We thank both reviewers for their comments and helpful suggestions.

We are aware that the focus in this ESSD paper should be strongly on what was measured with less emphasis on derived parameters and interpretation. This is for this reason that the dataset reported does not include derived quantities, such as fCO_2 and pH. However, we feel strongly that for qualifying the data and to better position this dataset relative to other data, it is important to consider fCO_2 and pH, which are the most currently reported parameters for which trends and variability (both seasonal and interannual) are estimated in this region. This is the main reason for which we presented the seasonal cycle and trends of derived fCO_2 and pH, alongside with the other (measured) variables. We expose this now explicitly in the paper. We have also added information on the associated T and S variables, as suggested by reviewer 2, referring for that mostly to ESSD paper

https://doi.org/10.5194/essd-10-1-2018,

which describes monthly binned reconstructed T and S dataset along the same transect. On the other hand, we do not think (as pointed out by reviewer 1) that the focus of the paper should be on interpreting the variability in the data. Thus we refrain in this paper to explore further the observed variability in DIC, fCO_2, nutrients, or pH. This will be the focus of further papers.

Response to Anonymous Referee #1

Received and published: 11 July 2018

SURATLANT: a 1993–2017 surface sampling in the central part of the North Atlantic subpolar gyre

Gilles Reverdin et al.

Reviewer 1: General comments.

This manuscript describes a data collection gathered from a dynamic region in the North Atlantic Ocean over a period stretching from 1993 to 2017 and includes observations from all seasons. An impressive achievement which merits publication in ESSD. It should provide a clear overview on the data, its strengths and its weaknesses. A revision is needed for that. The aims and scope of ESSD state that "Any interpretation of data is outside the scope of regular articles ". This raises questions about including derived parameters, pH and fCO2 which are not in the data file, and associated discussion texts and figures. Other related matters are a) the binning of data for presentation and b) applying adjustments based on Alkalinity-Salinity rela-

tions, issues which this reviewer finds defendable for the presentation and discussion of the data. These aspects require editorial decisions which might substantially alter and shorten the presentation. Consequently a diminished effort is given in this review on sections dealing with interpretations and derived parameters. The research effort described has expanded in scope and with time, thus developed with time into a series as is described in section 2. The research has progressed without monitoring guidelines such as are available today. The results are nevertheless valuable.

Authors: We thank reviewer 1 for his comments and helpful suggestions.
We have above answered on the first part of the overall comment on the contents of the paper. Reviewer 1 then lists two other related matters
a: the binning of data for presentation.
b: applying adjustments based on Alkalinity-Salinity relations.

(a) The binning of data is a fairly straightforward way to present data distributed geographically. In this sense, we either include most of the domain as one bin for a large number of the plots, or have tried to separate the area in a limited number of smaller domains. The choice of the bin size is a compromise between having a minimum sampling so that the average seasonal cycle or trends can be estimated with some confidence, and not including too large hydro-climatic differences across the bin. This resulted in selecting separate bins on the shelves, close to the shelves, in the interior of the gyre, and for the near-Reykjanes ridge region. As these definitions clearly could vary somewhat in time, this is a minimalist approach, and there can be rather large variations in salinity and water masses within a bin.

(b) To minimize their effect when producing bin-averages, we have tried to establish how the different variables relate to salinity. This, for example, is what we present with a standard Alkalinity-salinity relationship adapted to this specific region. We have then adopted the approach recommended by Friis et al. (2003). We will now also provide the bin-averaged seasonal cycles under the seanoe doi.

Reviewer 1: The authors have clearly spent a considerable effort on flagging suspicious data using a variety of arguments. In some cases these seem speculative, e.g. for particulate phosphate contribution dissolved phosphate on page 16 line 424. What could be the source of particulate phosphate in January, mid winter? My point is that speculations on grounds

for flagging or eliminating data are less important than describing the criteria for including the data. In discussing data problems and adjustments to sections of the data, the text should also report on how flags are used in the data file. The methods and uncertainties are dealt with in appendices. There, this reviewer finds sections on temperature and salinity missing although they are mentioned in the main text on page 6. The important parameter salinity is probably the one with the best method consistency through the time of observations. Furthermore, this reviewer sees no good reason to deal with some parameters in appendices and others in the main text. It should not be a problem to incorporate all into the main text. The text on the different parameters could be shortened and condensed. It is hardly necessary to include details on numbered sample bottles as is done on page 17.

The authors:

We agree that the presentation of T and S should also have been done in the appendix at the same time as the other variables, and moved it there (from the main text).

We agree that S is the parameter for which the sampling and method has been most consistent in time.

We have tried to condense and shorten the presentation of the different parameters.

Notice also on former page 16, line 142, that the particulate phosphate was measured on those samples, and thus corresponded to a real measurement. It obviously does not originate from plankton (as the reviewer also comments) but most likely from the storage of the water in a dirty tank or pipes upstream of the faucet, unbeknownst to us. This is consistent with the erroneous salinities observed during part of this transect, suggesting that the water was not originating directing from the sea as we thought.

We are now also concentrating the presentation on how we add flags, and on how we validate the data. It however seems important to retain the most important issues we encountered. The one with some of the bottles for DIC/At falls probably in that category. Although we agree that this might be incidental and as such could be omitted, the issue of contamination with bottles is not something that we were aware of from other publications. The erroneous data for those bottles lasted for multiple reuses of the bottles despite prolonged rinsing between successive uses (every three months) and after changing caps, indicating that those bottles either became chemically reactive or harbored microbial strains resistant to the mercury chloride solution. This was disconcerting, and, as it raised issue on the overall quality of our analyses, we decided to get rid of the old bottles. In this case, it is thus important to point

what flags we set on these data (with the old bottles). We kept the remaining DIC/At data as 'good', despite the suspicion that this raises.

Reviewer 1:

Specific comments

Reviewer 1: Page Line 5 102 Please explain the role and training of "ship riders".

Authors: the ship rider is very instrumental as he/she is alone on board with the crew, and that remote communication is not very easy. The ship rider is issued a list of instructions and notes, and meet with the PI and the different laboratory assistants before boarding the ship. The ship rider is in charge at the end of providing written notes and log sheets and of storing the water bottles on board for later recovery in a port or to carry water samples back to the PI. The sentence has been modified, as: 'The sampling used here is largely based on surface samples collected every three months by a volunteer instructed at LOCEAN (or at LDEO before 1995).'

Reviewer 1: 6 109 Delete regularly.

Authors: Done

Reviewer 1: 6 112 "by one of the authors". Please explain.

Authors: Removed. GR was in charge of following the data collected on this line. (l. 109, there was an error in the number of probes: 1200 instead of 12000)

Reviewer 1: 6 115 T measured directly from a bucket. Is there an estimate of the measurement precision?

Authors: The direct comparison of T in a bucket to an intake temperature (associated with the TSG) was only done during one crossing in April 1994. Based on this set of 74 bucket measurements, the difference in T averaged -0.03°C except for three instances with high wind (up to force 8) and large air-sea temperature differences (larger than -6°C), when the bucket T was too low by close to -0.3°C. This suggests that for most of the bucket T measurements, the bucket T accuracy is better than 0.1°C, as it was operated mostly in weak to moderate winds and not large air-sea temperature differences. We are looking forward to carry further comparisons with the current (smaller) bucket when an intake T-measurement will again be available. In the past, much larger differences have been reported on bucket measurements, but this refers to routine marine meteorological measurements.

Reviewer 1: 7 141 "Units are standard ones" is meaningless without explanation.

Authors: thank-you. This was replaced by 'all data are expressed in ‰ relative to the reference V-PDB (Vienna–PeeDee Belemnite) (Craig, 1957).'

Reviewer 1: 7 149 Reference is needed on the GISS database.

Authors: we added

Schmidt, G. A., Bigg, G. R., and Rohling, E. J.: Global Seawater Oxygen-18 Database, available at: http://data.giss.nasa.gov/o18data/, NASA Goddard Inst of Space Sci., New York, 1999

The database is currently being upgraded to a V2 version. However, upon recently checking, we found that the data we submitted two years ago have not yet been incorporated, but were told that this would be done soon.

Reviewer 1: 7 155 "The error in doing it has little impact on the computation". Add: for this region.

Authors: added

Reviewer 1: 8 175 Reference is missing for Cooperative

Global Atmospheric Data Integration Project (2017).

Authors: added

Dlugokencky, E.J., P.M. Lang, J.W. Mund, A.M. Crotwell, M.J. Crotwell, and K.W. Thoning (2017), Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2016, Version: 2017-07-28, Path:ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/.

Reviewer 1: 9 205 1-2-1 smoothing needs explanation.

*Authors:*This is a running mean average with coefficients ¼, ½, ¼ for successive months

Reviewer 1: 10 259 Could there be other non-siliceous organisms than calcifying ones?

Authors: sure, even in the late spring-summer period. We changed the sentence there to be more general ('non-siliceous organisms').

Reviewer 1: 10 263 Reference is needed for denitrification on Arctic shelves.

Authors: we added a reference to McTigue et al (2016)

McTigue, N. D. *et al*. Biotic and abiotic controls on co-occurring nitrogen cycling processes in shallow Arctic shelf sediments. *Nat. Commun.* **7**, 13145 doi: 10.1038/ncomms13145, 2016.

Reviewer 1: 16 436 How was the Certified Reference Material used?

Authors: The CRM are used at regular times (at least twice a day) during analysis runs that last often two to three days, to standardize the hydrochloric acid and to provide enough statistical information on the uncertainty for a particular analysis run. This uncertainty was reported to us, and has varied in time, but we don't think that it should be indicated in the paper/dataset, as it is probably not the main source of error. The different CRM batches used at LOCEAN since 2004 are numbers 64,66,71,73,77,83,88,97,99,104,116,119,122,128,131,139,155,165.

We have added:

'Calibrated Certified Reference Material (CRM) provided by A. Dickson (Scripps Institution of Oceanography, San Diego, USA) are regularly analyzed to standardize the hydrochloric acid and to provide an analytical uncertainty on the data.'

Reviewer 1: 17 443-445 Here a correction for mercuric chloride dilution is reportedly not applied for some samples but other samples are usually with the correction applied. Is it possible to be consistent and apply the correction to all results?

Authors: We had not applied a correction, as it was sometimes uncertain if the ship rider had used the specified amount of poison, or the filled volume in the DIC/A_t bottle sometimes differed from the recommended volume. However, we agree that this should be done, assuming that the solution has no DIC and no At.

Upon your request, we implemented the proposed correction in the data file (and have added the flag code 'dubious' for data in February 2005, for which the volume of saturated mercuric chloride solution was highly uncertain). Because of this and a few other added flags on A_t data, we have estimated an updated A_t-S relationship, excluding the data flagged dubious (this is the one now reported in the paper).

Reviewer 1: 17 468 Here the correction for the mercuric chloride dilution could and should be applied to simplify the comparison.

Authors: this is now done, and the table has been updated.

Reviewer 1: 18 477 "recent international inter-comparison",

reference or further information lacking.

Authors: we added

'A recent international inter-comparison on two shared water batches (spring 2017) suggests that the LOCEAN analysis presents a small negative bias both for A_t and DIC (Bockmon E. and A. Dickson, 2018, Personal Communication), but not in a very similar range of values to the ones observed during SURATLANT.

Reviewer 1: 20 550 Explain what IRMS is.

Authors: Now explained. IRMS means 'isotope ratio mass spectrometer'

Reviewer 1: 21 586 "The Nondal et al (2009) relation underestimates At for low S, but

...

". Note that Nondal et al. describe two relationships, one for S>34.5 and another for Polar Water with S<34.5. Calculating At for S<34.5 is very imprecise.

Authors: we agree. We use the Nondal et al relation for S > 34.5. The Polar water relation of Nondal et al for S<34.5 does not fit well the relation we find in the waters on the Labrador shelf influenced by the Canadian Arctic, which is not surprising as this relation was based on data mostly along north-east Greenland.

Reviewer 1: 35 1104 It appears unlikely that all listed laboratories use a potentiometric method for the determination of DIC but none the more common coulometric DIC determination.

Authors: Correct (table also updated, with additional comparisons in 2015-2016 that had been omitted in previous draft).

Reviewer 1: 47 1483 Comparing Figures 2 and 11 on data distribution reveals data points in fig. 11 which seem outside the subpolar (SURATLANT) region, north of 65N and as far north as 68N, in the Iceland Sea.

These data are in the data file and might deviate from the subpolar gyre relationships.

Authors: this is correct. There are 17 points collected north of 65°N and seen in Fig. 11 (including two which are way to the north or north-east of Iceland). They do not appear on Figure 2, as data were not including 2017 to establish the relationship discussed At-S, and

they are not included in the figures discussed in figures 3 to 10 because of the spatial domain used. It is possible that the relationship in Figure 11 might deviate from the one in the subpolar gyre, in particular for the point at 68°N. On the other hand, the northernmost points are only in summer, and bear very little weight on the trend illustrated on the right panel of figure 11.

Response to Reviewer 2

Reviewer 2:

The paper presents a dataset of temperature, salinity, carbon dioxide variables, nutrients and water stable isotopes in surface waters between Iceland and Newfoundland since 1993. The procedure of validation, accuracy and characteristics of the data are discussed together with the resulting seasonal and interannual cycles.

Even when there are many adjustments in the data set, justified in the text, one data experimentally determined is of great interest and should be recorded and archived. One of my major concerns in this paper is why the authors do not show temperature data (unless a Hövmüller diagram), seasonal and interannual trends. Temperature controls most of the variables considered and discussed and it is complex to explain the results without that figure. Moreover, they compute variables of the carbonate system (pH and $f\mathrm{CO}_2$) where this variable is crucial and, when they also compare their estimation with VOS line data presenting both a great (even too high) agreement. Moreover, due to the range in salinity along the region and the seasonal variability, carbonate variables (Alkalinity and total dissolved inorganic carbon) should be presented normalized to a constant salinity. They provided a relation between AT and S that should be used as indicated in Friis et al., 2003, included in their manuscript. Therefore, present your discussion indicating the seasonal trends for the normalized values of DIC and At. I consider a moderate/major revision should be done in the manuscript before acceptance.

Authors: Binned time series of T and S along a close-by transect (AX02) are presented in another ESSD paper
https://doi.org/10.5194/essd-10-1-2018

The interannual deviations of SST is presented as a Hövmüller diagram in Fig. C1 (right panel) along the transect. One clearly sees the period of large positive anomalies peaking near 2005, as well as interannual deviations. We also added one panel for winter T' in the comparison on Fig. C2, which hints to the quality of the analysis, and maybe its caveats tat times when SST data accuracy was insufficient. We did not present the average seasonal cycle of SST on Fig.3, as we thought that this was well known by lots of other means, and because the depiction we have with this data set is not that different from other ones. We agree that SST and the mixing processes associated with its seasonal changes play a major role in controlling the seasonal variability of many of the variables presented. We have thus added in the manuscript a reference to the seasonal range of SST in the different bins/latitude ranges (ranging from -0.4 (Feb) to 11.9 (August) over the Newfoundland shelf (yellow box of fig. 3) to 5.7 (Feb) to 12.1 (Aug) just north of it (red box), and 7.0 (Feb) to 11.0 (Aug) in box 4 (blue box)).

For the seasonal cycle of DIC and A_t (Fig. 3) we have worked with deviations from relations to salinity, as discussed in the text (in the spirit of Friis et al, 2013). We also considered those for estimating long-time trends. However, as the objective of the paper is not in diagnosing this interesting variability, but in validating and qualifying the dataset, we preferred to present the trends estimated without separating the 'water mass' (or 'hydrological') component, so that they can be directly compared with other published results.

Finally, we are not sure of what is meant in the sentence: 'when they also compare their estimation with VOS line data presenting both a great (even too high) agreement…'

If this refers to the comparison with fCO2 (measured and estimated), we could just say that we were reassured by the comparison and that the values of the standard deviations in this set of comparisons fitted with our expectation (and are better than the worst-case scenario for them that we report in the manuscript).

Other comments of **reviewer 2:**

Reviewer 2: Line 65. 2001-2008 is an important number of years for some parameters. However, for those with an important seasonal variability and with a high error of estimation such as the pH in this case, consider including some indication about the short period of time in the referred paper.

Authors: the scatter was shown for some parameters. Indeed, because of that, interannual variability… 2001-2008 is a short period. We think that the way our sentence was formulated clearly hints to the duration issue 'As an extreme case, based on the winter 2001-2008 data…', We thus do not see what we should change in this sentence to make it clearer. The choice of winter season in the earlier paper was related to the lesser effect of seasonal variability (thus less scatter due to sampling in different months) when making that estimate.

Reviewer 2: Line 97. Here and along the text, the transect is defined as AX02 and in others AX2 (also in Figure 1). Homogenize

Authors: thank-you. We prefer to use AX02 (in line with the other ESSD paper). This was thus homogenized in this paper too.

Reviewer 2: Line 126. I do not consider this method of sampling as one to be used for calibrating any data to, at least, the required accuracy.

Authors: we fully agree.

We encountered difficulties to associate temperatures that would be good estimates of surface temperatures to the data. Either, when SST data are derived from a TSG with no (or unreliable) intake T measurements, as happened part of the time, and for which we tried to correct 'a warming' based on other data (typically, from XBTs dropped from the ship). We also used temperature measured from buckets at time. In 1993-1994, we had a very large canvas bucket which seemed to be rather clean and did not modify significantly the water temperature and salinity samples. For instance, a comparison in April 1994 of 74 bucket temperatures with intake temperatures and XBT temperatures, shows that except in three cases of high wind (force 8) and large air-sea temperature differences (larger than 8°C), the comparisons were indistinguishable from 'noise', with a possible small negative bias of 0.03°C in bucket temperatures. In 2016-2018, a smaller bucket is used. We do not have reliable T-comparisons for this bucket, but the comparisons for S suggests a small excess salinity (on the order of 0.01), which might be indicative of a small negative bias in temperature due to evaporation. We consider (except during big storms), these T to have an accuracy close to 0.1°C. Then, we added in some 'rare' cases with no available other data, T(5m) from nearby Argo floats, when S compared well. These are accurate, but of course there is the error due to interpolation and non-similarities (probably at the 0.5°C level). Finally there are the cases of co-localisation of S/SMI SST estimates. We used the best

available data, and tried to avoid vicinity of fronts or rain areas, but of course, there are possible biases, which we tried to take into account. These data should be considered to have errors at the 1.0°C level.

Reviewer 2: Line 143. Even when they tried to explain it, please remove psu in any figure and in salinity values.

Authors: We are aware that the use of *psu* is not recommended. The reason we would have preferred to indicate it on figure captions is that there starts to be more papers showing absolute salinity instead of practical salinity (the symbol *psu* clearly indicates that we use practical salinity). We have removed elsewhere *psu* in the text and figures.

Reviewer 2: Line 218. Does it mean you select the data giving you a better figure? If you have done an important data treatment to remove, improve and homogenize the experimental data, why did you use different data sets? Line 238. Present the data for At and DIC normalized to a constant salinity following Friis et al., 200

Authors: The word 'errors' was misused. What was meant was larger 'scatter', thus resulting in a large uncertainty in the 'trends' estimates. There is a range of T and S in the domain, thus a larger scatter when using all data without 'normalization'. This can be minimized indeed by normalizing to a constant salinity following Friis et al., This is what we have done on figure 3 for the seasonal cycle. This was also tested for long-time trends, but is not reported here, as the trend results could not be compared to other published papers, which was the aim here (validation of the data set). We have not removed large subsets of experimental data, except when appropriately flagged as for the most serious cases explicitly mentioned in the appendix.

Reviewer 2: Line 243. The formation of organic matter reduced dissolved nutrient concentrations but also DIC (not increase)

Authors: we agree, but the comment on line 243 was not for DIC, but for A_t. It is usually expected that total alkalinity increases during large phytoplankton spring blooms in the subpolar North Atlantic, because of the uptake of nitrate, phosphate, and because it usually involves siliceous organisms. We have added A_t in the sentence to make it clearer.

Reviewer 2: Line 244. Do you see a decrease in alkalinity following this same argument? There is a small decrease in AT in the figure, but without normalization and with the error in the estimation, it is difficult to assure this conclusion.

Authors: this is what we are pointing out. It is indeed quite small and at the limit of the uncertainty range, but seems to stand out from error estimates in most boxes. There is apparently some misunderstanding on the contents of the figure. This is done after normalization (that is removing for A_t the linear fit to S, as in Friis et al (2003); a constant is then added to present it as if it was at S=35), as explained in the figure caption, and in section 2.2.1

Reviewer 2: Line 252. What about deep convective mixing in the area with important interannual variability?

Authors: the interannual variability of mixing is certainly a strong contributor to interannual variability. Notice however that most of the transects are outside of the western Irminger and other areas of deep winter mixing, and are thus not directly influenced by this process. They are influenced on the other hand rather directly by interannual changes in the Reykjanes mode water formation.

Reviewer 2: Line 267. Clarify what Snorth means.

Authors: we do not find Snorth in our version.

Reviewer 2: Line 275. It is not clear what SST values were used to compute these values.

Authors: In each box, an averaged seasonal cycle is estimated for SST, as well as for SSS, DIC, A_t, nutrients, from which a seasonal cycle of fCO_2 and pH is estimated. This is what is used in figures 4 and 5. We compared those seasonal cycles with estimating for each individual sample fCO_2 and pH and then averaging them in each box to produce an average seasonal cycle with no significant difference. Explaining and better understanding the seasonal cycles, the decadal trends and interannual variability will be the focus of a future paper.

Reviewer 2: Line 281. Again, in this region SST plays a strong influence in fCO_2 seasonal variability. Therefore, it will be important to show SST to see how much of this variability is related to SST and how much is due to biological effects. You could also present the data following Takahashi et al., papers.

Authors: there is certainly a strong dependence on SST variability. On the other hand, from what we understand on ESSD publications (and from the comments of Reviewer 1), we are not expected to dwell much on scientific results, but on what contributes to

validating/qualifying the data and to show the prospect of using these data for scientific investigations.

Reviewer 2: Line 340. Higher trends are found in Bates et al, 2014 paper for Irminger region and Olafsson et al. 2010. Please compare. Could this be in part due to your lower DIC trends?

Authors: Thank-you. We have modified the discussion starting on line 339:

'Similarly to fCO2 trend, this pH trend for the NASPG is close to the mean global ocean estimate of -0.0018 yr^{-1} (Lauvset et al., 2015) and comparable to other trends evaluated in the North Atlantic polar waters, ranging between -0.0017 yr^{-1} and -0.0026 yr^{-1} depending on the periods, seasons and regions (Bates et al., 2014; Lauvset and Gruber, 2014; Lauvset et al., 2015; Olafsson et al., 2009). Compared to the Irminger Sea, the difference of pH trends is mainly explained by observed DIC trends (0.7 µmol/kg/yr for NASPG against 1.6 µmol/kg/yr for Irminger Sea, Bates et al 2014).'

Reviewer 2: Line 344. How do you calculate the significance? You say in line 346 large uncertainties, small number of years ….

Authors: There are two main sources of uncertainties: the one due to the sampling (scatter due to small time and space variability, and unresolved seasonal variability), and the one associated with the method (for example, the average seasonal cycle is estimated with the same data as the trends, and there are systematic shifts in time of the seasonal sampling from year to year that could result in mixing up the seasonal cycle and trends), as well as with the systematic corrections of the data (as here for $\delta^{13}C_{DIC}$). The significance refers to the first source of uncertainty (and is reported now for each estimate of a trend), whereas the uncertainties on line 346 are of the second kind, which by nature we cannot quantify.

We summarize below all the trends (slightly revised, as we did not take into account data now listed with QC flagged 3 or 4) and their uncertainty (one standard deviation)

Figure 6 : At2 = f(t) : -0.036 (+/- 0.058) (not significative)
Figure 7 : DIC winter : +0.787 (+/- 0.081) ; DIC summer : +0.765 (+/- 0.133)
Figure 8 : fCO2 winter: + 1.757 (+/- 0.123) ; fCO2 summer: +2.060 (+/- 0.155)
Figure 9 : fCO2 all : +1.946 (+/- 0.116) ; delta fCO2 all : +0.002 (+/- 0.116)
Figure 10: pH all : -0.00206 (+/- 0.0001)

Figure 11 : C13 all : -0.0189 (+/- 0.0036) ; C13 winter: -0.0136 (+/- 0.0031); C13 summer: -0.0420 (+/- 0.0032)

Reviewer 2: Line 355. Please include reference for Alert station.

Authors: the added reference is White et al (2015)

White, J.W.C., B.H. Vaughn, and S.E. Michel (2015), University of Colorado, Institute of Arctic and Alpine Research (INSTAAR), Stable Isotopic Composition of Atmospheric Carbon Dioxide (13C and 18O) from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1990-2014, Version: 2015-10-26, Path: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2c13/flask/

Reviewer 2: Line 385 and the full paragraph. I do not see this as a conclusion. Please, move up

Authors: we agree that this is not so much a conclusion as a perspective. However, we think that this perspective has its place here, and reformulated it as:

'Part of the scatter we find in the discrete sample data set results however from insufficient sampling of the seasonal variability. To provide a more complete analysis, it will be important to combine with other data, either from the same ships of opportunity (operated mostly by NOAA/AOML) or from other platforms. These include near surface temperature and salinity operated near-continuously from TSGs, and also at times pCO_2 measured with equilibrator systems. Notice also that the information on instantaneous mixed layer depth and stratification information was provided by near-simultaneous XBT profiles. The investigation should also include the compilations of stations data in GLODAP (Olsen et al., 2016) or the very rich SOCAT (Bakker et al., 2016) data base. There is also a large array of complementary observations, such as from the Argo and the bio-Argo profiling platforms (Organelli et al., 2017). For example, these data suggest blooms in mid-winter that could have a depletion impact on net production and export of nutrient and carbon from the surface layer already in March (Lacour et al., 2017). Notice however that the bio-Argo floats have mostly sampled the rim of the subpolar gyre and provide only indirect evidence on near-surface carbon and nutrient that they did not measure. Thus combining the different in situ cruise data sets with the Argo data will provide other challenges.'

Reviewer 2: Line 459. You used CRMs for your analysis. Why should your values be adjusted by this important value? What about alkalinity values?

Authors: we do not understand the reference to line 459. Maybe this refers to line 436.

If this is the case, notice that we have changed this sentence, as: 'Calibrated Certified Reference Material (CRM) provided by A. Dickson (Scripps Institution of Oceanography, San Diego, USA) are regularly analyzed to standardize the hydrochloric acid and to provide an analytical uncertainty.'

Reviewer 2: The analysis

Line 480-498. The interannual trend in DIC of 0.7 in a period of 24 year means an increase of 16.8 μmol/kg in the full period, i.e., less than twice the indicated errors in the DIC values. Moreover, with an error of 10 units in DIC, computed pH values could be affected with an error as high as 0.02-0.03 units while in fCO2 could be close to 30 μatm. The resulting very low average difference between computed and VOS data and high error (line 493, -3.6 +- 12.4) indicates positive and negative deviations in DIC and/or different sign in the deviation of At and DIC. Conclude from this intercalibration exercise that used only experimental values what it is indicated in line 497 looks too much.

Authors: We agree that the camparison with the VOS data suggests a much smaller average error on DIC and At than the 10 μmol kg^{-1} that we mention as an upper limit. Indeed, this is more likely to be usually a factor of 2 or three smaller. On the other hand, the comparison with the VOS fCO2 data might not have captured rare but possible cases when error would have been larger, such as we identified in some transects that were flagged from the dataset. We thus added the sentence:

'This also suggests that the random error in DIC and A_t is much less than the worst-case scenario mentioned above of 10 μmol kg^{-1}.'

Authors: with respect to the interannual trend of DIC and fCO_2, we decided to change the text when discussing fCO_2 trend (former lines 339-343), as:

'Similarly to fCO2 trend, this pH trend for the NASG is close to the mean global ocean estimate of -0.0018 yr-1 (Lauvset et al., 2015) and comparable to other trends evaluated in the North Atlantic polar waters, ranging between -0.0017 yr-1 and -0.0026 yr-1 depending on the periods, seasons and regions (Bates et al., 2014; Lauvset and Gruber, 2014; Lauvset et al., 2015; Olafsson et al., 2009). Compared to the Irminger Sea, the difference of pH trends is mainly explained by observed DIC trends (0.7 µmol/kg/yr for NASPG against 1.6 µmol/kg/yr for Irminger Sea, Bates et al 2014).'

Reviewer 2: Line 508. The LOCEAN DIC values were always lower than other values, as indicated above. Have you corrected the data for this bias?

Authors: We have not corrected the data for this probable bias. However, we have now corrected all recent data for the dilution with the mercuric chloride solution, assuming fixed volume of 450 ml in all bottles, that the solution contains no DIC nor contributes to alkalinity, and that the recommended solution volume was added to the bottles.