Sampling uncertainties of particle size distributions and derived fluxes

K. M. Bisson[1], R. Kiko[2], D.A. Siegel[3], L. Guidi[2], M. Picheral[2], Emmanuel Boss[4] and B. B. Cael[5]

[1]Department of Botany and Plant Pathology, Oregon State University, Corvallis, USA

[2]Laboratoire d'Océanographie de Villefranche-sur-mer, Sorbonne Université, Villefranche-sur-mer, France

[3] Earth Research Institute and Department of Geography, University of California, Santa Barbara, Santa Barbara, USA

[4] School of Marine Sciences, University of Maine, Orono, Maine 04469, USA

[5] National Oceanography Centre, European Way, Southampton, SO143ZH, UK

Corresponding author: Kelsey Bisson (bissonk@oregonstate.edu)

Key Points:

- We model uncertainty in UVP5-derived PSDs and fluxes via Bayesian Poisson statistics and a truncated power law distribution.
- When sampling volume isn't sufficient to characterize the PSD (and rare particles), we recommend modeling the data with a truncated power law.
- Modeled carbon flux calculations have an uncertainty of ~50% arising from sampling uncertainty alone.

Abstract

Here we provide a method to quantify the uncertainty associated with sampling particle size distributions (PSD), using a global compilation of Underwater Vision Profiler observations (UVP, version 5). The UVP provides abundant in situ data of the marine PSD on global scales and has been used for a diversity of applications, but the uncertainty associated with its measurements has not been quantified, including how this uncertainty propagates into derived products of interest. We model UVP sampling uncertainty using Bayesian Poisson statistics and provide formulae for the uncertainty associated with a given sampling volume and observed particle count. We also model PSD observations using a truncated power law to better match the low concentration associated with rare large particles as seen by the UVP. We use the two shape parameters from this statistical model to describe changes in the PSD shape across latitude band, season, and depth. The UVP sampling uncertainty propagates into an uncertainty for modeled carbon flux exceeding 50%. The statistical model is used to extend the size interval used in a PSD-derived carbon flux model, revealing a high sensitivity of the PSD-derived flux model to the inclusion of small particles (80-128 μm). We provide avenues to address additional uncertainties associated with UVP-derived carbon flux calculations.

1 Introduction

In the ocean, an extraordinary range of particle sizes (from < 1 μm to 30 m; including non-living dust particles, detrital matter, bacteria, phytoplankton, zooplankton including salp chains, whales and many others) influence ecosystem structure and function, net primary production, particle sinking, and carbon flux (Sheldon et al., 1972, White et al., 2015, Alldredge and Gotschalk, 1988, Siegel et al., 2014). Over the last decade, bio-optics has enabled the characterization of portions of the particle size distribution (PSD) (Boss et al., 2001, Slade and Boss, 2015, Dall'Olmo et al., 2009, Reynolds et al., 2010, Chase et al., 2020, Stemmann and Boss, 2012, Cael and White 2020, Giering et al., 2020 and refs therein), especially through the advancement of in situ imaging technologies.

In order to use PSD observations in the most meaningful way in analyses and models, the uncertainty associated with the observations must be clearly quantified. The primary focus of our analysis herein is to provide a method to characterize PSD uncertainty arising from sampling volume. In situ observations of the PSD are a function of both the true size structure of the particle

assemblage and of the measurement method. In this study we focus on PSD data collected from the Underwater Vision Profiler (UVP, Gorsky et al., 2000, Picheral et al., 2010), which 'sees' a narrow size range (60 μm – 20,000 μm capabilities, Lombard et al., 2019) of living and non-living particles which are imaged within a small fraction of the water column (anywhere from 0.28 L to 10.5 L depending on the UVP version, Guidi et al., 2008). The surface area of pixels containing a particle is converted into an assumed equivalent spherical diameter using instrument specific calibrations (Picheral et al., 2010), no matter how a particle is shaped or oriented (introducing error into the retrieved particle size e.g., Karp-Boss et al., 2007).

Uncertainties in particle detection are propagated downstream into calculations of carbon flux and other applications, which rely on both accurate PSD observations as well as appropriate modeling to convert standing stocks of PSD observations into rates of sinking carbon across the full range of depths and particle types in the ocean. When PSDs are not used directly for calculations of flux or other quantities of interest, PSDs are commonly described with a power law to reflect the rapid decline in particle concentrations with increasing particle size (e.g., Jonasz and Fournier, 2011). However, the power-law exponent estimation is sensitive to the abundance of rare large particles, and the behavior of power-law distributed quantities (e.g., carbon flux) is sensitive to the exact values of the power-law exponent. In any natural system, a power law is a simplistic generalization of the true particle size distribution and is only applicable over a finite size range (and this size range must be adequately accounted for). A secondary aspect of our analysis herein is to model PSD observations with a truncated power law rather than a power law to better account for rare instances of large particles observed by in situ instruments. Moreover, a truncated power law distribution has an extra parameter about the particle size range for which power-law behavior holds, which offers more information about the shape of the particle size distribution than a scaling exponent alone.

In this study we quantified the sampling uncertainty associated with UVP observations as well as the error associated with extrapolation to other size classes. As a test of how UVP sampling uncertainties propagate into derived properties, we calculated carbon flux using both observed and modeled UVP particle concentrations over various size intervals. We discuss implications for the 2 retrieved parameters of the truncated power law distribution and we provide recommendations for future flux modeling of the PSD.

2 Materials and Procedure

2.1 UVP Data

Profiles of PSD observations used in this study come from Kiko et al., 2021, which synthesized observations from the UVP5 models only (Figure 1A). This dataset underwent very little processing prior to our analysis. All data were already binned to 5m vertical bins, and the reported particle concentrations are within standardized and consistent size bins, starting at 128 μm for this dataset. For each depth we multiplied the particle concentration (# L^{-1}) by the sampling volume specific to each depth in order to retrieve $\boldsymbol{N(d_i)}$, or the total number of particles for a reported equivalent diameter size range (given by the $\boldsymbol{i}$'th size bin). The PSD data reported here includes all living and non-living particles, and all data are inter-calibrated according to procedures described in Kiko et al., 2021.

Since its invention, the UVP has undergone design improvements so that its size and sampling speed are compatible with a standard CTD rosette. The UVP5 (Picheral et al., 2010) has an image acquisition frequency varying between about 3 to 20 Hz depending on versions and particle load of the water column (higher loads require more processing time and therefore a lower acquisition frequency). During normal CTD deployments with speeds up to 1 m/s, this allows imaging of up to 1 L/image at the highest frequency of 20 Hz and 20 L/m on a vertical profile at 1 m/s. The surface area of particles is converted from pixel counts (using instrument settings), and the equivalent spherical diameter (ESD) is calculated following $\boldsymbol{aa} *$

$number\_of\_pixels^{b}$ where aa and b were determined through calibration casts in the bay of Villefranche. Hereafter any use of the UVP is implied to mean UVP5 in our study.

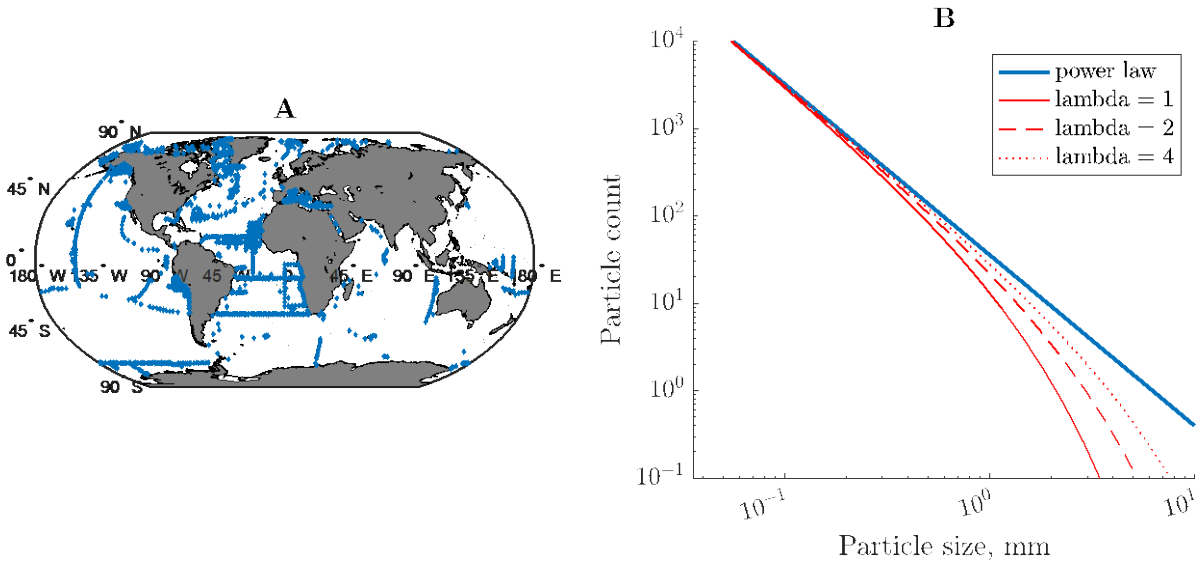

Figure 1. A. Location of UVP5 observations (blue). B. Comparison of a power law (blue) with a power law with an exponential cutoff of various λ values (red lines). All lines share the same α. The power law is of the form $N(d) = d^{-\alpha}$ while the truncated power laws follow $\sim d^{-\alpha} * e^{\frac{-d}{\lambda}}$.

2.2 Extrapolation and sampling uncertainty calculations

A UVP measurement of the PSD is an estimate of the true particle population in the water column. If a UVP samples N particles within a range of sizes with average diameter d in a volume V, intuitively the best estimate for the concentration of d-sized particles is N/V, and the larger V and/or N the better an estimate this will be — but what is the uncertainty associated with this estimate, and how does it depend on N and V? How do these uncertainties ultimately propagate into uncertainty in estimated flux? This problem is intractable to quantify perfectly, but may be substantially simplified by assuming Poisson statistics, which describe the probability of a certain number of events (e.g. particle counts) occurring in a fixed interval of space if such events occur at an average rate (e.g. the particle concentration). If some very large volume V has N particles of a given size in it (or within a given size range), the concentration of particles of that size in that volume is C = N/V. If some smaller volume *v* of that large volume V is being sampled, each particle has a probability of *v*/V of being sampled. If *v* is much smaller

than V, as is often the case with oceanographic measurements where we take small-volume measurements as representative of much larger oceanic volumes, and/or N/V is small so that a small number of particles n is sampled in v, then the uncertainty in the true concentration N/V based on the measured concentration n/v may be appreciable.

Assuming Poisson statistics, we take a Bayesian approach to finding the best estimate and uncertainty in the true concentration given the measured concentration. The Poisson distribution expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur independently with a known rate (Haight, 1967). Thus, if the probability of individual particles being sampled by the UVP is independent, and the concentration of particles of mean size d is some concentration C, then the sampled concentration follows the Poisson distribution. In Bayesian inference, the conjugate prior for the rate parameter of the Poisson distribution is the gamma distribution (Fink, 1997). This means that given a same sample of N measured particles of size d, and assuming a prior of Gamma(k, θ) the posterior distribution is $C \sim$ Gamma($k + N$, θ). (In Bayesian statistics, the prior is an assumption that quantifies prior knowledge about a quantity before evidence is taken into account, and the posterior distribution quantifies that same quantity after taking evidence into account.) In our case we have little information with which to form a prior, so the best prior is the maximum entropy (i.e. least informative) Jeffreys prior – Gamma(1/2,0) (Lunn et al., 2012). Altogether this means that if we measure N particles in a volume V, we get a posterior distribution for the concentration C of $C \sim$ Gamma($N + \frac{1}{2}$, $1/V$). This distribution has a mean of $\frac{N}{V}$, matching our intuition, and a standard deviation of $\sqrt{N}/V$. Sample volume and sampling uncertainty are thus inversely related, and for the same sample volume, the relative uncertainty is larger for lower measured concentrations.

We can also use this distribution to estimate how sample uncertainty propagates into estimated fluxes or parameters of a truncated power law using its posterior predictive distribution — the distribution of possible unobserved values conditional on the observed values — which in this case is the negative binomial (NB) distribution (Gelman et al., 2014). If we measure N particles in a volume V, then the distribution of possible unobserved values that accounts for uncertainty in the true concentration given these measured values is NB(N+1/2,1/(V+1)).

To estimate uncertainty in the fitted α and λ values from $N(d_i) = C_1 * {d_i}^{-\alpha} * e^{\frac{-d_i/d_0}{\lambda/\lambda_0}}$ (where λ_0=1 mm, d_0 = 1 mm, C_1 is concentration scaled at 1 mm, and i is the discrete size bin), and in the modeled carbon fluxes, we thus draw 100 random samples from NB(N+1/2,1/(V+1)) for each particle size class at each sampled place and time (using 'nbinrnd' function in MATLAB), effectively generating 100 PSDs from each observed PSD. These calculations were run at all places for depths 50 and 300 m to retrieve α and λ and the coefficient of variation of each. We also calculate carbon flux (described in section 2.4) for each of the 100 simulated PSDs. The coefficient of variation is reported as the standard deviation (σ) normalized by the mean, and relative error ($error_{rel}$) is given by the σ divided by the simulated $\boldsymbol{N(d_i)}$, multiplied by 100%, or $error_{rel} = \frac{\sigma}{N(d_i)}$.

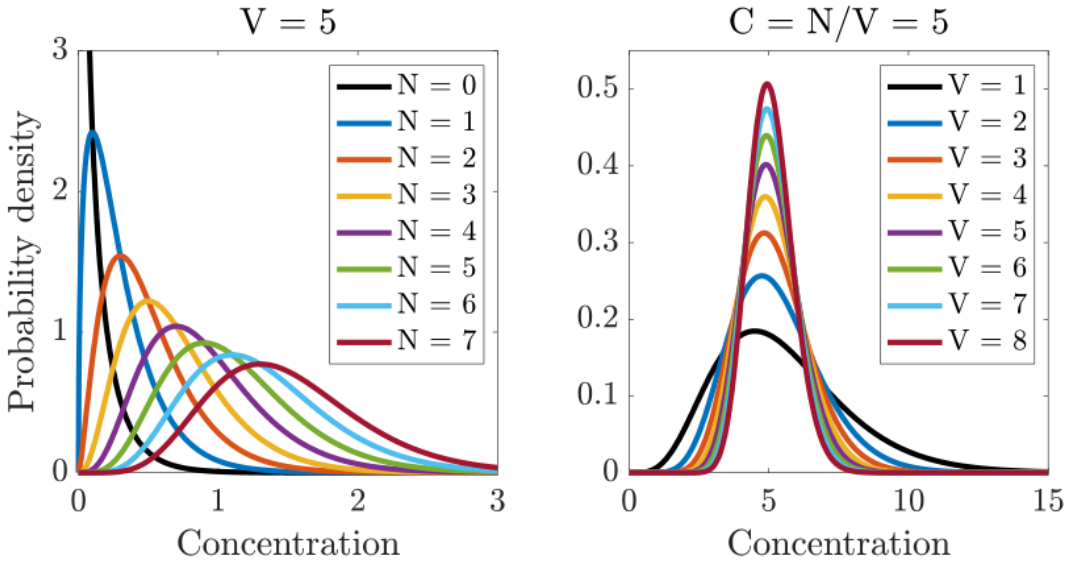

Figure 2. Theoretical probability (Poisson distribution) of particle concentration (N/V) based on observed particle number (N) or sampling volume (V).

The uncertainty associated with sampling volume is visualized using probability densities for particle concentrations for either fixed or variable sampling volumes (Figure 2). When sampling volume is fixed (and assumed to be 5 L), the width of the probability distribution increases substantially as particle count increases for an arbitrary size class (compare maroon line to black line, Figure 2, left plot). Essentially, if the observed particle count is 5, the true concentration in the water column is likely to be between 0.5 and 2 (green line). For fixed concentrations

(assumed to be 5 particles per L, Figure 2, right plot) and variable sample volumes, the probability that the true concentration of particles is accurately measured by the UVP scales with sampling volume. Higher sampling volumes (8 L, maroon line, Figure 2, right plot) result in narrow probability distributions that give higher fidelity to the observed particle concentration. Lower sampling volumes (black line, 1 L, Figure 2, right plot) have a wider probability distribution, where it is evident that the true particle concentration can be a factor of 2 (and greater) different than what was observed.

2.3 Modeling the observed PSD

Here we modeled observed PSD from the UVP (Figure 1) using a truncated power law, *i.e.* a power law with an exponential cutoff, which is simply a power law multiplied by an exponential function, or

$$\boldsymbol{N(d_i) = C_1 * d_i^{-\alpha} * e^{\frac{-d_i}{\lambda}}}. \quad [1]$$

We chose a truncated power law functional form because continuous PSD observations are commonly modeled with a power law, but UVP data commonly have a drop off in observations at larger particle sizes due to the relatively small sampling volume. This drop off is better captured with a truncated power law. We did not consider other functional forms in this study. $\boldsymbol{N(d_i)}$ is the number of particles within a given size bin (i) and normalized by the bin width (because the bin width is not the same width for each size across the size spectrum, and because normalizing by bin width allows the PSD to be independent of bin width and thus comparable across instruments), d is the equivalent spherical diameter, and α and λ are free parameters. $\boldsymbol{N(d_i)}$ is a discrete form in practice as $\boldsymbol{d}$ is not continuous but rather represents discrete size bins. It is implied that both d and λ are normalized by $d_0 = \lambda_0 = 1$ mm, everywhere d and λ are operated on in this text. The leading coefficient C is the concentration at d=1 mm divided by $e^{\frac{1}{\lambda}}$ for each PSD. The available sizes for d range from 10^{-3} to 26 mm, but operationally, the minimum observed particle size from the UVP5 falls into the 128-161 μm size class. Conceptually, α is a typical power law scaling exponent and λ is the upper limit until which the particle size distribution is well-described by a power law. High values of α are associated with a steep PSD slope, or a particle assemblage dominated by many small particles relative to larger

ones. Low values of λ are associated with a steep decline in $\boldsymbol{N(d_i)}$ earlier in the size spectrum (Figure 1b, solid red line compared to red dashed line).

Prior to model fitting, UVP observations of particle concentration (# L^{-1}) were multiplied by the sampling volume (L) specific to each depth, log10-transformed, and normalized by the bin width (mm) of each size class. We performed a weighted nonlinear optimization of the truncated power law parameters by minimizing the following cost function,

$$\boldsymbol{cost} = -\sum_{\boldsymbol{i=1}}^{\boldsymbol{i=n}} \boldsymbol{W_i} * \left|\boldsymbol{log10}\left(\boldsymbol{N(d_i)}\right) - \boldsymbol{log10}\left(\boldsymbol{PSD_i^{obs}}\right)\right| \quad [2]$$

where W_i is the weight for each bin (i) is the sampling volume divided by the relative sampling error of each size bin, or

$$\boldsymbol{W_i} = \frac{\boldsymbol{V}}{\boldsymbol{error_{rel}}} \quad [3]$$

Note we do not log-10 transform the weights because the weights are a function of relative and not absolute error. The model fitting was performed in MATLAB (using 'fminsearchbnd') over the observed particle size interval for each specific instance depending on the depth and location of observations. We constrained α to be between 0 and 6 (and note there was only 1 instance out of nearly 70,000 where the retrieved α value was a boundary value). The α range extends slightly beyond the range of observed power law scaling exponents for PSDs (Diehl and Haardt, 1980, Buonassissi and Dierssen 2010), in order to reduce boundary effects during fitting, and we constrained λ to be contained within the bounds of the smallest and largest observed particle size for a particular $\boldsymbol{PSD}$, with 0 instances where λ equaled a boundary value. Because λ spans several orders of magnitude, any reported λ averages for the remainder of this text were calculated using log10 transformed λ values and those averages are then converted into non-log transformed values that are simpler conceptually. We performed this model fit for all 7808 independent locations at the mean of depth bins, or 7.5, 22.5, 47.5, 97.5, 147.5, 222.5, 297.5, 497.5, and 997.5 m (hereafter expressed as 10, 25, 50, 100, 150, 225, 300, 500, and 1000 m).

The truncated power law model is a slightly better fit to the data than a power law, with an improved adjusted R-squared (accounting for free parameter differences, 0.96 for a truncated

power law versus 0.95 for a power law), relative percent error (24% for a truncated power law versus 27% for a power law), and relative bias (i.e., $100\% \times \frac{1}{n}\sum_1^n \left(|N(d_i) - PSD_i^{obs}|\right) \Big/ PSD_i^{obs}$, 7% for a truncated power law versus 9% for a power law) across all depths. In this study we choose a truncated power law because of the higher overall performance, and because the truncated power law parameters offer more insights about the observed PSD shape than a power law alone.

2.4 Application to carbon flux calculations

The applications of measured in situ PSDs introduce additional uncertainty and error into the derived measurements of interest, including quantifications of carbon flux (Guidi et al., 2008; 2016) and aggregate formation (Guidi et al., 2009). PSDs are ingested within a power law approximation to calculate carbon flux via

$$\boldsymbol{F = \int_{d\_min}^{d\_max} N(d) * Ad^b \, dd}, \quad [4]$$

where F is carbon flux (mg m^{-2} d^{-1}), $N(d)$ is the concentration of particles (# mm^{-1} L^{-1}) with a mean equivalent spherical diameter d (mm), A = 12.5, and b = 3.81 are free parameters that were first optimized in Guidi et al. (2008) using all available UVP versions with a shared size interval of 250 µm to 1.5 mm, and with sampling volumes ranging from 0.28 to 10.5 L. The function is integrated over the range of size classes available. While A and b are empirically derived, they conceptually arise from a general mechanistic model that incorporates sinking velocity (via Stokes' law, a power law) and carbon content of a particle (modeled as a power law). The product of sinking velocity ($w(d) = \beta d^2$) and carbon content ($m(d) = \alpha d^3$, both power laws) are modeled as a power law, providing the $\boldsymbol{Ad^b}$ term in equation 4. Given typical power law fits for $\boldsymbol{N(d_i)}$, equation 4 implies an infinite flux with increasing particle size, as well as a consistent size-to-flux relationship for equally sized cells, which will be violated for cells of different density and/or lability. We argue here that any PSD-derived flux formula must be aligned with the known uncertainties of the PSD observations. Particularly, the value of d_max is important if a power law $\boldsymbol{N(d_i)}$ is selected because the counts of N(d_max) become negligible due to sampling. The value of d_max is also important when comparing across different UVP versions with different size ranges.

We calculated flux using direct observations of UVP PSD via equation 4, as well as using the modeled PSD derived from equation 1, or

$$\boldsymbol{F_m = \sum C_1 * d_i^{-\alpha} * e^{\frac{-d_i}{\lambda}} * A d_i^{b}}. \quad [5]$$

Where F_m denotes flux from the modeled number distribution. We tested different values of $\boldsymbol{A}$ and $\boldsymbol{b}$ to reflect the different values used (Kiko et al., 2017, Kriest, 2002, Alldredge, 1998), where $\boldsymbol{A}$ = 2.8 and $\boldsymbol{b}$ = 2.24, noting that the latter **b** value might be a more realistic size-sinking scaling exponent than $\boldsymbol{b}$ = 3.81 (Cael et al., 2021), because the value of b in the Guidi et al., (2008) work is not the size-sinking speed relationship, but rather the optimized value when compared to observations over a defined size range. We note that we are less interested in the specific values of $\boldsymbol{A}$ and $\boldsymbol{b}$, but rather how the fundamental characteristics of flux's functional form affect its outcome given modeled sampling uncertainties, as in Cael and Bisson, (2018). For the remainder of this paper we use $\boldsymbol{A}$ = 2.8 and $\boldsymbol{b}$ = 2.24 because those values are meant to represent flux more realistically across the range of sizes and particles thought to contribute to flux. We note that the different values of specific choices of A and b tested in our work are those regularly used in the literature, but we do not aim to give a recommendation for a certain formulation. The specific choice of $\boldsymbol{A}$ and $\boldsymbol{b}$ in our work is somewhat arbitrary as we focus on the mathematical form and assumptions of the flux model moreso than the resulting flux numbers and how they may or may not represent flux observations.

One advantage of the modeled PSD in this study is that it can be used to extract the particle number outside the range of observed particle sizes. To quantify the sensitivity of the flux relationship to different sizes, we included bins two sizes smaller than the first size bin observed, as well as two bins larger than the last size bin observed, for each PSD model. Operationally this meant including 80 μm to anywhere from 1 mm to 26 mm (depending on the specific observation) for the size interval. The objective was not to extrapolate widely beyond what has been observed, but rather to include size classes within neighboring bins relative to what was actually seen by the UVP, in order to assess the sensitivity of flux derived from the UVP. Carbon flux calculated using a wider interval for particle sizes was compared to flux calculated from the observed PSD size range. In this study we are not concerned about the

performance of the flux model (as has been done in other studies, Guidi et al 2008, Fender et al., 2019). We instead ask, 'how does using an extended PSD affect flux calculations?'

4 Assessment and Discussion

4.1 Global α and λ values

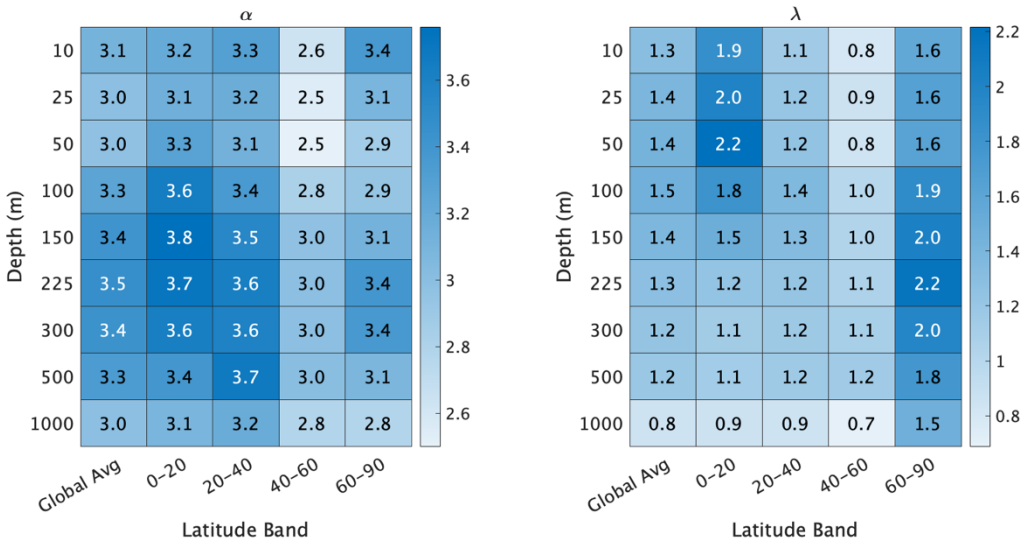

Figure 3. Heatmap of α (left) and λ (right, normalized to λ_0 = 1 mm) based on latitudinal bands and depth.

Global values of retrieved α and λ reveal patterns across space and depth (Figure 3, Supplementary Figures 1,2,8). The highest average α values (3.8) are in moderate depths (100-300 m) for places equatorward of 20 degrees. In general, α varies throughout the water column, with larger values between depths of 150 and 500 m, and lower values at the surface and at 1000

m depth. λ generally decreases with depth, where the global average λ decreases from 1.3 at 10m to 0.8 at 1000 m, and in all places the surface λ value exceeds λ at 1000 m, if only slightly.

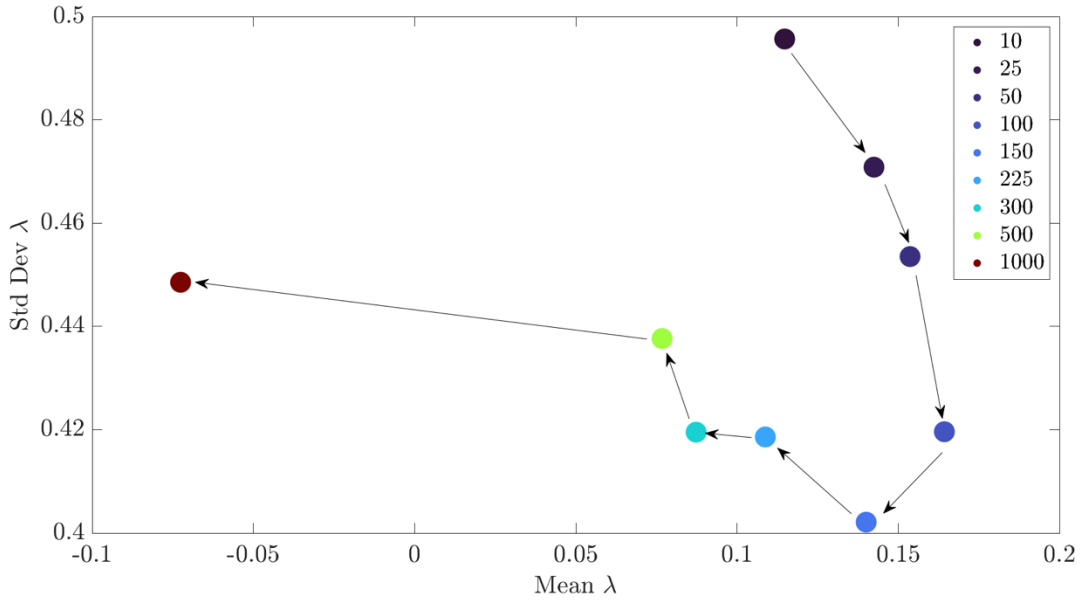

Figure 4. Standard deviation in log10 (λ/λ_0) plotted against the mean log10-transformed λ/λ_0 (where $\lambda_0 = 1$ mm) based on depth (colors). Arrows denote the transition from shallow (top right) to deep (top left) samples.

The standard deviation of λ is highest for the surface ocean and deepest observations at 1000m (Figure 4). In between the surface and depth, λ standard deviation and λ mean (all log10 transformed) have a qualitative clockwise trend (Figure 4) where the average λ changes only subtly at depths < 1000 m while the standard deviation decreases from ~$10^{0.5}$ to ~ $10^{0.4}$. The variability (given by the standard deviation) in λ decreases with depth up to 150 m but thereafter increases to a smaller degree. These results do not necessarily mean there are no big particles (or 'dragon kings', Bochdansky et al., 2016) in the deep, but rather the UVP5 is not observing them.

The model parameters α and λ from the truncated power law fit to observed PSD reflect the relative dominance of small versus large particles, and also indicate the heavy (or not) tail-ness of the size distribution. In essence, α mostly quantifies the mid-range behavior in the PSD and λ mostly quantifies the upper-range behavior. Although the model is statistical in nature, quantifying the PSD slope and size interval where a power law is applicable gives more information about the shape of the PSD than slope alone. In other words, conventional power

law fits to PSD assume that a power law is appropriate over the entire size distribution, and the 1 parameter power law model may not be ideal for characterizing the PSD shape from the UVP.

Lower values of α indicate a higher contribution of large particles relative to small ones, and lower λ values indicate that the power law breaks down at smaller particle sizes (and therefore we expect very few larger particles in the PSD compared to higher λ). It follows, then, that places with shifts in α or λ indicate shifts in the shape of the PSD that may be biogeochemically important. Without coincident observations of particle composition, it is not sensible to say whether or not changes in the PSD shape may specifically be due to e.g., aggregation/disaggregation, ingestion/egestion and vertical transport of zooplankton, bacterial remineralization processes, and so on. However, the clear decrease in global average λ with depth implies that there are fewer large particles at deeper depths in the ocean on average (as observed by the UVP). We note that the particle module on Ecotaxa does not discriminate living from non-living particles, so it is possible that changes in λ will scale with changes in zooplankton abundance and size.

Trends in α are less straightforward. In nearly every latitudinal band, α increases at moderate depths, indicating a higher prevalence of small particles, then decreases at deeper depths. The reported α and λ values here may be useful in future studies to guide improvements to the PSD-derived flux relationship. More work is needed to investigate how the shapes of the PSD (including statistics for the observed PSD's tail as described here) influence carbon flux. For example, can variations in λ values across depth/season/place be used to predict aggregation/disaggregation, or the sinking of fecal pellets? How might variations in α and/or λ along isopycnals (or depth) inform improved parameterizations for the PSD-derived carbon flux model?

4.2. Extrapolation and sampling uncertainties

There are particles that contribute to flux that are not captured by the UVP's sampling volume and specifications. Under what conditions or assumptions are the observed particles sufficiently representative of the total particle population's flux? Figure 5 shows how particles outside the UVP5-observed size range contribute to total flux, for a truncated power-law particle size distribution and a power-law size-flux relationship. If b is the exponent dictating how sinking and mass (or carbon or other elemental content) together scale with particle size, and α is

the exponent dictating how abundance scales with particle size within the power-law scaling range, the contribution to flux by particles of a given size will be determined by their difference, $b - \alpha$. The contribution of large particles will also be determined by λ, the particle size where the power law is truncated.

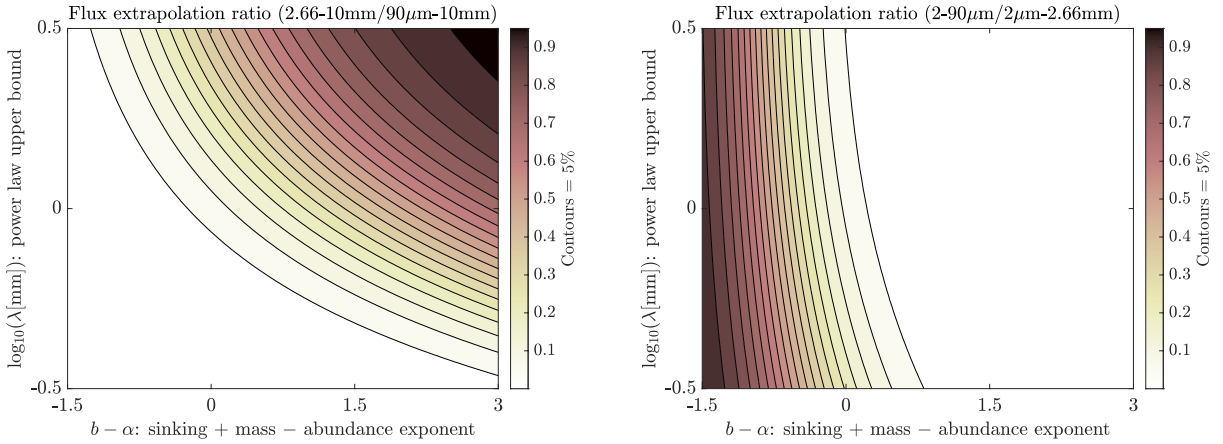

Figure 5. Theoretical flux extrapolation ratio as a function of the difference between b (sinking + mass, = 3.81 or 2.24) and alpha, and the upper bound particle size where a power law is appropriate.

Figure 5a shows the fraction of flux in the 90 μm-10 mm particle ESD range for which 2.66-10 mm particles are responsible, as a function of $b - \alpha$ and λ. We note that this is a conceptual figure that serves only to illustrate under what conditions particle fluxes can be dominated by particles of different sizes. Clearly both parameters play a role; when λ and $b - \alpha$ are both small, meaning large particles are rare and particles' sinking-and-mass size-dependence is weaker than particles size-abundance relationship, large particles contribute very little to total flux so almost none of the flux occurs in the 2.66-10 mm size range. When either λ or $b - \alpha$ are large, however, particles in this range do contribute appreciably to overall flux. When both λ and $b - \alpha$ are large, meaning the power-law distribution extends out to multi-millimeter particles and the sinking-and-mass size dependence of particles is strong relative to particles' size-abundance relationship, most of the flux actually can occur in this 2.66-10 mm size range. In contrast, Figure 5b shows the same but for small particles, comparing 2-90 μm particles against 2 μm-2.66 mm particles. In this case the dependence on λ is unsurprisingly very weak, but we do see that as long as approximately $b - \alpha$ <0, *i.e.* that particle abundance scales more strongly with size than particles' sinking and mass, much or even most of the flux occurs in particles <90 μm.

Although the UVP does not measure particles smaller than 90 μm, these figures underscore that accurate UVP-based flux estimates require understanding the controls on and variability of particles' sinking-size and mass-size relationships, the prevalence of large particles, and the slope of the particle size distribution. We include them to demonstrate that sampling uncertainty includes uncertainty due to particles outside the detection limit of the UVP or any PSD-resolving instrument.

There is high variability in retrieved α, λ, and carbon flux arising from the sample volume uncertainty using the observed size range from UVP observations, as calculated from 100 simulations with varying $\boldsymbol{N(d_i)}$ (informed by the observed $\boldsymbol{N(d_i)}$and sampling volume) for all locations in this study at 50 and 300 m (Figure 6, Supplementary Figure 3,7, see also section 2.2 for procedure). Across all three variates, the coefficient of variation is smallest for α at either depth than it is for λ and carbon flux. The median coefficient of variation for α is ~25% at both depths, while the median coefficient of variation for λ is nearly 60% for the 50 m case, and 55% for the 300 m case. The coefficient of variation for carbon flux arising from the sampling volume uncertainty is highest for deep particles (~65-70%) compared to the 50 m case (50-55%). The width of the coefficient of variation distributions varies for all three variates as well, with α showing the tightest range, followed by λ and carbon flux. We emphasize that the coefficient of variations reported here are due only to measurement error and not due to natural variability, which we could not fully characterize due to lack of repeat data (see Supplementary Figure 4). As a test of how larger sampling volumes may influence the coefficient of variation in α, λ, and carbon flux, we also ran the bootstrapping procedure using simulated sampling volumes that are double the observed sampling volume (but otherwise have the same $\boldsymbol{N(d_i)}$ used in the first simulation case). Doubling the sample volume reduces the coefficient of variation in carbon flux to a median of 56% compared to a median of 67% in the 300 m case (Supplementary Figure 5). Note that we did not adjust the $\boldsymbol{N(d_i)}$ (doing so would preserve the particle concentration) so that we could isolate the relative effect of enhancing sampling volume in a statistical sense.

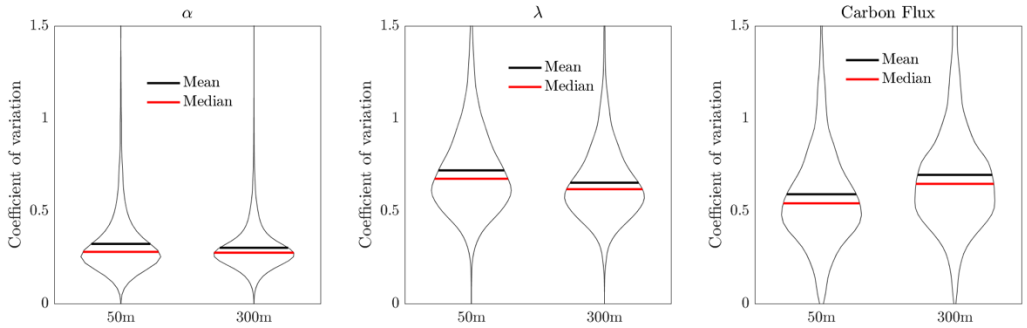

Figure 6. Violin plots for coefficient of variation in α (left), λ (middle) and carbon flux based on either 50 m or 300 m (resulting from the bootstrap procedure). The coefficient of variation reported in this figure is due only to the sampling volume uncertainty.

4.3 Sensitivity of modeled carbon flux to particle size

Carbon flux calculations using observed and modeled PSD (over a shared size range, i.e., the fixed lower limit of 128 μm for each profile and an upper limit dictated by the largest observed particle size) are well correlated as expected, agreeing within 10% for the majority of locations and depths (Figure 7A and Figure 7B). The flux relationship is highly sensitive to the inclusion of smaller particle sizes, which are not thought to contribute substantially to sinking carbon flux. In some cases, the ratio of flux calculated using two size bins smaller than observed to the flux calculated using only observed particle sizes is nearly 6. The high degree of sensitivity to small particles directly scales with the α value of the modeled PSD. For small α (indicating a dominance of larger particles relative to small), the inclusion of smaller size classes makes relatively little difference in the flux calculations (Figure 7C). However, for α larger than 2, the sensitivity of flux to smaller size classes is substantial, with relative differences exceeding

a factor of 2.

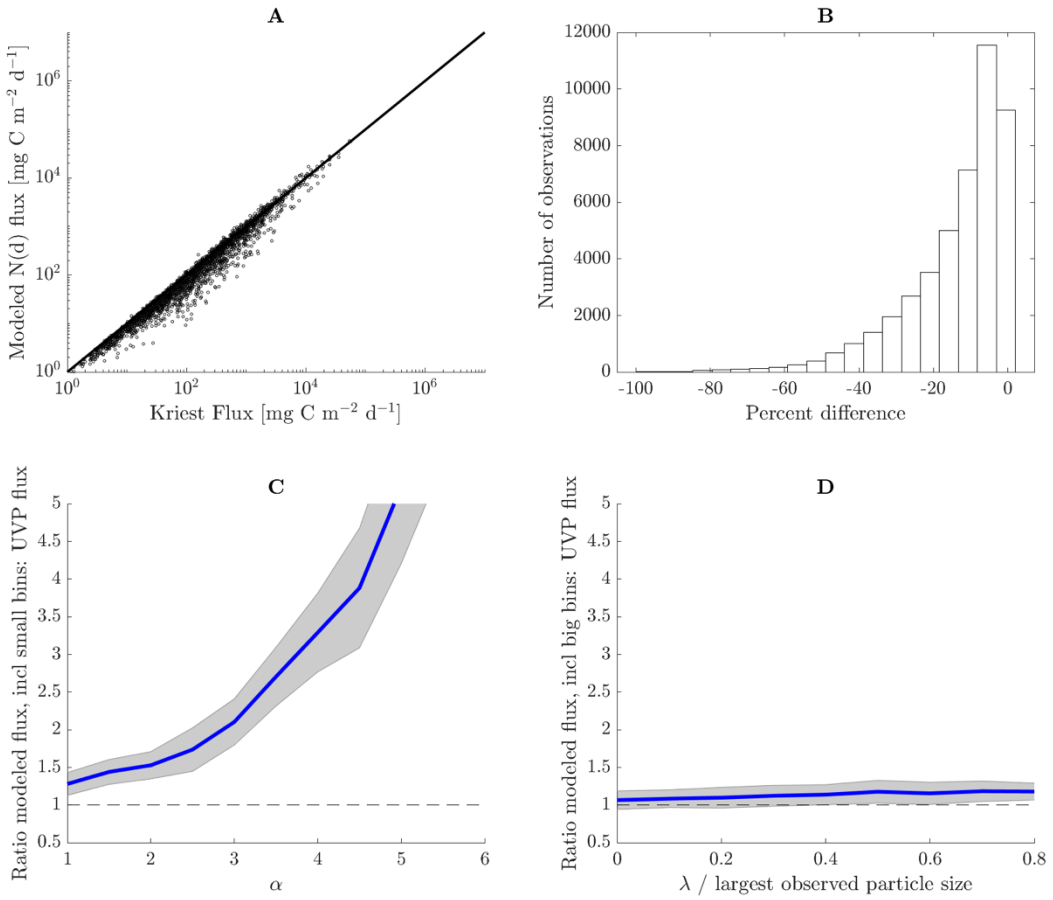

Figure 7. A. Flux comparisons between modeled $\boldsymbol{N(d_i)}$flux (y axis) and $\boldsymbol{N(d_i)}$ flux from UVP observations ('Kriest flux') over the same particle size range. B. Histogram of percent difference (relative to 'Kriest flux') between all flux determinations in A. C. The median ratio of modeled $\boldsymbol{N(d_i)}$ flux (including 2 size bins smaller than the observed size range) to observed $\boldsymbol{N(d_i)}$ flux as a function of depth (colors dots) and alpha value. Shaded area represents the interquartile range. Note this trend is an average of all depths considered in this study, and each specific depth exhibits the same trend. Black line represents equivalent fluxes. D. The median ratio of modeled $\boldsymbol{N(d_i)}$ flux (including 2 size bins larger than the observed size range) to observed $\boldsymbol{N(d_i)}$ flux as a function of depth (colored dots) and lambda value relative to largest observed particle size. Black line represents equivalent fluxes, and shaded area represents the interquartile range. Note this trend is an average of all depths considered in this study, and each specific depth exhibits the same trend.

On the flip side, the flux relationship with PSD is much less sensitive to the inclusion of larger size bins relative to what was observed (Figure 7D). Recall that the modeled PSD in this scenario contains two size bins larger than what was observed. It is common for only a single

particle (or none at all) to be observed at the largest observable size class, so modeled concentrations are accordingly low at the high end of the particle size spectrum. The few large particles in bigger bin sizes stands in contrast to smaller sized particles, which become more numerous as size decreases. We choose to compare the two flux scenarios across the ratio of λ relative to the largest observed particle size (the x-axis in Figure 7D). When this ratio is small, the modeled PSD breaks down from a power law into an exponential decay function at lower size classes. When this ratio is 1, the entire PSD can be modeled using a power law. The relative difference of flux calculated using modeled PSD for larger size classes, to flux calculated using the observed PSD, is around 50% and there is no obvious relationship with the λ value. Although a change of flux by 50% is non-trivial, it is modest compared to changes in flux exceeding 5-fold, as is the case when including smaller particle sizes.

The sensitivity of PSD-derived flux to either small or large particles is robust to changes in the free parameter values of $\boldsymbol{A}$ and $\boldsymbol{b}$. When the Guidi et al., 2008 formulation was applied (Supplementary Figure 6), a similar sensitivity was observed, although the overall magnitude of flux was much enhanced compared to the Kriest et al., (2002) formulation. Ideally, any biogeochemical model will not be sensitive to the inclusion of either small (< 128 μm) or large size classes (> 1 mm) because small particles are not thought to contribute the bulk of carbon flux (see Michaels and Silver, 1988; Bopp et al., 2005), and because large particles are rarely observed and highly uncertain (and so any model relying on large particles for flux would be highly uncertain as well). Further, not all large particles can be assumed to sink at low Reynolds numbers. In this study we found that no matter which flux parameter values are used (i.e., the empirically derived Guidi et al., 2008 values or the more mechanistic Kriest et al., 2002 values) the power law flux relationship is still highly sensitive to the inclusion of small particles. This sensitivity is a surprising result, and likely arises because the high abundance of small particles overcomes their relatively small diameter (in this case, 80 μm) to contribute a large amount of flux (up to 6 times the amount of carbon flux calculated using a minimum particle size of 128 μm). If we included even smaller size classes we expect the flux to increase further. The idea that small particles can contribute substantial flux stands contrary to what is expected from observations of sinking particles in a natural setting (Cael et al., 2021 and refs therein). However, some flux models also predict a larger contribution of small sinking flux (e.g., Bisson et al., 2020, Siegel et al., 2014) than is expected (Cael and White, 2020, Cael et al., 2021). Previous

observations have recorded a substantial contribution (> 30%) of small (< 64 μm) cells to total flux (Durkin et al., 2015). We note that the small particles are not thought to contribute substantially to flux when it is assumed that these small particles are formed in the surface, because they are either suspended and not sinking, or they will be remineralized in their 100s of meters transit to depth. However, small particles may actually dominate flux in deeper waters through disaggregation processes (Kiko et al., 2017, Bianchi et al., 2018). Ultimately, whether or not the empirically derived or mechanistically-informed parameters are used, the sensitivity of UVP-derived flux to small particles underscores the statistical nature of the model, which is biased toward small particles.

In this study, we found that the flux relationship is moderately sensitive to the inclusion of larger particles. One reason for this is because the inclusion of larger bin sizes did not introduce many more particles within this size range, since modeled $\boldsymbol{N(d_i)}$ is often low (if not zero) for large particles. In the real ocean, rare large (> 1 mm) particles can contribute a substantial amount of flux (Bochdansky et al., 2016), but these particles may be missed by the UVP due to sampling volume limitations. We recommend accounting for uncertainty in larger particles based on the sampling volume.

4.4 Limitations of using the UVP to assess particle flux

The primary uncertainties associated with carbon flux derived from UVP observations are 1) assuming the parameters $\boldsymbol{A}$ and $\boldsymbol{b}$ in equation 4 are globally valid at all depths, for all UVP models, and across all size classes 2) the UVP's pixel-to-size uncertainty, 3) error associated with particle detection due to image contrast and porous aggregates (that may appear as many small particles separated by holes) 4) the sampling uncertainty of the PSD, 5) the size to sinking rate uncertainty, and 6) the size to carbon mass uncertainty. We note that the error associated with #3 is likely to be the smallest of all errors presented because the UVP is built to detect near transparent particles in water. In this study we focused on quantifying the sampling uncertainty of the PSD (#4 as described above), as well as how this uncertainty propagates into a commonly used carbon flux model. Ultimately, the sensitivity of the flux relationship to smaller particles

494 was a surprising result of this study, and it invites a re-examination to the flux model in order to
495 guide future work.
496 While some attention has been paid to optimizing the parameters ($\boldsymbol{A}, \boldsymbol{b}$) of the carbon flux
497 model (Guidi et al., 2008, Fender et al., 2019), it seems a larger problem is in the foundation of
498 the flux model itself. The power law formulation for carbon content assumes that all particles of
499 a given size have the same carbon content and sinking speed, which is a flawed assumption
500 given current understanding of particle characteristics. We calculated carbon fluxes
501 incorporating the reported Guidi et al., 2008 parameter standard deviations to learn how carbon
502 flux is uncertain based on parameter value uncertainty. Uncertainty in $\boldsymbol{A}$ resulted in a median
503 21% relative error in carbon flux while uncertainty in $\boldsymbol{b}$ resulted in median 19% relative error
504 using the profiles in this study at all depths. Using the Guidi parameters instead of the Kriest
505 parameters results in median differences approaching a factor of 2 (compared to factors of 6
506 when incorporating 80-128 μm particles). We also note that the Kriest parameters assume larger
507 particles (mm sized) sink at low Reynolds numbers. Approaching the UVP flux model from a
508 mechanistic perspective can be dangerous, as the current model is inherently an empirically
509 derived statistical model inspired by mechanistic ideas (i.e., Stokes' law) that may not be valid in
510 all situations. On the other hand, treating the UVP as a statistical model requires tuning the
511 model to a quality set of observations relevant for model application.
512 Given the large uncertainty also associated with sampling volume, we recommend
513 optimizing the flux model using the same UVP version, or by accounting for uncertainty directly
514 in the model optimizations (e.g., Bisson et al., 2018). We note that the Guidi et al., 2008 study
515 used a size interval of 250 μm to 1.5 mm in order to incorporate older UVP versions (with
516 sampling volumes ranging from 0.28 – 10.5 L) and did not use UVP-5 data. However, although
517 the Guidi et al., 2008 parameters were optimized over a different size interval than was used
518 here, the specific values of $\boldsymbol{A}$ and $\boldsymbol{b}$ will not modify the sensitivity of flux to small size classes
519 (compare Figure 7 with Supplementary Figure 6). Normally, $\boldsymbol{A}$ and $\boldsymbol{b}$ values are optimized
520 within the boundaries of the size spectrum imaged by the UVP, and therefore any regionally
521 optimized PSD-flux relationship is not necessarily problematic to use, even though the
522 theoretical underpinnings of such a relationship are imperfect.
523 It can be instructive to think of the flux model, as a transfer function ($\boldsymbol{Ad^b}$) that is
524 multiplied by $\boldsymbol{N(d_i)}$. The transfer function is a monotonic power law that grows substantially at

larger particle sizes. Therefore, if anything, the flux model is expected to be sensitive to rare instances of large particles, depending on $\boldsymbol{b}$ versus α. If particles in the ocean grew indefinitely, infinite flux would be expected from this relationship. On the other hand, infinite flux is possible with smaller particle sizes if the concentration of particles grows more than particle size decreases, as was the case in this study. If the true PSD were not monotonic (i.e., increasingly higher concentrations of particles at lower particle sizes), a monotonic flux model (such as the power law used here) may be sufficient. If one were to approach the flux model mechanistically, what instead might be needed is a transfer function that quantifies the probability of a given particle size to sink. Small sized particles would accordingly have low probability, as would larger particles that are ultimately living zooplankton (or fish in the extreme case). Medium to medium-large sized particles would have moderate to high probability of becoming carbon flux, which might yield a more realistic carbon flux model. More work is needed to improve the PSD-derived carbon flux relationship, and especially the size to sinking carbon uncertainty, which is outside the scope of this paper.

Finally, particles sized by the UVP include living and non-living particles, which adds uncertainty to flux calculations derived from PSD alone. If only non-living particles were assembled for use by modelers and the rest of the community, the uncertainty associated with ambiguity of large particles (*i.e.,* is it a zooplankter or aggregate?) would be reduced. Indeed, one study (Kiko et al., 2020) found reduced variability in PSD-derived carbon flux, during which living objects and artefacts with an equivalent spherical diameter larger than 1 mm were removed from the UVP5 image dataset so that only detrital particles were used to calculate flux in this size range.

4.5 Future applications of using a truncated power law to model PSD

In this study we found enhanced performance of the modeled PSD when using a truncated power law rather than a power law. Truncated power laws offer more information about a PSD distribution compared to a power law because the behavior of the distribution is characterized through two main parameters (α, λ) rather than just one (α, in the case of a power law). There are several applications to using a truncated power law besides what has been explored here. First, with an improved model for PSD, one could extrapolate the PSD to quantify the carbon content of particles in the particulate fraction globally. Second, extrapolating the PSD using a truncated power law may enable improved respiration rates as derived from UVP

observations, as current estimates are limited by the size resolved by the UVP (Kalvelage et al., 2015, Thomsen et al., 2019). Third, the current UVP data hosted by Ecotaxa includes both living and non-living particles. Future work may explore whether or not the λ values will be useful to identify when the PSD spectrum transitions from particles to larger zooplankton (Forest et al., 2012).

5 Recommendations for future work

Although we chose to focus on sampling uncertainties and how they influence carbon flux values, there are outstanding issues with the assumed size to sinking rate uncertainty, and size to carbon mass uncertainty. These uncertainties may be reduced in future work by using existing information from UVP images. Below, we mention a few possible avenues to address uncertainty associated with the UVP carbon flux model.

1. *Sampling uncertainty:* Future UVP designs can reduce sampling uncertainty by increasing the sampling volume of the instrument. Current UVP designs can reduce sampling uncertainty to some extent by performing multiple casts of repeat sampling. More work needs to be done in order to distinguish aggregates from living plankton for particles in the observable size range, preferentially down to 2 μm size.

2. *Size to sinking uncertainty*: Although unconventional, a UVP fastened to a Lagrangian sediment and/or gel trap that is oriented with a side-viewing camera may allow sinking speed to be assessed via several images, where sediment trap flux, particle sinking speeds from a gel trap, and particle size information would be coupled and coincident. Similarly, in situ sinking speeds could be obtained using Particle Imaging Velocimetry (Cartwright et al. 2013), optimally during the upcast of a CTD/UVP profile. This logistically less demanding approach could yield PSD observations over the entire water column and coincident particle sinking speed observations at different water depths. Targeting blooms of different organisms with UVP observations may also help to improve size-sinking relationships. Direct observations that better constrain the size-sinking scaling

583 relationship globally, in different environments, and/or for different particle types is
584 essential for improving uncertainties in UVP-derived fluxes. (Cael et al., 2021).

585 3. *Size to carbon content uncertainty:* Dense particles may have a different reflectance than
586 less dense particles (based on the fractal dimension) which might provide a way to semi-
587 quantitatively assess particle composition from the contrast of the images. If such an
588 exercise is possible, the contrast of images may add information content to the flux
589 relationship so particle size and concentration are not the only variables. Further
590 classification of particle images into e.g. fecal pellets, marine snow and other types of
591 detrital matter and the application of class specific size to carbon ratios might also reduce
592 the errors in flux calculation (Durkin et al., 2021). Finally, due to remineralization,
593 carbon content might also decrease over depth without large changes in size or
594 appearance of the particles. Therefore, further work is needed to characterize the carbon
595 to size relationship of detrital particles at different depths, especially if the UVP flux
596 model is used in a mechanistic framework.

597 Although we did not investigate the uncertainty associated with the conversion of UVP pixels to
598 a particle size, more work is needed to characterize any error and uncertainty arising from
599 particle shape differences and assumed spherical diameters. Improved edge detection of pixels is
600 needed, as well as a sensitivity analysis of how threshold values for edge detection affect particle
601 size (as is also advised by Giering et al., 2020).

602 **6 Summary**

603 In this study we focused on UVP sampling uncertainties and how they propagate into
604 derived estimates of carbon flux. The PSD observations from the UVP5 have a sampling
605 uncertainty $\sqrt{\boldsymbol{N}}/\boldsymbol{V}$. Ultimately, flux calculated from the UVP is highly sensitive to small
606 particles. Gel trap imaging also shows that carbon flux in the 10 to 64 μm range actually
607 contributes substantially to total flux (Durkin et al., 2015). Combined deployments of gel traps
608 and the UVP should help to further resolve the relative importance of different size fractions for
609 carbon flux. The sampling uncertainty of PSD observations results in an uncertainty slightly
610 greater than 50% for carbon flux. The extrapolated carbon flux from the UVP is based on a
611 relationship that is highly sensitive (up to 4-fold different on average) to the inclusion of

612 particles slightly smaller than what was observed. We advocate for a revised carbon flux
613 relationship that is possibly non-monotonic and considers the probability of a given particle to
614 enter the sinking pool and contribute to carbon flux. In the absence of an improved carbon flux
615 relationship, carbon flux calculations should be made using parameters specific to a particular
616 region and depth to prevent large errors. Future work may benefit by using UVP data in
617 unconventional ways, such as coupling a UVP and sediment trap in the same water mass, and/or
618 by performing image analysis on the particular pixels comprising a particle.

619 **Acknowledgments**

620 Cael was supported by the National Environmental Research Council (NE/R015953/1) and the
621 Horizon 2020 Framework Programme (820989, project COMFORT, our common future ocean
622 in the Earth system—quantifying coupled cycles of carbon, oxygen, and nutrients for
623 determining and achieving safe operating spaces with respect to tipping points). The work
624 reflects only the authors' view; the European Commission and their executive agency are not
625 responsible for any use that may be made of the information the work contains.
626 RK acknowledges support via a "Make Our Planet Great Again" grant of the French National
627 Research Agency within the "Programme d'Investissements d'Avenir"; reference "ANR-19-
628 MPGA-0012", the BMBF funded project "CUSCO" and funding from the European Union's
629 Horizon 2020 research and innovation programme for the TRIATLAS project under
630 grant agreement No 817578. We are grateful to our reviewers, who provided comments and
631 criticisms that improved the clarity of our manuscript.

632 **Data Availability Statement**

633 The UVP data used in this study can be accessed via PANGAEA,
634 https://doi.pangaea.de/10.1594/PANGAEA.924375. Log ins are required for downloading, and
635 the data span 2008-2020. More processing details are in Kiko et al., 2021.

636 **Author Contribution Statement**

637 KB, DS, and BBC conceived the study. KB led the analysis with input from BBC and all
638 authors. KB led the writing with input from all authors.

639

640 **References**

641 Alldredge, A. L., & Gotschalk, C. (1988). In situ settling behavior of marine snow 1. *Limnology*
642 *and Oceanography*, *33*(3), 339-351.
643
644 Alldredge, A. (1998). The carbon, nitrogen and mass content of marine snow as a function of
645 aggregate size. *Deep Sea Research Part I: Oceanographic Research Papers*, *45*(4-5), 529-541.
646
647 Bianchi, D., Weber, T. S., Kiko, R., & Deutsch, C. (2018). Global niche of marine anaerobic
648 metabolisms expanded by particle microenvironments. *Nature Geoscience*, *11*(4), 263-268.
649
650 Bisson, K. M., Siegel, D. A., DeVries, T., Cael, B. B., & Buesseler, K. O. (2018). How data set
651 characteristics influence ocean carbon export models. *Global Biogeochemical Cycles*, *32*(9),
652 1312-1328.
653
654 Bisson, K., Siegel, D. A., & DeVries, T. (2020). Diagnosing mechanisms of ocean carbon export
655 in a satellite-based food web model. *Frontiers in Marine Science*, *7*, 505.
656
657 Bochdansky, A. B., Clouse, M. A., & Herndl, G. J. (2016). Dragon kings of the deep sea: marine
658 particles deviate markedly from the common number-size spectrum. *Scientific reports*, *6*(1), 1-7.
659
660 Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms
661 distribution to global warming and potential implications: A global model study. *Geophysical*
662 *Research Letters*, *32*(19).
663
664 Boss, E., Twardowski, M. S., & Herring, S. (2001). Shape of the particulate beam attenuation
665 spectrum and its inversion to obtain the shape of the particulate size distribution. *Applied*
666 *Optics*, *40*(27), 4885-4893.
667
668 Buonassissi, C. J., & Dierssen, H. M. (2010). A regional comparison of particle size distributions
669 and the power law approximation in oceanic and estuarine surface waters. *Journal of*
670 *Geophysical Research: Oceans*, *115*(C10).

Cael, B. B., Cavan, E. L., & Britten, G. L. (2021). Reconciling the size-dependence of marine particle sinking speed. *Geophysical Research Letters*, *48*(5), e2020GL091771.

Cael, B. B., & White, A. E. (2020). Sinking versus suspended particle size distributions in the North Pacific Subtropical Gyre. *Geophysical Research Letters*, *47*(15), e2020GL087825.

Cael, B. B., & Bisson, K. (2018). Particle flux parameterizations: Quantitative and mechanistic similarities and differences. *Frontiers in Marine Science*, *5*, 395.

Cartwright, G. M., Friedrichs, C. T., & Smith, S. J. (2013). A test of the ADV-based Reynolds flux method for in situ estimation of sediment settling velocity in a muddy estuary. *Geo-Marine Letters*, *33*(6), 477-484.

Chase, A. P., Kramer, S. J., Haëntjens, N., Boss, E. S., Karp-Boss, L., Edmondson, M., & Graff, J. R. (2020). Evaluation of diagnostic pigments to estimate phytoplankton size classes. *Limnology and Oceanography: Methods*, *18*(10), 570-584.

Dall'Olmo, G., Westberry, T. K., Behrenfeld, M. J., Boss, E., & Slade, W. H. (2009). Significant contribution of large particles to optical backscattering in the open ocean. *Biogeosciences*, *6*(6), 947.

Diehl, P., & Haardt, H. (1980). Measurement of the spectral attenuation to support biological-research in a plankton tube experiment. *Oceanologica Acta*, *3*(1), 89-96.

Durkin, C. A., Buesseler, K. O., Cetinić, I., Estapa, M. L., Kelly, R. P., & Omand, M. (2021). A visual tour of carbon export by sinking particles. *bioRxiv*.

Durkin, C. A., Estapa, M. L., & Buesseler, K. O. (2015). Observations of carbon export by small sinking particles in the upper mesopelagic. *Marine Chemistry*, *175*, 72-81.

Fender, C. K., Kelly, T. B., Guidi, L., Ohman, M. D., Smith, M. C., & Stukel, M. R. (2019). Investigating particle size-flux relationships and the biological pump across a range of plankton ecosystem states from coastal to oligotrophic. *Frontiers in Marine Science*, *6*, 603.

Fink, D. (1997). A compendium of conjugate priors. *See http://www. people. cornell. edu/pages/df36/CONJINTRnew% 20TEX. pdf*, *46*.

Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., & Babin, M. (2012). Size distribution of particles and zooplankton across the shelf-basin system in southeast Beaufort Sea: combined results from an Underwater Vision Profiler and vertical net tows. *Biogeosciences*, *9*(4), 1301-1320.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). Bayesian Data Analysis, volume third ed.

Giering, S. L., Cavan, E. L., Basedow, S. L., Briggs, N., Burd, A. B., Darroch, L. J., ... & Lindsay, D. J. (2020). Sinking organic particles in the ocean—flux estimates from in situ optical devices.

Guidi, L., Jackson, G. A., Stemmann, L., Miquel, J. C., Picheral, M., & Gorsky, G. (2008). Relationship between particle size distribution and flux in the mesopelagic zone. *Deep Sea Research Part I: Oceanographic Research Papers*, *55*(10), 1364-1374.

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., ... & Gorskya, G. (2009). Effects of phytoplankton community on production, size, and export of large aggregates: A world-ocean analysis. *Limnology and Oceanography*, *54*(6), 1951-1963.

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., ... & Coelho, L. P. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, *532*(7600), 465-470.

Gorsky, G., Picheral, M., & Stemmann, L. (2000). Use of the Underwater Video Profiler for the study of aggregate dynamics in the North Mediterranean. *Estuarine, Coastal and Shelf Science*, *50*(1), 121-128.

Haight, Frank A. *Handbook of the Poisson distribution*. No. 519.23 H3. 1967.

Jonasz, M., & Fournier, G. (2011). *Light scattering by particles in water: theoretical and experimental foundations*. Elsevier.

Karp-Boss, L., Azevedo, L., & Boss, E. (2007). LISST-100 measurements of phytoplankton size distribution: Evaluation of the effects of cell shape. *Limnology and Oceanography: Methods*, *5*(11), 396-406.

Kalvelage, T., Lavik, G., Jensen, M. M., Revsbech, N. P., Löscher, C., Schunck, H., ... & Kuypers, M. M. (2015). Aerobic microbial respiration in oceanic oxygen minimum zones. *PloS one*, *10*(7), e0133526.

Kiko, R., Brandt, P., Christiansen, S., Faustmann, J., Kriest, I., Rodrigues, E., ... & Hauss, H. (2020). Zooplankton-mediated fluxes in the eastern tropical North Atlantic. *Frontiers in Marine Science*, 358.

Kiko, Rainer; Picheral, Marc; Antoine, David; Babin, Marcel; Berline, Leo; Biard, Tristan; Boss, Emmanuel; Brandt, Peter; Carlotti, F; Christiansen, Svenja; Coppola, Laurent; de la Cruz, Leandro; Diamond-Riquier, Emilie; de Madron, Xavier Durrieu; Elineau, Amanda; Gorsky, Gabriel; Guidi, Lionel; Hauss, Helena; Irisson, Jean-Olivier; Karp-Boss, Lee; Karstensen, Johannes; Kim, Dong-gyun; Lekanoff, Rachel M; Lombard, Fabien; Lopes, Rubens M; Marec, Claudie; McDonnell, Andrew; Niemeyer, Daniela; Noyon, Margaux; O'Daly, Stephanie; Ohman, Mark; Pretty, Jessica L; Rogge, Andreas; Searson, Sarah; Shibata, Masashi; Tanaka, Yuji; Tanhua, Toste; Taucher, Jan; Trudnowska, Emilia; Turner, Jessie S; Waite, Anya M; Stemmann, Lars (2021): The global marine particle size distribution dataset obtained with the Underwater Vision Profiler 5 - version 1. PANGAEA https://doi.pangaea.de/10.1594/PANGAEA.924375

765 Kiko, R., Biastoch, A., Brandt, P., Cravatte, S., Hauss, H., Hummels, R., ... & Stemmann, L.
766 (2017). Biological and physical influences on marine snowfall at the equator. *Nature*
767 *Geoscience*, *10*(11), 852-858.
768
769 Kriest, I. (2002). Different parameterizations of marine snow in a 1D-model and their influence
770 on representation of marine snow, nitrogen budget and sedimentation. *Deep Sea Research Part I:*
771 *Oceanographic Research Papers*, *49*(12), 2133-2162.
772
773 Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., ... & Appeltans, W.
774 (2019). Globally consistent quantitative observations of planktonic ecosystems. *Frontiers in*
775 *Marine Science*, *6*, 196.
776
777 Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). The BUGS Book: A
778 Practical Introduction to Bayesian Analysis (CRC, Boca Raton, FL).
779
780 Michaels, A. F., & Silver, M. W. (1988). Primary production, sinking fluxes and the microbial
781 food web. *Deep Sea Research Part A. Oceanographic Research Papers*, *35*(4), 473-490.
782
783 Munk, W. H. (1966, August). Abyssal recipes. In *Deep Sea Research and Oceanographic*
784 *Abstracts* (Vol. 13, No. 4, pp. 707-730). Elsevier.
785
786 Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., & Gorsky, G. (2010). The
787 Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of
788 particle size spectra and zooplankton. *Limnology and Oceanography: Methods*, *8*(9), 462-473.
789
790 Reynolds, R. A., Stramski, D., Wright, V. M., & Woźniak, S. B. (2010). Measurements and
791 characterization of particle size distributions in coastal waters. *Journal of Geophysical Research:*
792 *Oceans*, *115*(C8).
793
794 Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., & Boyd, P. W.
795 (2014). Global assessment of ocean carbon export by combining satellite observations and food-
796 web models. *Global Biogeochemical Cycles*, *28*(3), 181-196.

Sheldon, R. W., Prakash, A., & Sutcliffe Jr, W. (1972). The size distribution of particles in the Ocean 1. *Limnology and oceanography*, *17*(3), 327-340.

Slade, W. H., & Boss, E. (2015). Spectral attenuation and backscattering as indicators of average particle size. *Applied optics*, *54*(24), 7264-7277.

Stemmann, L., & Boss, E. (2012). Plankton and particle size and packaging: from determining optical properties to driving the biological pump. *Annual Review of Marine Science*, *4*, 263-290.

Thomsen, S., Karstensen, J., Kiko, R., Krahmann, G., Dengler, M., & Engel, A. (2019). Remote and local drivers of oxygen and nitrate variability in the shallow oxygen minimum zone off Mauritania in June 2014. *Biogeosciences*, *16*(5), 979-998.

White, A. E., Letelier, R. M., Whitmire, A. L., Barone, B., Bidigare, R. R., Church, M. J., & Karl, D. M. (2015). Phenology of particle size distributions and primary productivity in the North Pacific subtropical gyre (Station ALOHA). *Journal of Geophysical Research: Oceans*, *120*(11), 7381-7399.