Interactive comment on "A novel inter-comparison of nutrient analysis at sea: recommendations to enhance comparability of open ocean nutrient data" *by* Triona McGrath et al.

Anonymous Referee #1

Received and published: 4 May 2018

Review of McGrath et al It is good to see this rather unusual detailed inter-comparison of sweater nutrient analyses methods at sea. I think the results will be useful for others and worth publishing although I would suggest some minor amendments. My only substantial suggestion is to make slightly more of the points touched on in the conclusions. The issue is not absolute accuracy and precision in its own right but data that is fit for purpose. In the deep ocean waters where nutrients are used to trace water flows and understand ocean scale biogeochemistry, these "fit for purpose" mean something very different to surface waters, where nutrient concentrations may be used to estimate biological production and its controls. The errors in quantifying these latter processes using the methods here, where detection limits are such that surface

water concentrations will be at or far below measurable concentrations, will be large. Improving sensitivity is then the issue. In deep waters the bias and accuracy issues are key. Additional specific minor points. Were reagents prepared freshly every day or less frequently? Line 213-5 I did not really understand how blanks were run on the Dal system with a Milli-Q wash. The blank is then the wash I assume. In addition at low concentration, how the zero is calculated becomes key, and this could be better explained. Line 270-9 I would assume drift samples were used to correct for drift, but here they seem to be being used as a quality control criteria and no drift assumed. Is this right and if so how long a batch run of samples was usually conducted? Line 276-7 I do not understand what is meant here by "but between drift samples" Section 3.1 I understand the point about why you might get slightly different results using different standard concentration ranges, but the issue is in part I would think about how you fit the best fit graph and whether you force the line through zero, so perhaps this should be clarified. The paper tends to assume the reader is very familiar with the Hydes et al paper and that may not always eb the case. Section 3.3 I wonder if the use of percentages here may not convey all the useful possible information since these can be misleading particularly at low concentrations. Discussion It seems to me that the Hydes et al suggestion that comparability of better than 1% is an extremely ambitious target. This might be realised by groups working in the laboratory and sharing common standards but that is not how ocean nutrient data is actually collected. A useful conclusion of this paper might be a more realistic evaluation of this comparability target by quoting the kind of comparability they see in different concentration ranges. Physical oceanographers can achieve agreement to 1% for salinity perhaps, but I think it is unrealistic for nutrient analyses and probably not necessary for most purposes. One last point. There is an implication that samples were stored frozen and then analysed on land, it would be interesting to know how these results compared. It may well be possible to achieve better comparability on land, but that would require that storage is effective and the work here might reveal if that is the case. A Minor details in the Phosphate method, what is FFD6?

C3