought.

Response: "*empirically*" has been removed.

p. 5, line 29 – I suggest "measured at both adjacent stations located 15 nm away along the cruise track"

Response: "*although they were measured* *for the nearby stations, 15 nm (nautical miles) away from either direction*" has been changed to "*although they were measured* *at both adjacent stations located 15 nm away along the cruise track*".

p. 5, line 32 – Heating decreases the solubility - this results in the gases leaving solution into the headspace over a long period of time, but also increases the fractionation between the purging gas bubbles and the sample.

Response: Thanks. We changed the sentence to make this clear. "*Before measurement, each ampoule sample was immersed in a warm water bath at 65 °C* *overnight to support the transfer of the gases into the headspace and* *to enhance the purging efficiency*."

p. 6, line 32 – the precisions seem reasonable (Table 1). Are they based upon only one duplicate pair of ampoules? (A separate comment on Table 1 is that the number of significant figures seems rather high for the detection limits given the precision)

Response: Yes, the first "precision" column showed only one duplicate pair of ampoules from the Baltic Sea. But now we have added the second "precision" column and presented four duplicated pairs of ampoules from the tropical Atlantic Ocean. We have deleted one number of significant figures for detection limits.

p. 7, line 10 – sentence needs an "and" for the final phrase

Response: "*and*" has been added.

p. 7, line 13 – Interpretation of Transient Tracer Distributions?

Response: "*Transient tracer interpreting methods*" has been changed to "*Interpretation of transient tracer distributions*".

p. 7, line 16 – At the older end of the time range, the detection limit and precision are also important.

Response: Yes, that is obviously correct. We added the sentence "*For older waters, and low tracer concentrations, the precision and detection limits will be limiting factors (Tanhua et al., 2008; Stöven et al., 2015)".*

p. 7, line 21 – Others publications refer to the tracer age as the partial pressure age to distinguish it from other possible ages defined by the tracers.

Response: "*Tracer age*" has been changed to "*Tracer age* *(or partial pressure age)*" on the first line of Sect. 4.2.

p. 8, line 14 – add "and SF6 concentrations"

Response: We added this.

p. 8, line 19+ - I thought Atlantic Water entered the Mediterranean Sea at the surface. It is not clear to me why it should take a long time to equilibrate with the atmosphere. The typical reasons for undersaturation are entrainment of waters from below the mixed layer or cooling of the surface layer, with n=both processes occurring at rates faster than gas exchange can re-establish equilibrium.

Response: We do not really know why we see a constant undersaturation of 6% over time, while in the Atlantic we see time-dependent saturation. But we are offering a possible explanation that is now better explained: "*For CFC-12, this is different from the situation in the North Atlantic Ocean (Tanhua et al., 2008), and could be an indication of the different oceanographic settings where the inflowing Atlantic Water (to the Mediterranean Sea) take a long time to equilibrate with the atmosphere. The constant undersaturation through time is then mainly a function of rapid cooling during winter and entrainment of water from below, rather than a rapid change of the atmospheric concentration that can drive undersaturation that varies over time.*".

p. 8, line 27+ - I suggest the authors apply more rigor to the QC process. What is the definition of "inconsistent"? Greater than 3 x precision difference? As presented, it seems arbitrary. Are there correlations between the concentrations of CFC-12 and the Medusa tracers in the samples labeled as 5 that could help identify the issues? When I plot the Aqua Medusa tracers vs the Aqua Medusa CFC-12, there is significant overlap between good (flag=2) and bad (flag=5) data for HCFC141a and HCFC142b (i.e. the concentration ratios are consistent). There is less correlation for the other medusa tracers.

Response: We added a new section (Sect. 3.5) to illustrate the QC procedure. We calculated the three times the standard deviations (3σ) of CFC-12 observations for profile pairs 51-53, 83-85, and 105-107. The "inconsistent" means that the misfit of CFC-12 concentrations measured by the Medusa-Aqua system and PT-GC-ECD is more than the 3σ.

While we agree that it is indeed possible, and not uncommon, to be able to reliably measure one compound correctly and others not, we prefer to be on the safe side in this manuscript. That means that we used the criteria of "large" deviations from expected CFC-12 concentration to flag ALL data for those samples. Note that we still show samples flagged "5" in the figures, although with a different symbol. The inconsistency between ECD and Medusa measurements for CFC-12 does indicate an issue somewhere along the chain of operations, likely in sampling or storing that would affect tracers somewhat equal.

p. 8, line 34 – The data in the supplemental spreadsheet (Table S6) are reported in concentration units. These values are for the partial pressures. You should make this explicit in the manuscript.

Response: The last sentence in Sect. 3.4 has been changed to "*The observations of CFC-12, HCFC-22, HCFC-141b, HCFC-142b, HFC-134a and HFC-125 (in pmol kg^{-1}) measured by the Medusa-Aqua system in seawater from both cruises are shown in Table S6 with quality flags marked*." by adding "*(in pmol kg^{-1})*".

p.9, line 2 – consistent with your expectations of tracer concentrations in the deep waters.

Response: "*consistent with the well-ventilated Mediterranean deep waters*" has been changed to "*consistent with our expectations of tracer concentrations, and indicating the well-ventilated Mediterranean deep waters as interpreted by CFC-12 and SF_6 (Li and Tanhua, 2020)*".

p. 9, line 4 – This sentence needs to be re-written for clarity. Proximity to equilibrium with the atmosphere is not a factor – perhaps an indicator?

Response: Rewritten as "*The surface saturation of seawater can be an indicator of the stability of a compound in surface seawater or the correctness of the seawater solubility function*".

p. 9, line 12 – annual basis

Response: “*annual base*” has been changed to “*annual basis*”.

p. 9, line 13 - Why are the CFC-12 saturations so high in the Baltic Sea? Why do the saturations of the other gases vary so greatly? Without some explanation, I have little confidence in the data quality. Entrainment and surface warming would affect all of the tracers equally.

Response: We are not quite sure why we see these high values in the Baltic Sea for CFC-12. This is an active area, and we do agree that the saturation should be similar for all tracers, so that the saturation of HCFC-142b and HFC-134a looks “about right”. We realize that that is not a very satisfactory statement. As mentioned above, we have initiated trials to determine the solubility coefficient of these compounds.

The experiments in the laboratory in January 2021 (Table A1) indicate CFC-12 concentrations very close to saturation, as an indication of reliable measurements. We agree that the variability of the three samples is higher than hoped for.

Table A1. Concentrations of CFC-12, HCFC-22, HCFC-141b, HCFC-142b and HFC-134a (in pmol kg^{-1}) in seawater samples equilibrated with the atmosphere in the laboratory conducted in January 2021

Ampoule number	CFC-12	HCFC-22	HCFC-141b	HCFC-142b	HFC-134a
310	1.39	20.13	0.99	0.78	1163.50
105	1.50	20.90	1.04	0.84	1305.40
134	1.45	20.44	0.96	0.80	1205.50
372	1.51	21.02	1.03	0.84	1324.60
Average	1.46	20.62	1.00	0.81	1249.75
Precision (%)	3.7	2.0	3.5	3.4	6.2
Expected conc. [a]	1.38	9.70	0.77	0.35	2.01

[a] The expected concentrations of tracers are calculated based on their extrapolated atmospheric mole fractions in the Northern Hemisphere background sites in the year 2021.125 (only CFC-12 in the year 2021.5) and seawater solubility functions (Li et al., 2019), although the seawater samples were contaminated after equilibrated with the laboratory air.

P. 9-10 – If the 1-G TTD was representative of the processes controlling the distributions of the tracers in the Mediterranean Sea, all of the data plotted in the panels of Fig. 10 should fall along the same Δ/Γ line. Even the SF6 - CFC-12 data fall into the region of $\Delta/\Gamma>1.8$.

Response: We were aware that IG-TTD is not so suitable to constrain ventilation in the Mediterranean Sea (Stöven and Tanhua, 2014). There are two reasons for this as described in that paper: 1) It is most likely a mix of two water masses so that a "2IG-TTD" would be needed to describe the TTD, and 2) the time-variant ventilation of the Mediterranean Sea that invalidated the TTD concept. Indeed, we were not able to fit the SF_6 and CFC-12 measurements to the IG-TTD successfully from that cruise. However, as the first step, we would like to see the results based on different Δ/Γ ratios under IG-TTD for various tracer pairs. We do expect that the tracer pairs should fall within, or at least, close to the "validity area", just as the SF_6/CFC-12 pair.

We realize that the actively ventilated Mediterranean Sea is interesting from a ventilation perspective, see for instance Li and Tanhua (2020), and that we can expect to find these tracers through the water column. But that the time-variant ventilation and mixing of two or three major water masses with different ages and Δ/Γ makes the constraint difficult. So there are pros and cons of this area.

I agree with some of what the authors conclude about the feasibility of using these tracers in future oceanographic studies. The real issue for me is whether the ampoule sampling combined with the Aqua-Medusa system is capable of being utilized for these measurements. This is where I would focus my efforts in the future.

Response: To answer this question, we took seawater from the tropical Atlantic Ocean, and let it equilibrate with the atmosphere in the laboratory. The water was then sampled from Niskin bottles in the same way as during a cruise and flame sealed in ampoules, although for this experiment we used 300 mL samples. The reproducibility for CFC-12, HCFC-22, HCFC-141b, and HCFC-142b measurements from four duplicate samples are 3.7%, 2.0%, 3.5%, and 3.4%, respectively. The added results validate the feasibility of the analytical method of ampoule sampling combined with the Aqua-Medusa system and make our results more convincing.

References

Li, P., Mühle, J., Montzka, S. A., Oram, D. E., Miller, B. R., Weiss, R. F., Fraser, P. J., and Tanhua, T.: Atmospheric histories, growth rates and solubilities in seawater and other natural waters of the potential transient tracers HCFC-22, HCFC-141b, HCFC-142b, HFC-134a, HFC-125, HFC-23, PFC-14 and PFC-116, Ocean Sci., 15, 33–60, https://doi.org/10.5194/os-15-33-2019, 2019.

Li, P., and Tanhua, T.: Recent changes in deep ventilation of the Mediterranean Sea; evidence from long-term transient tracer observations, Frontiers in Marine Science, 7, 594, https://doi.org/10.3389/fmars.2020.00594, 2020.

Stöven, T., and Tanhua, T.: Ventilation of the Mediterranean Sea constrained by multiple transient tracer measurements, Ocean Sci., 10, 439–457, https://doi.org/10.5194/os-10-439-2014, 2014.