

### Review #3

M. Vázquez-Rodríguez, F. F. Pérez, A. Velo, A. F. Ríos, and H. Mercier (2012), Observed trends of anthropogenic acidification in North Atlantic water masses. *Biogeosciences Discuss.* (bg-2012-83) 9:3003-3030

#### General Comments

The authors have assembled a vast dataset of carbonate system constraints measured over the last three decades. They have verified that the first set of quality control checks have been performed upon this data and have taken some efforts to see that the assembled data are intercomparable. They have then used this dataset to determine the evolution of relationships between various parameters and the measured or calculated pH. Finally, they have applied the determined relationships to an annually averaged property distribution to estimate the temporal evolution of the pH distribution from the temporal evolution of the regression of pH against other properties. The authors' main contribution in this paper is a set of rates of acidification for a number of isopycnal layers in the North Atlantic.

This paper is making a comment on the acidification of deeper and intermediate waters that is important and likely correct, and it does so concisely. However, the paper has several large flaws in its current state: the description of the dataset quality control efforts is incomplete, the description of and the justification for the analysis used is lacking, and the conclusions are essentially a restatement of other work published by the same researchers given the assumptions implicit in the methods used. Given that the same dataset is used for the earlier published paper, this paper can also not be viewed as a confirmation of the earlier findings.

#### Scientific Comments

Many of my concerns have already been brought up by Anonymous Reviewer #1, with whom I broadly agree. One exception is that I am not worried by potential errors created by the use of the SWS scale for pH. It is important that a single scale be used throughout, as the authors have done, and there should be essentially no uncertainties in any trends found that originate from the choice of scale. The author's comment that a single scale must be used to avoid errors from the choice of scale (on the order of 0.01 pH units) is correct, but could perhaps be more clearly phrased.

The problems with this manuscript largely come from the choice of pH as a variable with which to assess long term trends. This variable is problematic for trend analysis for two reasons:

1. First (and I disagree with the authors on this point) pH measurements have large *accuracy* uncertainties (though excellent *precision* when using Clayton and Byrne's [1993] spectrophotometric analytical methods).

Yao et al. [2007] have demonstrated that impurities in meta-cresol dyes typically used for spectrophotometric analyses can lead to significant differences in pH values measured for a single sample of seawater when dye from different manufacturers is used. This work has been expanded upon by the Byrne lab [Liu et al., 2011], who show that the pH measured by individual batches of dye can be in disagreement with one another by as much as  $\pm 0.01$  pH units. It is possible that steps have been taken to adjust measured pH data for this - or that only a single batch of dye has been used for all measurements - but this large potential source for error needs to be addressed.

The lack of an appropriate reference material is an additional concern for detection of trends in pH. Tris buffers, rarely used at sea, have been shown to have uncertainties on the order of  $\pm 0.005$  in pH [Clayton and Byrne, 1993; DelValls and Dickson, 1998; Buck et al. 2002; Nemzer and Dickson, 2005; Marion et al. 2011]. Certified seawater reference materials (CRMs) have not

been available for the entire sampling period considered in this paper, so the author's claim that they have been used to check these measurements is confusing. Furthermore, CRMs are not certified for pH measurements, and therefore the use of these seawater samples for pH calibration must rely upon a calculation of pH from carbonate constants and the certified values of alkalinity,  $C_T$ , and nutrient concentrations. The uncertainties on pH values estimated from carbonate constant sets are considerable. Consider that 6 commonly-used sets of carbonate constants will yield pH values that range by as much as 0.03 in pH for seawater of standard composition [Hansson 1973; Mehrbach et al. 1973; Dickson and Millero 1987; Goyet and Poisson 1989; Roy et al. 1994; Millero et al. 2006]. This suggests that the author's unsubstantiated estimate of  $\pm 0.0034$  uncertainty is too small for these values. Indeed, their uncertainty estimate is smaller even than the pH uncertainty that would be expected ( $\sim \pm 0.008$  pH units) if their reported typical analytical accuracy uncertainty for total alkalinity ( $\pm 4 \mu\text{mol kg}^{-1}$ ) were the *only* uncertainty on pH, and there were no uncertainty in the carbonate constants themselves. Finally, even if one were to trust a CRM pH value, this would allow for only a single point calibration at the approximate pH of the surface of the North Pacific gyre, and the Clayton and Byrne [1993] method has been shown to have a pH dependence to its disagreement with pH values estimated from carbonate system equations [McElligott et al., 1998].

This trend analysis needs to demonstrate that non-systematic accuracy uncertainties are not biasing the results. The switch in methods of pH estimation from mostly calculated to mostly spectrophotometrically measured (table 1) suggests that systematic errors for these two approaches could also bias the early record relative to the late record, and could therefore lead to errors in trend estimates. The differences in methods could result in the drops in pH seen between the first two pH maps in Figure 2 (where the first is calculated and the second is determined spectrophotometrically), just as changes in the manufacturer of dye used could result in the changes seen between the second map and each subsequent map (provided this was not adjusted for).

2. Secondly, pH trends are not well-represented by multiple linear regression analysis (MLR): pH does not mix linearly, nor does it respond linearly to changes in total dissolved inorganic carbon ( $C_T$ ). Therefore - even assuming Redfield ratios of remineralization are appropriate - pH would not be expected to co-vary linearly with apparent oxygen utilization (AOU) or with physical mixing variables such as potential temperature or salinity. MLR is only appropriate when attempting to account for processes that change the properties considered in a linear way (see discussion by Li and Peng [2002] for more on how MLR implicitly represents linear mixing and remineralization). To see how non-linearities could be problematic for this analysis, consider seawater in a small container in which organic matter of a constant elemental ratio is being uniformly remineralized at a constant rate. A MLR analysis relating the various properties of this seawater at different points in time would obtain a consistent relationship between AOU and  $C_T$ , but the relationship between AOU and pH would be constantly changing as the  $p\text{CO}_2$  increases and the Revelle factor changes. This could lead to differences in MLR coefficients simply from the sampling pattern of each cruise. When you also allow for non-linear changes in pH from essentially linear mixing of potential temperature, alkalinity, and  $C_T$ , the MLR outputs become very difficult to correctly interpret.

Another problem with MLR analysis is the need to show that long term trends in water mass temperature and salinity do not also play a role. Long term warming or freshening of water masses will "move" isopycnals to deeper depths where, typically, there is more  $C_T$  and pH is lower [Wong et al., 2001; Aoki et al., 2005]. Therefore all trends demonstrated here could be instead interpreted as long term temperature and salinity trends biasing MLR results. While this

seems unlikely to be the dominant source for long term pH changes in these water masses, this source of uncertainty should be addressed.

Collectively these problems with trend detection in pH raise the question “why examine trends in pH instead of  $C_T$ ?” Ignoring for the moment that  $C_T$  has been done for this region, doing pH separately can be motivated by the importance of pH and carbonate ion availability for ecosystem dynamics. However, the specific connection between deepwater acidification and the health of deep-water benthic and pelagic organisms (deep specifically) is not made until the discussion section. When one considers that the anthropogenic  $C_T$  trend *has* been estimated [Pérez et al., 2010], it becomes less clear why anthropogenic pH trends need also be assessed. While  $C_T$  and pH do contain independent information, the determination of trends in pH from trends in  $C_T$  may be done provided one assumes that there are no long term trends occurring in total alkalinity, salinity, nutrients, or temperature. All of these assumptions are explicitly or implicitly made by this paper (implied by the choice of MLR properties). If finding independent information is the central aim, then one might look for long term trends in alkalinity. This variable responds linearly to mixing and, arguably, remineralization, and would therefore be a better choice for MLR analysis. It is also more orthogonal to  $C_T$  with respect to the carbonate system, and therefore would contain more independent information than pH given that trends in  $C_T$  have already been documented. Given the authors’ assumptions of constant alkalinity, temperature, and salinity, and given the findings of Pérez et al., [2010], the main findings of this paper simplify to “all other things being equal, seawater pH decreases when  $C_T$  increases.”

On a final note, the statistical requirements for demonstrating that observed trends are indeed the result of long-term (relative to the length of the timeseries) trends rather than decadal variability are steep. See a paper by Weatherhead et al. [1998] for in-depth analysis of what is required to demonstrate the presence of a long-term trend. Until such considerations have been taken into account, this discussion should be rephrased to allow for the possibility that the trends observed are the results of decadal variability in the relationships between the tracers used.

### Specific Comments

Format: Page Number.Line Number – “Text being commented upon” Comment

3004.13 – “Unprecedented” This is a strong word that would require justification that is beyond the scope of this paper.

3004.14 to 15 – “The latter ... water mass.” This sentence only makes sense after the next sentence is read.

3004.18 – “Lineal” Should be linear.

3005.5 – “30% reduction in pH.” Should be “30% reduction in  $[H^+]$ .”

3005.11 – “medley of juxtaposed” This is confusing.

3005.16 to 19 – “It has ... capacity.” This is misleading since it implies that the same amount of  $CO_2$  is being added everywhere. The  $pCO_2$  of the atmosphere is changing, and regions with higher Revelle factors will change their  $C_T$  less in response to a change in  $pCO_2$ . It would perhaps be best to simply mention that at high Revelle factors, a given change in  $pCO_2$  will produce a larger change in pH. However, the most ecologically relevant quantity, the change in the saturation of calcite or aragonite, is actually smaller for a given change in  $pCO_2$  in colder higher-Revelle-factor waters.

3005.24 to 27 – “The negative feedback is that as... 2010).” This sentence needs rewriting, I believe.

3006.3 and 15 – “synthetic ocean data” Synthetic should be replaced: e.g. “and ocean data syntheses (or an ocean data synthesis).”

3007.13 – “precision limits of 0.002 pH units.” The excellent precision of spectrophotometric pH measurements is only a fringe benefit for this study (by perhaps reducing the random noise around your regressions). It is important to note that the large accuracy problems for past pH measurements can greatly complicate trend detection.

3008.1 to 2 – “The estimated... is  $\pm 0.0034$ .” This comment requires justification.

3009.16 to 19 – “To be consistent ... whenever possible.” Why not neutral density... simply for compatibility? What are the limits used? Where was using the cited proposed limits not possible and why? How sensitive are the results to the choice of limits?

3010.Paragraph 1 – This paragraph is vague. It is not clear what is meant by “considering the signature” until the next paragraph, and there is no indication that the signature considered should be that of a process that changes pH linearly.

3012. 14 to 18 – “Deeper and older... organic matter.” As written, this implies that pH is lower because there are lower oxygen concentrations.

3013.23 to 29 – “To estimate the acidification... surface ocean.” This paragraph is very vague about what was done. Also, it is not clear that all natural variability was removed. If a long term trend in pH were occurring naturally, would your methods have removed it successfully? If yes, how would it not have also removed any anthropogenic trend?

3014.13 – “specific alkalinity.” “Alkalinity” would make more sense here, I believe.

3015.8 – “such trend appears.” The language is off here.

3015.8 to 9 – “rapidly formed upper LSW (uLSW) meant...” The language is off here. It implies the water meant something.

3015.20 to 23 – “In summary... faster than it would occur...” Is this a new result? If not, should it not be put in the context of other papers with similar findings for this and other water masses.

3015.24 – “plausible” This should be “possible.”

3017.14 – “worst-case” This scenario name will likely be clear to Earth scientists, but perhaps be more specific here.

3018.4 to 7 “The shoaling of the ... lysocline shoaling.” I was unsure what was meant by this or how this conclusion was reached from the data presented.

#### References:

Aoki, S., N. L. Bindoff, and J. A. Church. 2005. Interdecadal water mass changes in the Southern Ocean between 30° E and 160° E, *Geophys. Res. Lett.*, 32, L07607.

Buck, R. P., S. Rondinini, A. K. Covington, F. G. K. Baucke, C. M. A. Brett, M. F. Camões, M. J. T. Milton, T. Mussini, R. Naumann, K. W. Pratt, P. Spitzer, and G. S. Wilson. 2002. Measurement of pH. Definition, standards, and procedures. *Pure Appl. Chem.*, 74: 2169–2200.

- Clayton, T. D., and R. H. Byrne. 1993. Spectrophotometric Seawater pH Measurements - Total Hydrogen-Ion Concentration Scale Calibration of M-Cresol Purple and at-Sea Results. *Deep-Sea Res.* 40(10): 2115-2129.
- DelValls, T. A., and A. G. Dickson. 1998. The pH of buffers based on 2-amino-2-hydroxymethyl-1,3-propanediol ('tris') in synthetic sea water. *Deep-Sea Res.* 45: 1541-1554.
- Dickson, A. G., and F. J. Millero. 1987. A Comparison of the Equilibrium-Constants for the Dissociation of Carbonic-Acid in Seawater Media. *Deep-Sea Res.* 34(10): 1733-1743.
- Goyet, C., A. Poisson 1989. New Determination of Carbonic-Acid Dissociation-Constants in Seawater as a Function of Temperature and Salinity. *Deep-Sea Res.* 36(11): 1635-1654.
- Hansson, I. 1973. New Set of Acidity Constants for Carbonic-Acid and Boric-Acid in Sea-Water. *Deep-Sea Res.* 20(5): 461-478.
- Li, Y. H. and T. H. Peng (2002), Latitudinal change of remineralization ratios in the oceans and its implication for nutrient cycles, *Global Biogeochem. Cycles*, 16: 1130.
- Liu, X., Patsavas, M.C., Byrne, R.H., 2011. Purification and Characterization of meta-Cresol Purple for Spectrophotometric Seawater pH Measurements, *Environ. Sci. Techol.* doi.org/10.1021/es200665d
- Marion, G. M., F. J. Millero, M. F. Camões, P. Spitzer, R. Feistel, C.-T.A. Chen. 2011. pH of Seawater, *Mar. Chem.* Doi:10.1016/j.marchem.2011.04.002
- McElligott, S., R. H. Byrne, K. Lee, R. Wanninkof, F. J. Miller, R. A. Feely. 1998. Discrete water column measurements of CO<sub>2</sub> fugacity and pH in seawater: A comparison of direct measurements and thermodynamic calculations. *Mar. Chem.* 60:63-73
- Mehrbach C., C. H. Culberson, J. E. Hawley, and R. M. Pytkowicz. 1973. Measurement of apparent dissociation-constants of carbonic-acid in seawater at atmospheric-pressure. *Limnol. and Oceanogr.* 18(6): 897-907.
- Millero, F. J., T. B. Graham, F. Huang, H. Bustos-Serrano, D. Pierrot. 2006. Dissociation constants of carbonic acid in seawater as a function of salinity and temperature. *Mar. Chem.* 100(1-2): 80-94.
- Nemzer, B.V. and A.G. Dickson. 2005. The stability and reproducibility of Tris buffers in synthetic seawater. *Mar. Chem.* 96: 237-242.
- Pérez, F. F., Vázquez-Rodríguez, M., Louarn, E., Padín, X. A., Mercier, H., and Ríos, A. F.:Temporal variability of the anthropogenic CO<sub>2</sub> storage in the Irminger Sea, *Biogeosciences*,5, 1669-1679, doi:10.5194/bg-5-1669-2008, 2008.
- Roy, R. N., L. N. Roy, K. M. Vogel, C. Porter Moore, T. Pearson, C. E. Good, F. J. Millero, D. M. Campbell 1994. The Dissociation-Constants of Carbonic-Acid in Seawater at Salinities 5 to 45 and Temperatures 0 °C to 45 °C. *Mar. Chem.* 45(4): 337-337.
- Weatherhead, E. C., G. C. Reinsel, G. C. Tiao, X. Meng, D. Choi, W. Cheang, T. K. J. DeLuisi, D. J. Wuebbles, J. B. Kerr, A. J. Miller, S. J. Oltmans, and J. E. Frederick. 1998, Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *J. Geophys. Res.* 103: 17,149-17,161

Wong, A. P. S., N. L. Bindoff, and J. A. Church. 2001. Freshwater and heat changes in the North and South Pacific Oceans between the 1960s and 1985–94.

Yao, W. S., X. W. Liu, and R. H. Byrne. 2007. Impurities in indicators used for spectrophotometric seawater pH measurements: Assessment and remedies. *Mar. Chem.* 107: 167–172.