Interactive comment on "Carbon Flux Explorer Optical Assessment of C, N and P Fluxes" *by* Hannah L. Bourne et al.

Hannah L. Bourne et al.

hbourne@berkeley.edu

Received and published: 12 November 2018

We found an error in our response to detailed comments 1 and 18 below. Corrected responses are shown in italicized red in these sections
Response to Reviewer 2
Thank you for your review. We appreciate your insight. Below, in black are your review comments and our individualized responses are found in italicized blue.

Comments on Bourne et al. paper

SUMMARY In this manuscript the authors address a current critical research field aiming at better estimating the Biological Carbon Pump (BCP) in the ocean by the use of

autonomous in situ floats. These devices allow particle flux observations at very high spatio-temporal resolutions essential to capture the rapid ecological changes responsible in a large part for the BCP efficiency variations. In particular, this study targets a calibration between a proxy of particle concentrations in the water column, the volume-attenuance (VA) measured with a Lagrangian float-deployed imaging sediment trap, the Carbon Flux Explorer (CFE), and particle bulk chemical composition in Particulate Organic Carbon (POC), Particulate Nitrogen (PN) and Particulate Phosphorus (PP) measured on the same particles previously imaged and collected with a novel particle sampler added to the CFE (the whole instrument being named CFE-Cal). The ultimate goal of this calibration is an accurate estimation of element fluxes directly from particle imaging which thus would offer large potential in term of flux data collection which are still today and despite intensive efforts poorly spatially and temporally resolved. After detailing thoroughly the material and methods employed for particle imaging, collection and analyses the authors present results from 15 deployments of the CFE-Cal which lasted 18 to 24 hours near 150 m depth in four different locations in the California Current system selected for their contrasting primary productivity features. Results show good correlations between particle content in C and N (but not P) and VA, promising perspectives of using this autonomous in situ imaging to estimate the fluxes of these elements. Each result is discussed (Results and Discussion grouped in the same section) and focus is put on results not meeting authors expectations or not agreeing with the literature. For the results that deviate from expectations, the authors suggest possible explanations from either material malfunctions or the characteristics inherent to the different environment sampled.

GENERAL COMMENTS AND RECOMMENDATIONS This manuscript is well-written and leaves the reader with the general impression of a solid piece of work. Each section is correctly articulated and information are in general presented where they are expected. Overall, the figures and table shown are clear and deliver well the message intended. The objective tackled here is with no doubt one of the main current and future challenges in BCP research studies (converting particle flux from in situ

imaging to biogeochemical fluxes) and I am always pleased to read about work that try to push further our methods to measure these complex and very dynamic ecological processes that drive the BCP with technical innovations. Even if not realising a major advance in the field and presenting results that could be argued, especially in term of potential bias, dataset size and finding significance this study is worthy being published in BG because it is an attempt to a step forward and will certainly interest the research community working on ocean particle fluxes. However, even if I acknowledge the work done, its quality and how it is presented I have some concerns about this manuscript that lie mostly on a lack of details about the limitations of the method employed, that also reflect in the results, discussion and the general conclusion made by the authors.

(I) To obtain a good conversion from particle images to POC, PN and PP content, two key parameters have to be carefully considered: (1) the conversion from particle 2D images obtained by the CFE to their 3D volume (detailed in Bishop et al., 2016). Briefly, in Bishop et al. (2016), aggregate (including those of phytodetrital and fecal origin) volume was inferred from cross sectional area converted to equivalent circular diameter and then to volume using an empirical relationship between aggregate thickness and their equivalent circular diameter reported in Bishop et al. (1978); (2) The conversion from particle volume to their chemical content. For that Bishop et al. (2016) used a published value for aggregate dry-weight density (0.087 g cm-3; Bishop et al., 1978), and an estimated fraction of organic matter of 60 % in total dry weight. Finally, Bishop et al. (2016) uses an OM:C ratio of 1.88 to convert the estimated OM weight to POC.

Following Bishop et al. (2016), our optical results are calculated as Volume attenuance (mAtn-cm2). Bishop et al. (2016) had to estimate POC in the CFE imagery as no calibration samples for POC and PN were collected. In this study, we were able to collect corresponding POC samples using the two CFE-Cals. We therefore did not need to do any conversions of 2D images to 3D volumes to estimate carbon as was done in Bishop et al. (2016). In this paper, we report that the estimate of VA:POC ratio estimated by Bishop et al. using a 2D to 3D volume conversion formula and

C3

information on particle density and carbon contents from Bishop et al. (1978), differed by a factor of 3 from the directly calibrated VA:POC relationship. This is shown in figure 6.

(II) The authors highlighted clearly the problem of using these literature-based calibration factors as they are often applicable only in the limited spatio-temporal context of their formulation. But I hardly understand the aim of the present calibration if it is not to finally succeed at reconstructing the flux from images alone and from a large range of environment and ecosystem structures. From the way it is currently presented, the manuscript suggests that the authors are trying to establish a library of relationships between VA and C, N, P contents. If it is the case it should be clearly stated.

Yes, our goal is to reconstruct flux from images. In the future, using the CFE-Cals to collect samples from more regions will make the calibration more robust and widely applicable. We will emphasize these goals more clearly.

(III) I would have found very interesting to see in the discussion and conclusion sections some perspectives on how to improve the calibration presented here. In particular, the combined acquisition of particle images and measurements of POC, PN and PP done here offers the great opportunity of estimating the quality of a traditional flux reconstruction (i.e. by inferring its value from the images using published volume to organic contents conversion factors as done before) by comparing it to the real values measured here (as explored in Estapa et al., 2017). I assume that the final goal is to estimate the POC flux from images alone and much work is yet to be done by the community to understand how to translate small differences in image detection to potentially large differences in chemical contents.

A comparison between the calibration method developed here and other methods that try to convert images to elemental fluxes should have been made. The use of Polyacrylamide or Cryogels sediment traps to collect particles and then use image

C4

analysis and published values of organic content to convert the images to fluxes is a very close approach to this study. The major advance that the present work could have brought is by extending the use of particle images to push further the estimation of their organic content from the image analysis. It is a bit disappointing to finally realise that this study has the great potential of presenting both the images and the "true" values of their content usable to further our understanding of observed discrepancies but that unfortunately this opportunity was not seized by the authors.

We appreciate the reviewer's suggestion. Such a study of gel trap samples and CFE imagery needs to be done on the same platform and same time and with careful attention to the photography and illumination of gel trap samples. We'd love to facilitate such a comparison in the future but this activity is beyond the scope of this paper. The best same place, same platform, same time study is that of Estapa et al. (2017) who used 2D size distributions in gel trap samples to estimate carbon. They did not extend the analysis of gel trap imagery to attenuance units. The main requirement is that the illumination source intensity (without particles present in the gel) must be mapped precisely as we do with the CFE. Variations in gel homogeneity and thickness degrade particle detection.

Also, the limited number of results obtained due to device malfunction or inherent to the properties of the particle flux collected (i.e. presence of swimmers), or the corrections of POC, PN and PP values obtained from the CFE004 dives due to a discrepancy between images and sampling, should have led the authors to much more caution in their conclusion.

Instead, the authors claim "strong calibrations" between the VA and POC-PN contents for a dataset on which many values have been removed or multiplied by an empirically-determined factor; in this case the 1.45 times factor representing the difference of abundance of ovoid pellets in the sampler and from the images.

Based on all these general remarks, I still recommend this article for publication in BG

but after significant changes have been made to the Results, Discussion and Conclusion sections and substantial evidences provided where required.

In particular, I strongly advise the authors to focus on the general issues mentioned above and summarised as follows:

(A) Add to the manuscript a comparison with other techniques of image conversion to biogeochemical fluxes (e.g. gel sediment trap analyses).

This recommendation is addressed in (III) above. This is worth-while to do when such particle collections are made simultaneously, ideally on the same platform. We do not agree that comparison to gel trap sample size distribution analysis should be brought into the present discussion.

(B) Use the dataset presented to explore further the known discrepancies between image analysis and inferred organic content. The authors could investigate if a reconstruction of the fluxes measured by the sampler here would be feasible by using the corresponding images and by applying various volume to chemical content relationships to the different particle types identified (e.g. different relationships for fecal pellets, marine snow, etc.). New insights informing on why we struggle at inferring the flux from images would certainly increase significantly the impact that this manuscript will have on the research community.

In this paper we show that the VA:POC and VA:PN were largely independent of particle size distributions and class (Figs. 3 and 5). What we meant was that we were impressed by the fact that Anchovy pellet dominated samples fell on the same line for C and N as the amorphous "snow-like" aggregates. We do not think a discussion applying volume-carbon relationships for discrete particle classes to compare with our direct calibration results should be added because (1) such construction of morphologically distinct particle class size distributions from imagery is highly time consuming and (2) models are limited. For these reasons, our group last published on particle flux derived from known particle class concentration size distributions 30

years ago. For example, Bishop et al. (2016) used equations relating aggregate particle size to volume, particle density, and carbon content from particles sampled in the Cape Basin in the South Atlantic (Bishop et al., 1978) to derive particle fluxes in the California coastal waters. We show in this paper that results of predicted VA:POC from Bishop et al. (2016) were a factor of three different then the directly calibrated relationship. We are not sure what volume metric transformations will add to this paper. We are encouraged by our correlations of sample loading and VA.

(C) Depending on the modifications made after (1) and (2), moderate if needed the stated significance of the results and discuss it more objectively and into details. Especially, the term "strong correlation" can hardly be used with such confidence knowing that the dataset has been trimmed and partly multiplied by an empirically-determined factor, and that authors seem themselves unsure about potential contaminations of their samples.
These points are addressed below in the detailed comments section.

Additionally, below are more detailed comments on the manuscript including technical and typographical corrections that will need particular attention before publication. I advise the authors to give a special attention to the four questions/comments below marked with an asterisk (*) as their response should influence the final decision for publication.

DETAILED COMMENTS

1) Page 1, Line 17: please add to the R2 the size of the sample included in the fit (n) and the p-value.
Yes, here they are below, and they will be added to the manuscript
POC, n=13, p=5.96x10^{-8}
PN, n=16, p=1.91x10^{-10}

Should be:
Yes, here they are below, and they will be added to the manuscript
POC, n=12, p=5.2^{-8}
PN, n=15, p=5.7^{-10}
The n values for POC and PN should be 12 and 15 respectively as shown in the reviewed manuscript figure 6 and table 1. The point that was not shown in the figure (CFE-Cal2 dive 111) was eliminated due to the presence of an actively swimming cm-sized gelatinous organism in all images for this dive. All images showed particles and the organism moving around and attenuance estimates ranged by a factor of two. The 16th point was eliminated for these reasons.

2) Same line: "...was not sensitive to environment or classes of particles sampled." I assume this statement is used as a proof of applicability of the current calibration to many different ecological contexts. But, it could also suggest that the environment where the deployments were made was not contrasted enough for this calibration.
Figure 3 and 5 were intended to contrast particle size distributions and classes found at each location. We will add clarifications of details of the variability of the environments and size distribution at each collection site.

3) Page 1, line 21: a space is missing between "Approximately" and "10".
Yes, will fix.

4) Page 3, line 11: change "our 2.8 conversion factor" to "the 2.8 conversion factor obtained by Bishop et al. (2016)". I understand it is the same team but "our" would mean version Discussion paper a factor inferred in the current study and it is not the case.
We will change this.

5) Page 4, line 9: the glass stage appears quite small and subjected to overload if a single cm-sized particle (or a few mm-sized particles) happened to enter in the trap. What is the diameter of the opening?

The diameter of the high aspect ratio funnel is 15.4 cm, the stage diameter is 2.5 cm. The funnel is covered by 1-cm hexagonal light baffle cells. Bishop et al. (2016) provides this full detail. We address sample overloading in the response to point (6).

6) Page 4, line 13: The time of 25 min seems critical. Is there a threshold at which volume attenuance can be biased by particle overload (particles accumulating over previously deposited particles on the stage)? How did the author choose this time and the time of 1.8h mentioned below (line 14) as it seems dependent upon the amplitude of the particle flux at the time and depth of the deployment?

The times were chosen to allow for high temporal resolution of flux and since we were not power limited (CFE can operate for 8 months at hourly frequency) we chose by default cleaning times to be the same as times used by Bishop et al. (2016). Overlapping particles do not bias attenuance flux as their contributions are additive. (see discussion in Bishop et al., 2016 and also our response to Reviewer 1 on this point). Implied is the question of whether or not attenuance is saturated (e.g. transmission is 0). The highest fluxes recorded by the CFE were those from January 2013 discussed in Bishop (2016). Even at this high rate of flux, attenuance was not saturated. We have not seen evidence of saturating attenuance in our images from CCE-LTER.

7) Page 6, line 5: do particles larger than the size limit of 3 mm can get stuck inside the openings?

Aggregates and rare gelatinous organisms larger than 3 mm do not get stuck. In one of our 23 deployments, a larval crab became stuck in the stage area and had to be removed after recovery. We will add a discussion of this in the text.

8) Page 6, lines 7-8: I assume the CFE-Cal has not yet been used for trace metal studies then (this intended use is mentioned above in the manuscript)?

Yes, we designed the CFE-Cal so that in future it can be used for trace metal studies. More refinement is required.

9) Page 7, line 3: peri slides. Please correct.

This will be fixed.

10) Page 7, line 16 and below: why giving results in the Material and Methods section?

We will move this to the results section.

11) * Page 7, lines 19-22: this will need clarification as it seems to be a very serious issue. How could the process blanks be higher than the samples themselves even in case of accidental collection or contamination? Over the 6 replicates of process blanks, how many were contaminated? How did the authors deal with this issue as blanks have to be subtracted from sample values? Are the negative values on Fig. 5A a result of this correction?

We will modify this section to provide further clarification. We did subtract the blank value from the sample values, as stated in equation 1. This drove one sample negative, though not negative within error. This one sample that was driven negative was collected from location 3. Fluxes at location 3 were very low - an order of magnitude lower than samples collected in other regions. We note that the N for this region was not negative.

12) Page 7, lines 24-25: "... which we assume is based on sample heterogeneity". Do the authors have evidence to support this assumption?

The samples were punched from the filter as described in the methods. We retrieve

punches evenly distributed across the filter, but inevitably as there are discrete particles on the filter, there is some heterogeneity between the sub-samples. The sample is not homogenized and then sub-sampled. As this same process is done for every sample, we applied the RSD in the error calculation. We will clarify this in the text. Figure A3 depicts particles dispersed across filters.

13) Page 8, lines 24-25: please mention what would have been the total number of samples in case of no malfunction or swimmers and give a percentage of "fail". My point is that it is hard to estimate the robustness of the CFE-Cal without a proper estimation of its percentage of fail (how many successful dives/samples over the total number intended?).

Of the 60 dives, there were 8 dives that failed due to a malfunction (either the float not diving, or the sampler not working). These malfunctions occurred almost entirely at the beginning of the cruise. One dive had a gelatinous swimmer and one had the larval crab. As all instrument malfunctions were resolved, future deployments will be far more robust. We will add a discussion of this to the text.

14) Page 8, line 27: this is not really a measure of "collection efficiency" (only assumed) but more a measure of transfer efficiency between the imaging stage and the bottles.

We will change the heading to transfer efficiency.

15) Page 8, line 29: "...close...", please give a precise number.

Counts were within error of being the same. We will describe this result statistically.

16)* Page 8 line 30: again this is a very worrying result that needs more investigation as it suggests a real issue with the collection and/or transfer method employed. * Page 9, line 1 and lines 3-4: the authors first state that they do not fully understand the issue and then claim to have addressed the problem by solving a software issue.

Please bring clarification on this.

We do understand the issue better than our very brief description implied. We describe the software problem and how we fixed it below.

The optical encoder that provides feedback to the motor which turns the sample selector valve was programmed to time out after a certain period of time if the correct "home" position could not be found. During lab tests, where the CFE-Cal was deployed in a large water tank and monitored by video, the sampler would find its correct position in the allowable time. However, these tests were done in a lab, with room temperature water and not in $10°C$ water under pressure.

We discovered midway through during deployments that the sampler on CFE4 consistently failed to find the "home" position during pre-deployment checkout. From the "home" position, programming logic advances the sampler by counting encoder pulses to reach open and closed positions for particular sample bottles. This seems to explain the over transfer of particles to this CFE4-Cal. Once this issue was discovered, we increased the time significantly. This issue was not encountered again once it was addressed. We will provide more details on the issue in the text.

17) Page 9, line 12: what does this time of 2 minutes sample collection time refer to? (how is it different to the 25 min imaging sequential time?). Is it the duration of particle transfer to the bottles?

This is the duration of particle transfer time during a cleaning cycle. A cleaning cycle takes 2 minutes to complete. The 25-minute timing is the time interval between successive image sets after a cleaning cycle. We will clarify this.

18) Page 9, line 17: please add the sample size (n) and p-values for each regression fits. Page 9, line 17-25: being "not typical of sinking particles" is certainly not a valid reason to exclude these values from the dataset. Authors are required to provide valid reasons here (e.g. why these C/N ratios would make these particles not wanted in this dataset?).

We will add regression statistics for all 16 POC points (the 3 excluded ones are already shown in the figure) and report the slope, n and p in the text. See also response to 19 below.

Should be:
We will add regression statistics for all 15 POC points (the 3 excluded ones are already shown in the figure) and report the slope, n and p in the text. See also response to 19 below. The total number of POC points is 15, not 16. See (1) above for detail.

19) * Page 9, lines 26-32: again this is a very serious issue. If the sampler building material is potentially responsible for contaminating the samples, how can the authors be confident that not all their POC and PN values are biased by chemicals leaked from this 3D printed part?
Any DOC or DON leaked from the 3D printed part would not be retained in the collected sample. The reason that we excluded the 3 POC points from the regression (see point 18) is that (a) we found visually obvious particles on the filter that were the support material used in the 3D printing process. (b) The C/N of the other 12 samples were all consistent with natural populations. These are in line with C/N that have been found in the region previously (C/N = 11.1 at 100m, Stukel, 2013). As the 3D printed material contains no nitrogen, C/N values would elevated if they were contaminated. In decades of particle collection, Bishop has not found natural material with C:N ratios as high as 20. We will clarify the text.

20) Page 10, lines 1-4: this also seems to be pure speculation without any evidence of TEP presence in samples.
This was a hypothesis which we will remove.

21) Page 10, line 8: if I understand well, the objective of this calibration is to ul-

timately allow an estimation of biogeochemical fluxes from in situ imaging that could be applied to the largest range of particle types and flux amplitude. It seems very contradictory then to remove particles from the dataset because they are inherently different from the rest of the flux to improve the goodness of the fit. This is very troubling as it suggests that the authors don?t fully comprehend their ultimate goal here. Page 10, line 10-13: this is precisely why it seems so hard to reconstruct a biogeochemical flux from images alone. I strongly suggest that the authors use this example to illustrate the difficulty of meeting the challenge addressed here and impartially discuss their results following the approach of reconstructing the flux from its various particle types having contrasted chemical contents (see general remarks above).
In the lines you mention, we discuss how the high P content of anchovy fecal pellets, combined with the fast sinking rates, led to phosphorous loadings far higher than the other samples. Though the ultimate goal is to allow an estimation of biogeochemical fluxes based on image analysis, we concluded that we cannot predict particulate phosphorous based on VAF, precisely because of the highly heterogeneous nature of phosphorous in particles, such as the anchovy pellet. In other words, as P is highly labile, we found it was impossible to estimate PP based on in-situ imaging. Our intent was to show that even when eliminating that particular point, the relationship between PP and VAF was still far less robust than that of POC or PN. We will clarify this point more clearly in this section.

22) Page 11, line 26: please remove "strong" as it does not seem appropriate. Provide n and p-values.
We will change strong to well-correlated. n and p values are provided.

23) Page 11, lines 26-27: " that apply over a wide range of environments". This statement could be made with confidence only if the deployments were made in different oceanic regions, seasons and water column layers. It appears too early at

this stage to claim this. Page 11, line 31: "... insensitive to particle classes dominating export". This is not true and is directly contradicted by previous findings shown in this manuscript (see observations made by the authors about the anchovy fecal pellet flux). Please amend as required.

Yes, but we were impressed by the fact that the POC and PN of anchovy pellet dominated samples fell on the same line as the amorphous "snow-like" aggregates. We will modify this to better describe the specific range of environments and particle classes encountered as in P1 Line 17. We will clarify that in the future we hope to make the calibration more robust by further collecting samples from different regions, seasons and depths.

24) Page 12, lines 5-8: this is confusing and again suggests contradictory intentions of the authors. It is still unclear at this very end of the manuscript if the authors intend to establish a library of VA:element fluxes relationships for each environment and ecological settings sampled (the use of one specific slope would then be reusable to infer the biogeochemical fluxes from images taken in the corresponding region, time of the year and depth), or if they intend to find a general relationship usable in many oceanic regions, environments and ecosystem structures. In both cases, an extensive work remains to be done and it should be clearly stated.

Our results are a first step towards expanding the range of particle flux that may be retrieved through optical methods. We will rephrase the last sentence. We will also add a discussion of future work to be done.

C15