



**HAL**  
open science

## Sampling uncertainties of particle size distributions and derived fluxes

Kelsey Bisson, Rainer Kiko, David Siegel, Lionel Guidi, Marc Picheral,  
Emmanuel Boss, B. Cael

► **To cite this version:**

Kelsey Bisson, Rainer Kiko, David Siegel, Lionel Guidi, Marc Picheral, et al.. Sampling uncertainties of particle size distributions and derived fluxes. *Limnology and Oceanography: Methods*, 2022, 10.1002/lom3.10524 . hal-03847300v2

**HAL Id: hal-03847300**

**<https://hal.science/hal-03847300v2>**

Submitted on 24 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sampling uncertainties of particle size distributions and derived fluxes

Kelsey M Bisson<sup>1</sup>, Rainer Kiko<sup>2</sup>, David A. Siegel<sup>3</sup>, Lionel Guidi<sup>4</sup>, Marc Picheral<sup>5</sup>, Emmanuel Boss<sup>6</sup>, and B. B. Cael<sup>7</sup>

<sup>1</sup>Oregon State University

<sup>2</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany

<sup>3</sup>University of California, Santa Barbara

<sup>4</sup>French National Centre for Scientific Research (CNRS)

<sup>5</sup>CNRS/UPMC

<sup>6</sup>School of Marine Sciences, University of Maine, Orono, ME, USA

<sup>7</sup>National Oceanography Centre

November 24, 2022

## Abstract

The Underwater Vision Profiler (UVP) provides abundant in situ data of the marine particle size distribution (PSD) on global scales and has been used for a diversity of applications, but the uncertainty associated with its measurements has not been quantified. Here we use a global compilation of UVP (version 5) observations of the PSD to assess the sampling uncertainty associated with the UVP's sampling characteristics. 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 power law with an exponential cutoff 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 microns). We close with recommendations on how to revise the carbon flux model, and we provide avenues to address additional uncertainties associated with UVP-derived carbon flux calculations.

1           **Sampling uncertainties of particle size distributions and derived fluxes**

2   **K. M. Bisson<sup>1</sup>, R. Kiko<sup>2</sup>, D.A. Siegel<sup>3</sup>, L. Guidi<sup>2</sup>, M. Picheral<sup>2</sup>, Emmanuel Boss<sup>4</sup> and B. B.**  
3   **Cael<sup>5</sup>**

4   <sup>1</sup>Department of Botany and Plant Pathology, Oregon State University, Corvallis, USA

5   <sup>2</sup>Laboratoire d'Océanographie de Villefranche-sur-mer, Sorbonne Université, Villefranche-sur-  
6   mer, France

7   <sup>3</sup>Earth Research Institute and Department of Geography, University of California, Santa  
8   Barbara, Santa Barbara, USA

9   <sup>4</sup>School of Marine Sciences, University of Maine, Orono, Maine 04469, USA

10   <sup>5</sup>National Oceanography Centre, European Way, Southampton, SO143ZH, UK

11

12   Corresponding author: Kelsey Bisson ([bissonk@oregonstate.edu](mailto:bissonk@oregonstate.edu))

13   **Key Points:**

- 14       • We model uncertainty in UVP5-derived PSDs and fluxes via Bayesian Poisson statistics  
15       and a truncated power law distribution.
- 16       • Modeled carbon flux calculations have an uncertainty of ~50% arising from sampling  
17       uncertainty.
- 18       • The carbon flux model is highly sensitive (up to 6-fold differences) to the inclusion of  
19       80-128  $\mu\text{m}$  particles.

**20 Abstract**

21 The Underwater Vision Profiler (UVP) provides abundant in situ data of the marine particle size  
22 distribution (PSD) on global scales and has been used for a diversity of applications, but the  
23 uncertainty associated with its measurements has not been quantified. Here we use a global  
24 compilation of UVP (version 5) observations of the PSD to assess the sampling uncertainty  
25 associated with the UVP's sampling characteristics. We model UVP sampling uncertainty using  
26 Bayesian Poisson statistics and provide formulae for the uncertainty associated with a given  
27 sampling volume and observed particle count. We also model PSD observations using a power  
28 law with an exponential cutoff to better match the low concentration associated with rare large  
29 particles as seen by the UVP. We use the two shape parameters from this statistical model to  
30 describe changes in the PSD shape across latitude band, season, and depth. The UVP sampling  
31 uncertainty propagates into an uncertainty for modeled carbon flux exceeding 50%. The  
32 statistical model is used to extend the size interval used in a PSD-derived carbon flux model,  
33 revealing a high sensitivity of the PSD-derived flux model to the inclusion of small particles (80-  
34 128 microns). We close with recommendations on how to revise the carbon flux model, and we  
35 provide avenues to address additional uncertainties associated with UVP-derived carbon flux  
36 calculations.

**37 Plain Language Summary**

38 The size of a particle in the ocean influences its ecological role. Carbon included in bigger  
39 sinking particles are thought to be removed from the surface ocean and possibly sequestered  
40 from the atmosphere. The Underwater Vision Profiler (UVP) is a camera system that takes  
41 pictures of particles from the ocean's surface to depth. The UVP images a small portion of the  
42 water column (~1L at the highest frequency), and does not often capture rare large particles  
43 thought to be important for carbon storage. We use statistical models to assess the uncertainty in  
44 particle concentrations associated with the UVP, and we calculate the uncertainty of sinking  
45 carbon calculated from UVP observations. We find a formula for UVP sampling uncertainty that  
46 depends on particle counts and sampling volume. The associated sinking carbon rate uncertainty  
47 is ~50%. We also model UVP observations using a statistical model that captures rare, large  
48 particles better than a commonly used power law. We use this updated PSD model to 1) describe  
49 changes in the PSD shape across depth, time, and place, and 2) test how sinking carbon



50 calculations change when a different size range is used. The sinking carbon relationship is very  
51 sensitive to small particles.

## 52 **1 Introduction**

53 In the ocean, an extraordinary range of particle sizes (from  $< 1 \mu\text{m}$  to 30m; including non-  
54 living dust particles, detrital matter, bacteria, phytoplankton, zooplankton including salp chains,  
55 whales and many others) influence ecosystem structure and function, net primary production,  
56 particle sinking, and carbon flux (Sheldon et al., 1972, White et al., 2015, Alldredge and Gotschalk,  
57 1988, Siegel et al., 2014). Over the last decade, bio-optics has enabled the characterization of  
58 portions of the particle size distribution (PSD) (Boss et al., 2001, Slade and Boss, 2015, Dall’Olmo  
59 et al., 2009, Reynolds et al., 2010, Chase et al., 2020, Stemmann and Boss, 2012, Cael and White  
60 2020, Giering et al., 2020 and refs therein), especially through the advancement of in situ imaging  
61 technologies.

62 In order to use PSD observations in the most meaningful way in analyses and models, the  
63 uncertainty associated with the observations must be clearly quantified. In situ observations of the  
64 PSD are a function of both the true size structure of the particle assemblage and of the measurement  
65 method. In this study we focus on PSD data collected from the Underwater Vision Profiler (UVP,  
66 Gorsky et al., 2000, Picheral et al., 2010), which ‘sees’ a narrow size range (60 microns – 20,000  
67 micron capabilities, Lombard et al., 2019) of living and non-living particles which are imaged  
68 within a small fraction of the water column (anywhere from 0.28L to 10.5L depending on the UVP  
69 version, Guidi et al., 2008). The surface area of pixels containing a particle is converted into an  
70 assumed equivalent spherical diameter using instrument specific calibrations (Picheral et al.,  
71 2010), no matter how a particle is shaped or oriented (introducing error into the retrieved particle  
72 size e.g., Karp-Boss et al., 2007).

73 Uncertainties in particle detection are propagated downstream into calculations of carbon  
74 flux and other applications, which rely on both accurate PSD observations as well as appropriate  
75 modeling to convert standing stocks of PSD observations into rates of sinking carbon across the  
76 full range of depths and particle types in the ocean. When PSDs are not used directly for  
77 calculations of flux or other quantities of interest, PSDs are commonly described with a power law  
78 to reflect the rapid decline in particle concentrations with increasing particle size (e.g., Jonasz and  
79 Fournier, 2011). However, the power-law exponent estimation is sensitive to the abundance of rare

80 large particles, and the behavior of power-law distributed quantities (e.g., carbon flux) is sensitive  
81 to the exact values of the power-law exponent. In any natural system, a power law is only  
82 applicable over a finite size range and this size range must be adequately accounted for. We model  
83 PSD observations with a truncated power law rather than a power law to better account for rare  
84 instances of large particles observed by in situ instruments. Moreover, a truncated power law  
85 distribution has an extra parameter about the particle size range for which power-law behavior  
86 holds, which offers more information about the shape of the particle size distribution than a scaling  
87 exponent alone.

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

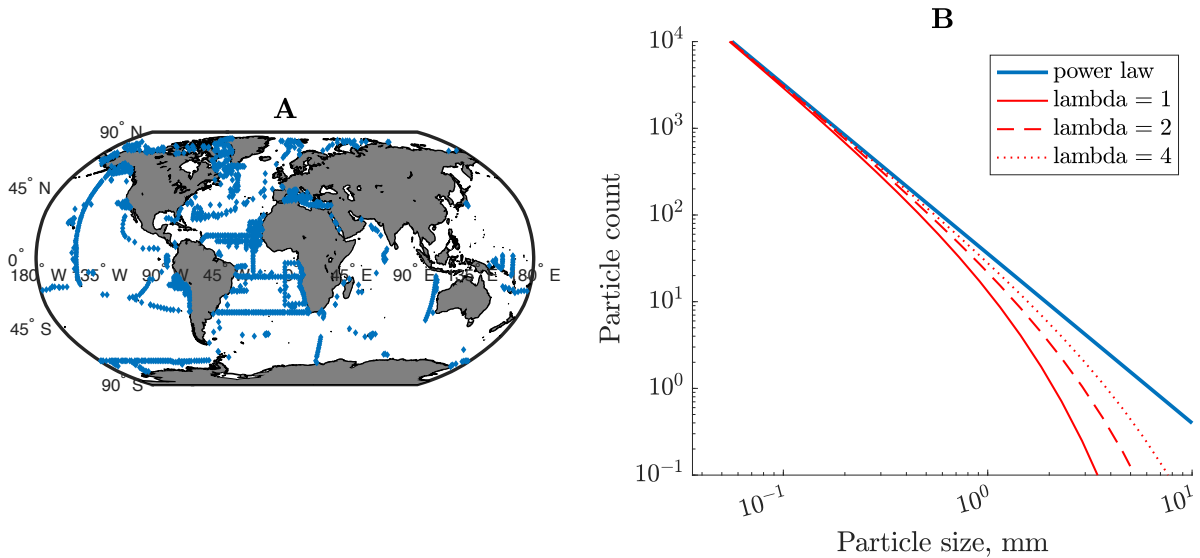
## 94 **2 Materials and Methods**

### 95 **2.1 UVP Data**

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

105 Since its invention, the UVP has undergone design improvements so that its size and  
106 sampling speed are compatible with a standard CTD rosette. The UVP5 (Picheral et al., 2010)  
107 has an image acquisition frequency varying between about 3 to 20 Hz depending on versions and  
108 particle load of the water column (higher loads require more processing time and therefore a  
109 lower acquisition frequency). During normal CTD deployments with speeds up to 1 m/s, this

110 allows imaging of up to 1L/image at the highest frequency of 20Hz and 20L/m on a vertical  
 111 profile at 1m/s. The surface area of particles is converted from pixel counts (using instrument  
 112 settings), and the equivalent spherical diameter (ESD) is calculated following  $aa *$   
 113  $number\_of\_pixels^b$  where  $aa$  and  $b$  were determined through calibration casts in the bay of  
 114 Villefranche. Hereafter any use of the UVP is implied to mean UVP5 in our study.  
 115



116  
 117 Figure 1. A. Location of UVP5 observations (blue). B. Comparison of a power law (blue) with a  
 118 power law with an exponential cutoff of various  $\lambda$  values (red lines). All lines share the same  $\alpha$ .  
 119 The power law is of the form  $N(d) = d^{-\alpha}$  while the truncated power laws follow  $\sim d^{-\alpha} * e^{-\frac{d}{\lambda}}$ .

120

## 121 2.2 Extrapolation and sampling uncertainty calculations

122

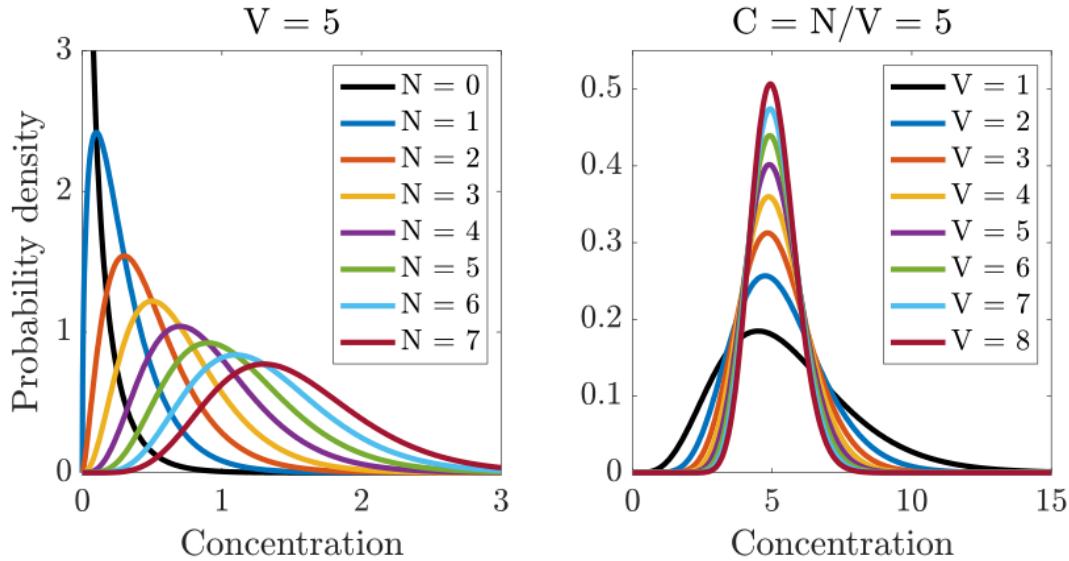
123 A UVP measurement of the PSD is an estimate of the true particle population in the water  
 124 column. If a UVP samples  $N$  particles within a range of sizes with average diameter  $d$  in a  
 125 volume  $V$ , intuitively the best estimate for the concentration of  $d$ -sized particles is  $N/V$ , and the  
 126 larger  $V$  and/or  $N$  the better an estimate this will be — but what is the uncertainty associated  
 127 with this estimate, and how does it depend on  $N$  and  $V$ ? How do these uncertainties ultimately  
 128 propagate into uncertainty in estimated flux? Because particle dis/aggregation is complex, this  
 129 problem is intractable to quantify perfectly, but may be substantially simplified by assuming  
 130 Poisson statistics, i.e. that if the true concentration is  $C$ , and a volume  $V$  is being sampled, the  
 131 likelihood of each particle being sampled is independently  $C \times V$ .

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

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

157 To estimate uncertainty in the fitted  $\alpha$  and  $\lambda$  values from  $N(d) = C * d^{-\alpha} * e^{\frac{-d/d_0}{\lambda/\lambda_0}}$  (where  
 158  $\lambda_0=1\text{mm}$  and  $d_0 = 1 \text{ mm}$ ), and in the modeled carbon fluxes, we thus draw 100 random samples  
 159 from  $\text{NB}(N+1/2, 1/(V+1))$  for each particle size class at each sampled place and time. These  
 160 calculations were run at all places for depths 50 and 300m to retrieve  $\alpha$  and  $\lambda$  and the coefficient

161 of variation of each. We also calculate carbon flux (described in section 2.4) for each of the 100  
 162 simulated PSDs. The coefficient of variation is reported as the standard deviation ( $\sigma$ ) normalized  
 163 by the mean, and relative error ( $error_{rel}$ ) is given by the  $\sigma$  divided by the  $N(d)$ , x 100%.



164

165 Figure 2. Theoretical probability of particle concentration ( $N/V$ ) based on observed particle  
 166 number ( $N$ ) or sampling volume ( $V$ ).  
 167

168 The uncertainty associated with sampling volume is visualized using probability densities for  
 169 particle concentrations for either fixed or variable sampling volumes (Figure 2). When sampling  
 170 volume is fixed (and assumed to be 5L), the width of the probability distribution increases  
 171 substantially as particle count increases for an arbitrary size class (compare maroon line to black  
 172 line, Figure 2, left plot). Essentially, if the observed particle count is 5, the true concentration in  
 173 the water column is likely to be between 0.5 and 2 (green line). For fixed concentrations  
 174 (assumed to be 5 particles per L, Figure 2, right plot) and variable sample volumes, the  
 175 probability that the true concentration of particles is accurately measured by the UVP scales with  
 176 sampling volume. Higher sampling volumes (8L, maroon line, Figure 2, right plot) result in  
 177 narrow probability distributions that give higher fidelity to the observed particle concentration.  
 178 Lower sampling volumes (black line, 1L, Figure 2, right plot) have a wider probability

179 distribution, where it is evident that the true particle concentration can be a factor of 2 (and  
180 greater) different than what was observed.

### 181 **2.3 Modeling the observed PSD**

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

$$185 \quad N(d) = C * d^{-\alpha} * e^{-\frac{d}{\lambda}}. \quad [1]$$

186

187  $N(d)$  is the number of particles within a given size bin and normalized by the bin width,  $d$  is the  
188 equivalent spherical diameter, and  $\alpha$  and  $\lambda$  are free parameters. It is implied that both  $d$  and  $\lambda$  are  
189 normalized by  $d_0 = \lambda_0 = 1\text{mm}$ , everywhere  $d$  and  $\lambda$  are operated on in this text. The leading  
190 constant  $C$  is the concentration at  $d=1$  mm divided by  $e^{\frac{1}{\lambda}}$ . The available sizes for  $d$  range from  
191  $1.0 \times 10^{-3}$  to 26mm, but operationally, the minimum observed particle size from the UVP5 falls  
192 into the 128-161 microns size class. Conceptually,  $\alpha$  is a typical power law scaling exponent and  
193  $\lambda$  is the upper limit until which the particle size distribution is well-described by a power law.  
194 High values of  $\alpha$  are associated with a steep PSD slope, or a particle assemblage dominated by  
195 many small particles relative to larger ones. Low values of  $\lambda$  are associated with a steep decline  
196 in  $N(d)$  earlier in the size spectrum (Figure 1b, solid red line compared to red dashed line).

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

201

$$202 \quad \text{cost} = - \sum_{i=1}^n W_i * [\log_{10}(N(d)) - \log_{10}(PSD_{obs})] \quad [2]$$

203

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

206

$$207 \quad W_i = \frac{V}{\text{error}_{rel}} \quad [3]$$

208

209 The model fitting was performed over the observed particle size interval for each specific  
 210 instance depending on the depth and location of observations. We constrained  $\alpha$  to be between 0  
 211 and 6. The  $\alpha$  range extends slightly beyond the range of observed power law scaling exponents  
 212 for PSDs (Diehl and Haardt, 1980, Buonassissi and Dierssen 2010), in order to reduce boundary  
 213 effects during fitting, and we constrained  $\lambda$  to be contained within the bounds of the smallest and  
 214 largest observed particle size for a particular  $N(d)$ . Because  $\lambda$  spans several orders of magnitude,  
 215 any reported  $\lambda$  averages for the remainder of this text were calculated using log10 transformed  $\lambda$   
 216 values and those averages are then converted into non-log transformed values that are simpler  
 217 conceptually. We performed this model fit for all 7808 independent locations at the mean of  
 218 depth bins, or 7.5, 22.5, 47.5, 97.5, 147.5, 222.5, 297.5, 497.5, and 997.5db (hereafter expressed  
 219 as 10, 25, 50, 100, 150, 225, 300, 500, and 1000m).

220 The truncated power law model is a better fit to the data than a power law, with an  
 221 improved adjusted R-squared (accounting for free parameter differences, 0.96 for a truncated  
 222 power law versus 0.95 for a power law), relative percent error (24% for a truncated power law  
 223 versus 27% for a power law), and relative bias (i.e.,  $(N(d) - PSD_{obs}) / PSD_{obs}$ , 7% for a  
 224 truncated power law versus 9% for a power law) across all depths. In this study we choose a  
 225 truncated power law because of the higher overall performance, and because the truncated power  
 226 law parameters offer more insights about the observed PSD shape than a power law alone.

227

## 228 2.4 Carbon flux calculations

229 The applications of measured in situ PSDs introduce additional uncertainty and error into  
 230 the derived measurements of interest, including quantifications of carbon flux (Guidi et al., 2008;  
 231 2016) and aggregate formation (Guidi et al., 2009). PSDs are ingested within a power law  
 232 approximation to calculate carbon flux via

$$233 \quad F = \int_{d_{min}}^{d_{max}} N(d) * A d^b \, dd, \quad [4]$$

234 where  $F$  is carbon flux ( $\text{mg m}^{-2} \text{d}^{-1}$ ),  $N(d)$  is the concentration of particles ( $\# \text{L}^{-1}$ ) with a mean  
 235 equivalent spherical diameter ( $d$ , mm), and  $A$  (12.5) and  $b$  (3.81) are free parameters that were  
 236 first optimized in Guidi et al. (2008) using all available UVP versions with a shared size interval

237 of 250 microns to 1.5mm, and with sampling volumes ranging from 0.28 to 10.5L. The function  
 238 is integrated over the range of size classes available. While  $A$  and  $b$  are empirically derived, they  
 239 conceptually arise from a general mechanistic model that incorporates sinking velocity (via  
 240 Stokes' law, a power law) and carbon content of a particle (modeled as a power law). The  
 241 product of sinking velocity ( $w(d) = \beta d^2$ ) and carbon content ( $m(d) = \alpha d^3$ , both power laws)  
 242 are modeled as a power law, providing the  $A d^b$  term in equation 4. We note that the power law  
 243 formulation for carbon content assumes that all particles of a given size have the same carbon  
 244 content and sinking speed, which is a flawed assumption given current understanding of particle  
 245 characteristics. Given typical power law fits for  $N(d)$ , equation 4 implies an infinite flux with  
 246 increasing particle size, as well as a consistent size-to-flux relationship for equally sized cells,  
 247 which will be violated for cells of different density and/or lability. We argue here that any PSD-  
 248 derived flux formula must be aligned with the known uncertainties of the PSD observations.  
 249 Particularly, the value of  $d_{\max}$  is important if a power law  $N(d)$  is selected because the counts  
 250 of  $N(d_{\max})$  become negligible due to sampling. The value of  $d_{\max}$  is also important when  
 251 comparing across different UVP versions with different size ranges.

252

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

255

$$256 \quad F_m = \sum C * d^{-\alpha} * e^{\frac{-d}{\lambda}} * A d^b. \quad [5]$$

257

258 Where  $F_m$  denotes flux from the modeled number distribution. We tested different values of  $A$   
 259 and  $b$  to reflect the different values used (Kiko et al., 2017, Kriest, 2002, Alldredge, 1998),  
 260 where  $A = 2.8$  and  $b = 2.24$ , noting that the latter  $b$  value might be a more realistic size-sinking  
 261 scaling exponent than  $b = 3.81$  (Cael et al., 2021), because the value of  $b$  in the Guidi et al.,  
 262 (2008) work is not the size-sinking speed relationship, but rather the optimized value when  
 263 compared to observations over a defined size range. We note that we are less interested in the  
 264 specific values of  $A$  and  $b$ , but rather how the fundamental characteristics of flux's functional  
 265 form affect its outcome given modeled sampling uncertainties, as in Cael and Bisson, (2018). For  
 266 the remainder of this paper we use  $A = 2.8$  and  $b = 2.24$  because those values are meant to

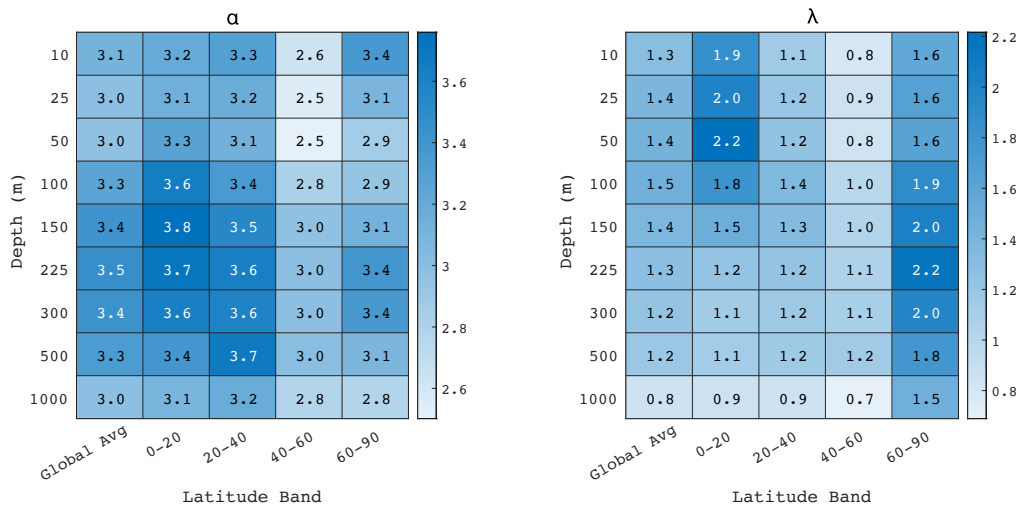


267 represent flux more realistically across the range of sizes and particles thought to contribute to  
 268 flux.

269 One advantage of the modeled PSD in this study is that it can be used to extract the  
 270 particle number outside the range of observed particle sizes. To quantify the sensitivity of the  
 271 flux relationship to different sizes, we included bins two sizes smaller than the first size bin  
 272 observed, as well as two bins larger than the last size bin observed, for each PSD model.  
 273 Operationally this meant including 80 microns to anywhere from 1 mm to 26 mm for the size  
 274 interval. The objective was not to extrapolate widely beyond what has been observed, but rather  
 275 to include size classes within neighboring bins relative to what was actually seen by the UVP, in  
 276 order to assess the sensitivity of flux derived from the UVP. Carbon flux calculated using a  
 277 wider interval for particle sizes was compared to flux calculated from the observed PSD size  
 278 range. In this study we are not concerned about the performance of the flux model (as has been  
 279 done in other studies, Guidi et al 2008, Fender et al., 2019). We instead ask, ‘how does using a  
 280 more complete PSD affect flux calculations?’

## 281 4 Results and Discussion

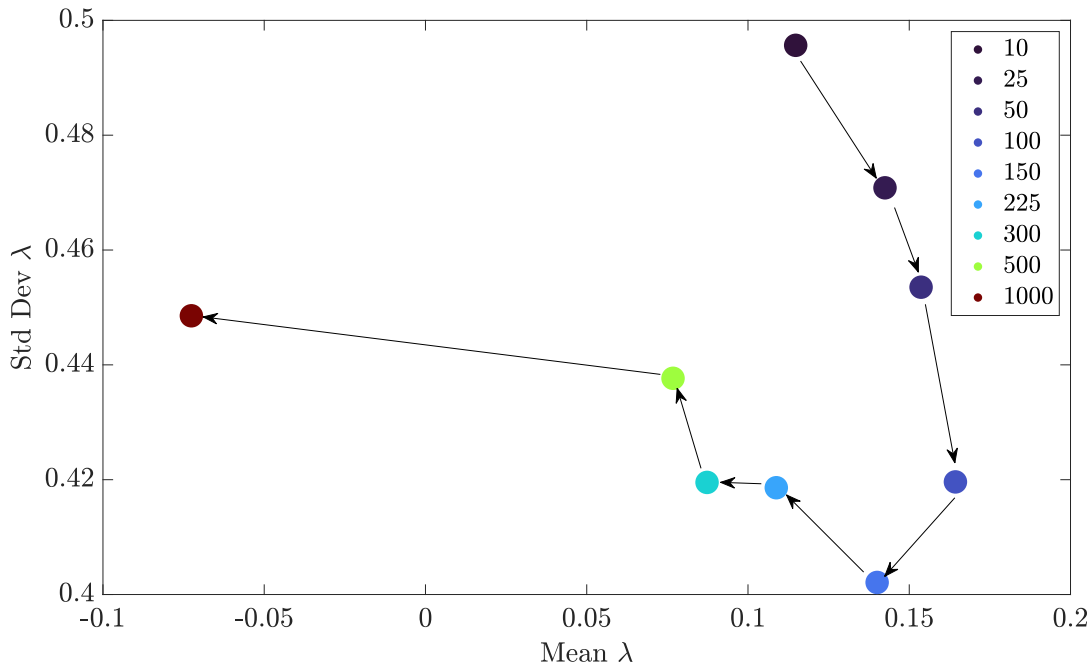
### 282 4.1 Global $\alpha$ and $\lambda$ values



283  
 284 Figure 3. Heatmap of  $\alpha$  (left) and  $\lambda$  (right, normalized to  $\lambda_0 = 1$  mm) based on latitudinal bands  
 285 and depth.

286 Global values of retrieved  $\alpha$  and  $\lambda$  reveal patterns across space and depth (Figure 3,  
 287 Supplementary Figures 1,2). The highest average  $\alpha$  values (3.8) are in moderate depths (100-300

288 m) for places equatorward of 20 degrees. In general,  $\alpha$  varies throughout the water column, with  
 289 larger values between depths of 150 and 500m, and lower values at the surface and at 1000m  
 290 depth.  $\lambda$  generally decreases with depth, where the global average  $\lambda$  decreases from 1.3 at 10m to  
 291 0.8 at 1000m, and in all places the surface  $\lambda$  value exceeds  $\lambda$  at 1000m, if only slightly.



292

293 Figure 4. Standard deviation in  $\log_{10}(\lambda)$  plotted against the mean  $\log_{10}$ -transformed  $\lambda/\lambda_0$   
 294 (where  $\lambda_0 = 1\text{mm}$ ) based on depth (colors). Arrows denote the transition from shallow (top right)  
 295 to deep (top left) samples.

296 The standard deviation of  $\lambda$  is highest for the surface ocean and deepest observations at  
 297 1000m (Figure 4). In between the surface and depth,  $\lambda$  standard deviation and  $\lambda$  mean (all  $\log_{10}$   
 298 transformed) have a qualitative clockwise trend (Figure 4) where the average  $\lambda$  changes only  
 299 subtly at depths  $< 1000\text{m}$  while the standard deviation decreases from  $\sim 10^{0.5}$  to  $\sim 10^{0.4}$ . The  
 300 variability (given by the standard deviation) in  $\lambda$  decreases with depth up to 150m but thereafter  
 301 increases to a smaller degree. These results do not necessarily mean there are no big particles (or  
 302 ‘dragon kings’, Bochdansky et al., 2016) in the deep, but rather the UVP5 is not observing them.

303 The model parameters  $\alpha$  and  $\lambda$  from the truncated power law fit to observed PSD reflect  
 304 the relative dominance of small versus large particles, and also indicate the heavy (or not) tail-  
 305 ness of the size distribution. In essence,  $\alpha$  mostly quantifies the mid-range behavior in the PSD  
 306 and  $\lambda$  mostly quantifies the upper-range behavior. Although the model is statistical in nature,

307 quantifying the PSD slope and size interval where a power law is applicable gives more  
308 information about the shape of the PSD than slope alone. In other words, conventional power  
309 law fits to PSD assume that a power law is appropriate over the entire size distribution, and the  
310 parameter power law model may not be ideal for characterizing the PSD shape from the UVP.

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

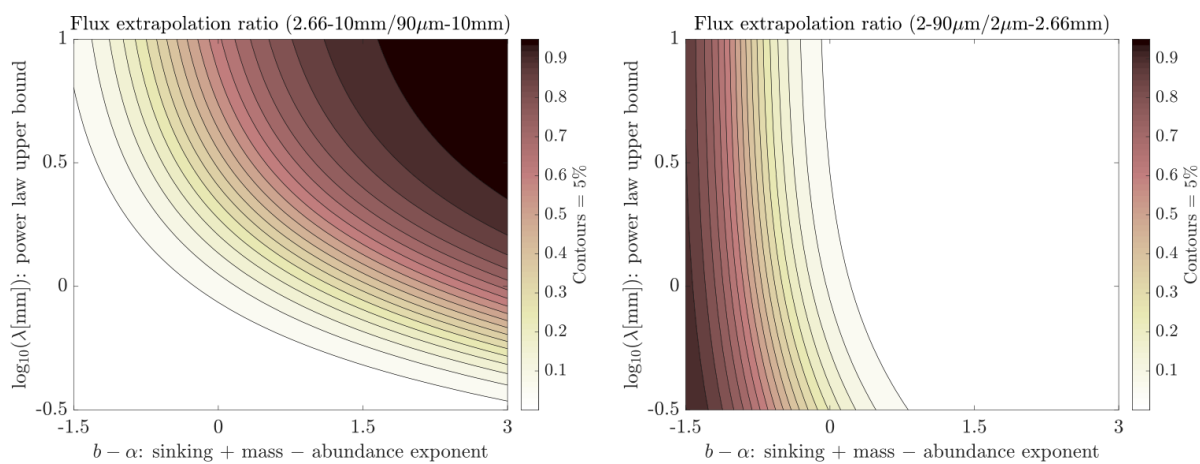
323 Trends in  $\alpha$  are less straightforward. In nearly every latitudinal band,  $\alpha$  increases at  
324 moderate depths, indicating a higher prevalence of small particles, then decreases at deeper  
325 depths. The reported  $\alpha$  and  $\lambda$  values here may be useful in future studies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

## 332 **4.2. Extrapolation and sampling uncertainties**

333 We acknowledge that it is not sensible to calculate flux in the surface ocean using PSD  
334 observations that are unlikely to comprise sinking particles, when considering either the UVP or  
335 another PSD-resolving instrument. Here we consider particles  $2\mu\text{m}$  and above, and we use the  
336 following scaling argument to estimate that any sinking by particles  $<2\mu\text{m}$  can be considered

337 negligible. For a particle to be considered as sinking, its vertical transport from sinking must be  
 338 greater than its vertical transport from ambient turbulent fluid motions. Balancing the two  
 339 processes, a particle's minimum sinking speed  $w_{min} \sim \sqrt{\kappa/\tau}$  will be  $\sim 3 \text{ m day}^{-1}$ , assuming  $\kappa \sim$   
 340  $10^{-4} \text{ m}^2 \text{ s}^{-1}$  [Munk, 1966] and  $\tau = 1 \text{ day}$  is a characteristic measurement, diel, and small particle  
 341 lifetime timescale. Then if particles' sinking speeds scale approximately as  $w(d) \sim 100 d^{-0.63} \text{ m}$   
 342  $\text{day}^{-1}$  (Kriest et al 2002, Cael et al, 2021),  $w \sim w_{min}$  when  $d \sim 2 \mu\text{m}$ . Note also  $2 \mu\text{m}$  is roughly  
 343 the smallest particle size that can be estimated by other instruments in a sinking or flux context  
 344 (Cael and White, 2020).

345 There are particles that contribute to flux that are not captured by the UVP's sampling  
 346 volume and specifications. Under what conditions or assumptions are the observed particles  
 347 sufficiently representative of the total particle population's flux? Figure 5 shows how particles  
 348 outside the UVP5-observed size range contribute to total flux, for a truncated power-law particle  
 349 size distribution and a power-law size-flux relationship. If  $b$  is the exponent dictating how  
 350 sinking and mass (or carbon or other elemental content) together scale with particle size, and  $\alpha$  is  
 351 the exponent dictating how abundance scales with particle size within the power-law scaling  
 352 range, the contribution to flux by particles of a given size will be determined by their difference,  
 353  $b - \alpha$ . The contribution of large particles will also be determined by  $\lambda$ , the particle size where  
 354 the power law is truncated.



355  
 356

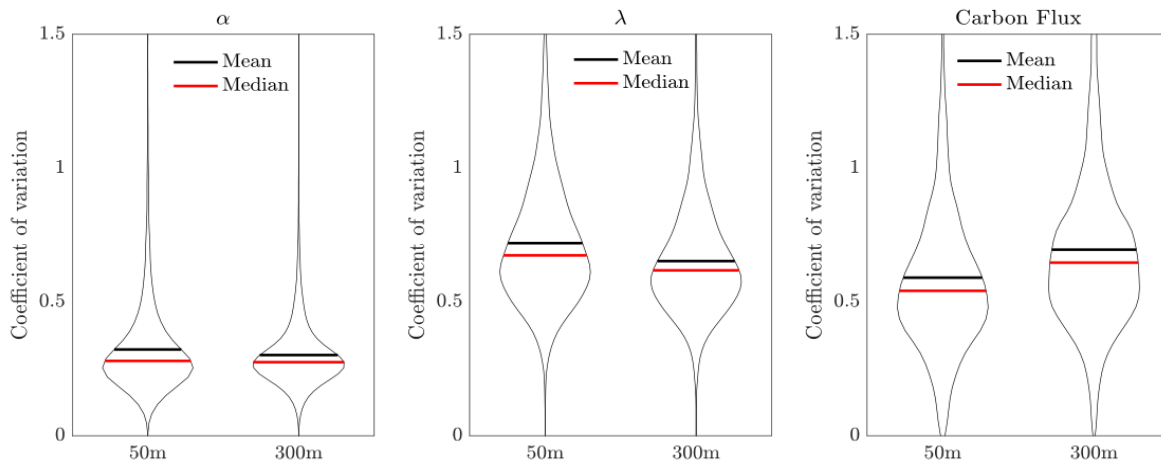
357 Figure 5. Theoretical flux extrapolation ratio as a function of the difference between  $b$  (sinking +  
 358 mass, = 3.81 or 2.24) and  $\alpha$ , and the upper bound particle size where a power law is  
 359 appropriate.

360

361 Figure 5a shows the fraction of flux in the 90 $\mu$ m-10mm particle ESD range for which  
362 2.66-10mm particles are responsible, as a function of  $b - \alpha$  and  $\lambda$ . We note that this is a  
363 conceptual figure that serves only to illustrate under what conditions particle fluxes can be  
364 dominated by particles of different sizes. Clearly both parameters play a role; when  $\lambda$  and  $b - \alpha$   
365 are both small, meaning large particles are rare and particles' sinking-and-mass size-dependence  
366 is weaker than particles size-abundance relationship, large particles contribute very little to total  
367 flux so almost none of the flux occurs in the 2.66-10mm size range. When either  $\lambda$  or  $b - \alpha$  are  
368 large, however, particles in this range do contribute appreciably to overall flux. When both  $\lambda$  and  
369  $b - \alpha$  are large, meaning the power-law distribution extends out to multi-millimeter particles and  
370 the sinking-and-mass size dependence of particles is strong relative to particles' size-abundance  
371 relationship, most of the flux actually can occur in this 2.66-10mm size range. In contrast, Figure  
372 5b shows the same but for small particles, comparing 2-90 $\mu$ m particles against 2 $\mu$ m-2.66mm  
373 particles. In this case the dependence on  $\lambda$  is unsurprisingly very weak, but we do see that as long  
374 as approximately  $b - \alpha < 0$ , *i.e.* that particle abundance scales more strongly with size than  
375 particles' sinking and mass, much or even most of the flux occurs in particles <90 $\mu$ m. Although  
376 the UVP does not measure particles smaller than 90  $\mu$ m, these figures underscore that accurate  
377 UVP-based flux estimates require understanding the controls on and variability of particles'  
378 sinking-size and mass-size relationships, the prevalence of large particles, and the slope of the  
379 particle size distribution. We include them to demonstrate that sampling uncertainty includes  
380 uncertainty due to particles outside the detection limit of the UVP or any PSD-resolving  
381 instrument.

382 There is high variability in retrieved  $\alpha$ ,  $\lambda$ , and carbon flux arising from the sample volume  
383 uncertainty using the observed size range from UVP observations, as calculated from 100  
384 simulations with varying  $N(d)$  (informed by the observed  $N(d)$  and sampling volume) for all  
385 locations in this study at 50 and 300m (Figure 6, Supplementary Figure 3, see also section 2.2 for  
386 procedure). Across all three variates, the coefficient of variation is smallest for  $\alpha$  at either depth  
387 than it is for  $\lambda$  and carbon flux. The median coefficient of variation for  $\alpha$  is ~25% at both depths,  
388 while the median coefficient of variation for  $\lambda$  is nearly 60% for the 50m case, and 55% for the  
389 300m case. The coefficient of variation for carbon flux arising from the sampling volume  
390 uncertainty is highest for deep particles (~65-70%) compared to the 50m case (50-55%). The  
391 width of the coefficient of variation distributions varies for all three variates as well, with  $\alpha$

392 showing the tightest range, followed by  $\lambda$  and carbon flux. We emphasize that the coefficient of  
 393 variations reported here are due only to measurement error and not due to natural variability,  
 394 which we could not fully characterize due to lack of repeat data (see Supplementary Figure 4).  
 395 As a test of how larger sampling volumes may influence the coefficient of variation in  $\alpha$ ,  $\lambda$ , and  
 396 carbon flux, we also ran the bootstrapping procedure using simulated sampling volumes that are  
 397 double the observed sampling volume. Doubling the sample volume reduces the coefficient of  
 398 variation in carbon flux to a median of 56% compared to a median of 67% in the 300m case  
 399 (Supplementary Figure 5). Note that we did not adjust the  $N(d)$  (doing so would preserve the  
 400 particle concentration) so that we could isolate the relative effect of enhancing sampling volume  
 401 in a statistical sense.

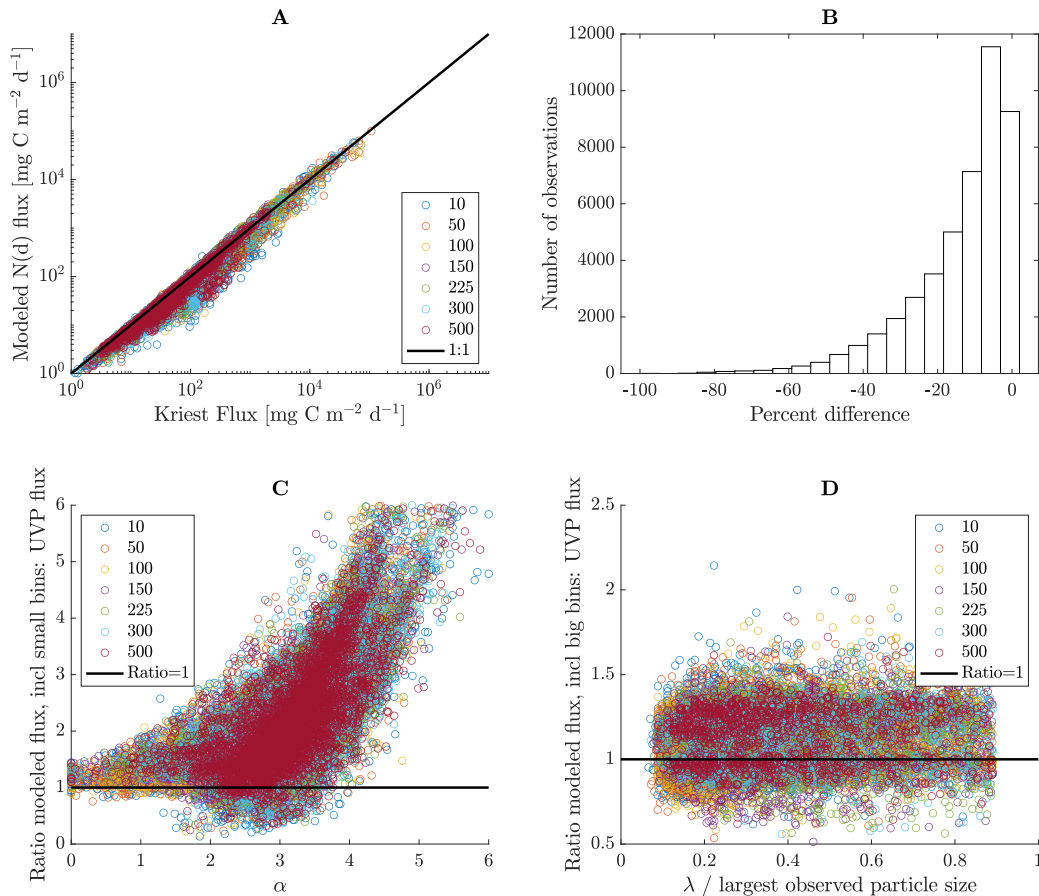


402  
 403 Figure 6. Violin plots for coefficient of variation in  $\alpha$  (left),  $\lambda$  (middle) and carbon flux based on  
 404 either 50 m or 300m (resulting from the bootstrap procedure). The coefficient of variation  
 405 reported in this figure is due only to the sampling volume uncertainty.  
 406

### 407 4.3 Sensitivity of modeled carbon flux to particle size

408 Carbon flux calculations using observed and modeled PSD (over a shared size range, i.e.,  
 409 the fixed lower limit of 128  $\mu\text{m}$  for each profile and an upper limit dictated by the largest  
 410 observed particle size) are well correlated as expected, agreeing within 10% for the majority of  
 411 locations and depths (Figure 7A and Figure 7B). The flux relationship is highly sensitive to the  
 412 inclusion of smaller particle sizes. In some cases, the ratio of flux calculated using two size bins  
 413 smaller than observed to the ratio of flux calculated using only observed particle sizes is nearly  
 414 6. The high degree of sensitivity to small particles directly scales with the  $\alpha$  value of the modeled

415 PSD. For small  $\alpha$  (indicating a dominance of larger particles relative to small), the inclusion of  
 416 smaller size classes makes relatively little difference in the flux calculations (Figure 7C).  
 417 However, for  $\alpha$  larger than 2, the sensitivity of flux to smaller size classes is substantial, with  
 418 relative differences exceeding a factor of 2.



419  
 420 Figure 7. A. Flux comparisons between modeled N(d) flux (y axis) and N(d) flux from UVP  
 421 observations ('Kriest flux') over the same particle size range. B. Histogram of percent difference  
 422 (relative to 'Kriest flux') between all flux determinations in A. C. The ratio of modeled N(d) flux  
 423 (including 2 size bins smaller than the observed size range) to observed N(d) flux as a function  
 424 of depth (colored dots) and alpha value. Black line represents equivalent fluxes. D. The ratio of  
 425 modeled N(d) flux (including 2 size bins larger than the observed size range) to observed N(d)  
 426 flux as a function of depth (colored dots) and lambda value relative to largest observed particle  
 427 size. Black line represents equivalent fluxes.

428 On the flip side, the flux relationship with PSD is much less sensitive to the inclusion of  
 429 larger size bins relative to what was observed (Figure 7D). Recall that the modeled PSD in this

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

442         The sensitivity of PSD-derived flux to either small or large particles is robust to changes  
443 in the free parameter values of **A** and **b**. When the Guidi et al., 2008 formulation was applied  
444 (Supplementary Figure 6), a similar sensitivity was observed, although the overall magnitude of  
445 flux was much enhanced compared to the Kriest et al., (2002) formulation. Ideally, any  
446 biogeochemical model will not be sensitive to the inclusion of either small (< 128  $\mu\text{m}$ ) or large  
447 size classes (> 1mm) because small particles are not thought to contribute the bulk of carbon flux  
448 (see Michaels and Silver, 1988; Bopp et al., 2005), and because large particles are rarely  
449 observed and highly uncertain (and so any model relying on large particles for flux would be  
450 highly uncertain as well). In this study we found that no matter which flux parameter values are  
451 used (i.e., the empirically derived Guidi et al., 2008 values or the more mechanistic Kriest et al.,  
452 2002 values) the power law flux relationship is still highly sensitive to the inclusion of small  
453 particles. This sensitivity is a surprising result, and likely arises because the high abundance of  
454 small particles overcomes their relatively small diameter (in this case, 80 microns) to contribute  
455 a large amount of flux (up to 6 times the amount of carbon flux calculated using a minimum  
456 particle size of 128 microns). If we included even smaller size classes we expect the flux to  
457 increase further. The idea that small particles can contribute substantial flux stands contrary to  
458 what is expected from observations of sinking particles in a natural setting (Cael et al., 2021 and  
459 refs therein). However, some flux models also predict a larger contribution of small sinking flux  
460 (e.g., Bisson et al., 2020, Siegel et al., 2014) than is expected (Durkin et al., 2015, Cael and



461 White, 2020, Cael et al., 2021). We note that the small particles are not thought to contribute  
462 substantially to flux when it is assumed that these small particles are formed in the surface,  
463 because they will be remineralized in their 100s of meters transit to depth. However, small  
464 particles may actually dominate flux in deeper waters through disaggregation processes (Kiko et  
465 al., 2017, Bianchi et al., 2018).

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

#### 473 **4.4 Limitations of using the UVP to assess particle flux**

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

486 While some attention has been paid to optimizing the parameters ( $\mathbf{A}$ ,  $\mathbf{b}$ ) of the carbon flux  
487 model (Guidi et al., 2008, Fender et al., 2019), it seems a larger problem is in the foundation of  
488 the flux model itself. We calculated carbon fluxes incorporating the reported Guidi et al., 2008  
489 parameter standard deviations to learn how carbon flux is uncertain based on parameter value  
490 uncertainty. Uncertainty in  $\mathbf{A}$  resulted in a median 21% relative error in carbon flux while

491 uncertainty in  $\mathbf{b}$  resulted in median 19% relative error using the profiles in this study at all  
492 depths. Using the Guidi parameters instead of the Kriest parameters results in median differences  
493 approaching a factor of 2 (compared to factors of 6 when incorporating 80-128 $\mu\text{m}$  particles).

494         Given the large uncertainty also associated with sampling volume, we recommend  
495 optimizing the flux model using the same UVP version, or by accounting for uncertainty directly  
496 in the model optimizations (e.g., Bisson et al., 2018). We note that the Guidi et al., 2008 study  
497 used a size interval of 250 microns to 1.5mm in order to incorporate older UVP versions (with  
498 sampling volumes ranging from 0.28 – 10.5L) and did not use UVP-5 data. However, although  
499 the Guidi et al., 2008 parameters were optimized over a different size interval than was used  
500 here, the specific values of  $\mathbf{A}$  and  $\mathbf{b}$  will not modify the sensitivity of flux to small size classes  
501 (compare Figure 7 with Supplementary Figure 6). Normally,  $\mathbf{A}$  and  $\mathbf{b}$  values are optimized  
502 within the boundaries of the size spectrum imaged by the UVP, and therefore any regionally  
503 optimized PSD-flux relationship is not necessarily problematic to use, even though the  
504 theoretical underpinnings of such a relationship are imperfect.

505         It can be instructive to think of the flux model as a transfer function ( $\mathbf{A}\mathbf{d}^{\mathbf{b}}$ ) that is  
506 multiplied by  $\mathbf{N}(\mathbf{d})$ . The transfer function is a monotonic power law that grows substantially at  
507 larger particle sizes. Therefore, if anything, the flux model is expected to be sensitive to rare  
508 instances of large particles, depending on  $\mathbf{b}$  versus  $\alpha$ . If particles in the ocean grew indefinitely,  
509 infinite flux would be expected from this relationship. On the other hand, infinite flux is possible  
510 with smaller particle sizes if the concentration of particles grows more than particle size  
511 decreases, as was the case in this study. If the true PSD were not monotonic (i.e., increasingly  
512 higher concentrations of particles at lower particle sizes), a monotonic flux model (such as the  
513 power law used here) may be sufficient. What instead might be needed is a transfer function that  
514 quantifies the probability of a given particle size to sink. Small sized particles would accordingly  
515 have low probability, as would larger particles that are ultimately living zooplankton (or fish in  
516 the extreme case). Medium to medium-large sized particles would have moderate to high  
517 probability of becoming carbon flux, which might yield a more realistic carbon flux model. More  
518 work is needed to improve the PSD-derived carbon flux relationship, and especially the size to  
519 sinking carbon uncertainty, which is outside the scope of this paper.

520         Finally, particles sized by the UVP include living and non-living particles, which adds  
521 uncertainty to flux calculations derived from PSD alone. If only non-living particles were

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

#### 528 **4.5 Future applications of using a truncated power law to model PSD**

529 In this study we found enhanced performance of the modeled PSD when using a  
530 truncated power law rather than a power law. Truncated power laws offer more information  
531 about a PSD distribution compared to a power law because the behavior of the distribution is  
532 characterized through two main parameters ( $\alpha$ ,  $\lambda$ ) rather than just one ( $\alpha$ , in the case of a power  
533 law). There are several applications to using a truncated power law besides what has been  
534 explored here. First, with an improved model for PSD, one could extrapolate the PSD to quantify  
535 the carbon content of particles in the particulate fraction globally. Second, extrapolating the PSD  
536 using a truncated power law may enable improved respiration rates as derived from UVP  
537 observations, as current estimates are limited by the size resolved by the UVP (Kalvelage et al.,  
538 2015, Thomsen et al., 2019). Third, the current UVP data hosted by Ecotaxa includes both living  
539 and non-living particles. Future work may explore whether or not the  $\lambda$  values will be useful to  
540 identify when the PSD spectrum transitions from particles to larger zooplankton (Forest et al.,  
541 2012).

542

#### 543 **5 Recommendations for future work**

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

- 549 1. *Sampling uncertainty*: Future UVP designs can reduce sampling uncertainty by  
550 increasing the sampling volume of the instrument. Current UVP designs can reduce

551 sampling uncertainty to some extent by performing multiple casts of repeat sampling.  
552 More work needs to be done in order to distinguish aggregates from living plankton for  
553 particles in the observable size range, preferentially down to 2  $\mu\text{m}$  size.

554 2. *Size to sinking uncertainty*: Although unconventional, a UVP fastened to a Lagrangian  
555 sediment and/or gel trap that is oriented with a side-viewing camera may allow sinking  
556 speed to be assessed via several images, where sediment trap flux, particle sinking speeds  
557 from a gel trap, and particle size information would be coupled and coincident. Similarly,  
558 in situ sinking speeds could be obtained using Particle Imaging Velocimetry (Cartwright  
559 et al. 2013), optimally during the upcast of a CTD/UVP profile. This logistically less  
560 demanding approach could yield PSD observations over the entire water column and  
561 coincident particle sinking speed observations at different water depths. Targeting  
562 blooms of different organisms with UVP observations may also help to improve size-  
563 sinking relationships. Direct observations that better constrain the size-sinking scaling  
564 relationship globally, in different environments, and/or for different particle types is  
565 essential for improving uncertainties in UVP-derived fluxes. (Cael et al., 2021).

566 3. *Size to carbon content uncertainty*: Dense particles may have a different reflectance than  
567 less dense particles (based on the fractal dimension) which might provide a way to semi-  
568 quantitatively assess particle composition from the contrast of the images. If such an  
569 exercise is possible, the contrast of images may add information content to the flux  
570 relationship so particle size and concentration are not the only variables. Further  
571 classification of particle images into e.g. fecal pellets, marine snow and other types of  
572 detrital matter and the application of class specific size to carbon ratios might also reduce  
573 the errors in flux calculation (Durkin et al., 2021). Finally, due to remineralization,  
574 carbon content might also decrease over depth without large changes in size or  
575 appearance of the particles. Therefore, further work is needed to characterize the carbon  
576 to size relationship of detrital particles at different depths.

577 Although we did not investigate the uncertainty associated with the conversion of UVP pixels to  
578 a particle size, more work is needed to characterize any error and uncertainty arising from  
579 particle shape differences and assumed spherical diameters. Improved edge detection of pixels is

580 needed, as well as a sensitivity analysis of how threshold values for edge detection affect particle  
581 size (as is also advised by Giering et al., 2020).

## 582 **6 Summary**

583 In this study we focused on UVP sampling uncertainties and how they propagate into  
584 derived estimates of carbon flux. The PSD observations from the UVP5 have a sampling  
585 uncertainty  $\sqrt{N}/V$ . The sampling uncertainty of PSD observations results in an uncertainty  
586 slightly greater than 50% for carbon flux. The extrapolated carbon flux from the UVP is based  
587 on a relationship that is highly sensitive (up to 6-fold different) to the inclusion of particles  
588 slightly smaller than what was observed. We advocate for a revised carbon flux relationship that  
589 is possibly non-monotonic and considers the probability of a given particle to become carbon  
590 flux. In the absence of an improved carbon flux relationship, carbon flux calculations should be  
591 made using parameters specific to a particular region and depth to prevent large errors. Future  
592 work may benefit by using UVP data in unconventional ways, such as coupling a UVP and  
593 sediment trap in the same water mass, and/or by performing image analysis on the particular  
594 pixels comprising a particle.

## 595 **Acknowledgments and Data Statement**

596 Cael was supported by the National Environmental Research Council (NE/R015953/1) and the  
597 Horizon 2020 Framework Programme (820989, project COMFORT, our common future ocean  
598 in the Earth system—quantifying coupled cycles of carbon, oxygen, and nutrients for  
599 determining and achieving safe operating spaces with respect to tipping points). The work  
600 reflects only the authors' view; the European Commission and their executive agency are not  
601 responsible for any use that may be made of the information the work contains.

602 RK acknowledges support via a "Make Our Planet Great Again" grant of the French National  
603 Research Agency within the "Programme d'Investissements d'Avenir"; reference "ANR-19-  
604 MPGA-0012", the BMBF funded project "CUSCO" and funding from the European Union's  
605 Horizon 2020 research and innovation programme for the TRIATLAS project under  
606 grant agreement No 817578. The UVP data used in this study can be accessed via PANGAEA,  
607 <https://doi.pangaea.de/10.1594/PANGAEA.924375>. Log ins are required for downloading, and  
608 the data span 2008-2020. More processing details are in Kiko et al., 2021.

609 **References**

- 610 Alldredge, A. L., & Gotschalk, C. (1988). In situ settling behavior of marine snow 1. *Limnology*  
611 *and Oceanography*, 33(3), 339-351.
- 612
- 613 Alldredge, A. (1998). The carbon, nitrogen and mass content of marine snow as a function of  
614 aggregate size. *Deep Sea Research Part I: Oceanographic Research Papers*, 45(4-5), 529-541.
- 615
- 616 Bianchi, D., Weber, T. S., Kiko, R., & Deutsch, C. (2018). Global niche of marine anaerobic  
617 metabolisms expanded by particle microenvironments. *Nature Geoscience*, 11(4), 263-268.
- 618
- 619 Bisson, K. M., Siegel, D. A., DeVries, T., Cael, B. B., & Buesseler, K. O. (2018). How data set  
620 characteristics influence ocean carbon export models. *Global Biogeochemical Cycles*, 32(9),  
621 1312-1328.
- 622
- 623 Bisson, K., Siegel, D. A., & DeVries, T. (2020). Diagnosing mechanisms of ocean carbon export  
624 in a satellite-based food web model. *Frontiers in Marine Science*, 7, 505.
- 625
- 626 Bochdansky, A. B., Clouse, M. A., & Herndl, G. J. (2016). Dragon kings of the deep sea: marine  
627 particles deviate markedly from the common number-size spectrum. *Scientific reports*, 6(1), 1-7.
- 628
- 629 Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms  
630 distribution to global warming and potential implications: A global model study. *Geophysical*  
631 *Research Letters*, 32(19).
- 632
- 633 Boss, E., Twardowski, M. S., & Herring, S. (2001). Shape of the particulate beam attenuation  
634 spectrum and its inversion to obtain the shape of the particulate size distribution. *Applied*  
635 *Optics*, 40(27), 4885-4893.
- 636
- 637 Buonassissi, C. J., & Dierssen, H. M. (2010). A regional comparison of particle size distributions  
638 and the power law approximation in oceanic and estuarine surface waters. *Journal of*  
639 *Geophysical Research: Oceans*, 115(C10).

- 640
- 641 Cael, B. B., Cavan, E. L., & Britten, G. L. (2021). Reconciling the size-dependence of marine  
642 particle sinking speed. *Geophysical Research Letters*, *48*(5), e2020GL091771.
- 643
- 644 Cael, B. B., & White, A. E. (2020). Sinking versus suspended particle size distributions in the  
645 North Pacific Subtropical Gyre. *Geophysical Research Letters*, *47*(15), e2020GL087825.
- 646
- 647 Cael, B. B., & Bisson, K. (2018). Particle flux parameterizations: Quantitative and mechanistic  
648 similarities and differences. *Frontiers in Marine Science*, *5*, 395.
- 649
- 650 Cartwright, G. M., Friedrichs, C. T., & Smith, S. J. (2013). A test of the ADV-based Reynolds  
651 flux method for in situ estimation of sediment settling velocity in a muddy estuary. *Geo-Marine*  
652 *Letters*, *33*(6), 477-484.
- 653
- 654 Chase, A. P., Kramer, S. J., Haëntjens, N., Boss, E. S., Karp-Boss, L., Edmondson, M., & Graff,  
655 J. R. (2020). Evaluation of diagnostic pigments to estimate phytoplankton size classes.  
656 *Limnology and Oceanography: Methods*, *18*(10), 570-584.
- 657
- 658 Dall'Olmo, G., Westberry, T. K., Behrenfeld, M. J., Boss, E., & Slade, W. H. (2009). Significant  
659 contribution of large particles to optical backscattering in the open ocean. *Biogeosciences*, *6*(6),  
660 947.
- 661
- 662 Diehl, P., & Haardt, H. (1980). Measurement of the spectral attenuation to support biological-  
663 research in a plankton tube experiment. *Oceanologica Acta*, *3*(1), 89-96.
- 664
- 665 Durkin, C. A., Buesseler, K. O., Cetinić, I., Estapa, M. L., Kelly, R. P., & Omand, M. (2021). A  
666 visual tour of carbon export by sinking particles. *bioRxiv*.
- 667
- 668 Fender, C. K., Kelly, T. B., Guidi, L., Ohman, M. D., Smith, M. C., & Stukel, M. R. (2019).  
669 Investigating particle size-flux relationships and the biological pump across a range of plankton  
670 ecosystem states from coastal to oligotrophic. *Frontiers in Marine Science*, *6*, 603.

- 671  
672 Fink, D. (1997). A compendium of conjugate priors. See [http://www. people. cornell.](http://www.people.cornell.edu/pages/df36/CONJINTRnew%20TEX.pdf)  
673 [edu/pages/df36/CONJINTRnew% 20TEX. pdf](http://www.people.cornell.edu/pages/df36/CONJINTRnew%20TEX.pdf), 46.  
674
- 675 Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., & Babin, M. (2012).  
676 Size distribution of particles and zooplankton across the shelf-basin system in southeast Beaufort  
677 Sea: combined results from an Underwater Vision Profiler and vertical net  
678 tows. *Biogeosciences*, 9(4), 1301-1320.  
679
- 680 Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014).  
681 Bayesian Data Analysis, volume third ed.  
682
- 683 Giering, S. L., Cavan, E. L., Basedow, S. L., Briggs, N., Burd, A. B., Darroch, L. J., ... &  
684 Lindsay, D. J. (2020). Sinking organic particles in the ocean—flux estimates from in situ optical  
685 devices.  
686
- 687 Guidi, L., Jackson, G. A., Stemmann, L., Miquel, J. C., Picheral, M., & Gorsky, G. (2008).  
688 Relationship between particle size distribution and flux in the mesopelagic zone. *Deep Sea*  
689 *Research Part I: Oceanographic Research Papers*, 55(10), 1364-1374.  
690
- 691 Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., ... & Gorsky, G.  
692 (2009). Effects of phytoplankton community on production, size, and export of large aggregates:  
693 A world-ocean analysis. *Limnology and Oceanography*, 54(6), 1951-1963.  
694
- 695 Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., ... & Coelho, L. P.  
696 (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, 532(7600),  
697 465-470.  
698
- 699 Gorsky, G., Picheral, M., & Stemmann, L. (2000). Use of the Underwater Video Profiler for the  
700 study of aggregate dynamics in the North Mediterranean. *Estuarine, Coastal and Shelf Science*,  
701 50(1), 121-128.



702

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

704

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

707

708 Karp-Boss, L., Azevedo, L., & Boss, E. (2007). LISST-100 measurements of phytoplankton size  
709 distribution: Evaluation of the effects of cell shape. *Limnology and Oceanography: Methods*,  
710 5(11), 396-406.

711

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

715

716 Kiko, R., Bianchi, D., Grenz, C., Hauss, H., Iversen, M., Kumar, S., ... & Robinson, C. (2020).  
717 Zooplankton and Nekton: Gatekeepers of the Biological Pump. *Frontiers in Marine Science*, 7,  
718 545.

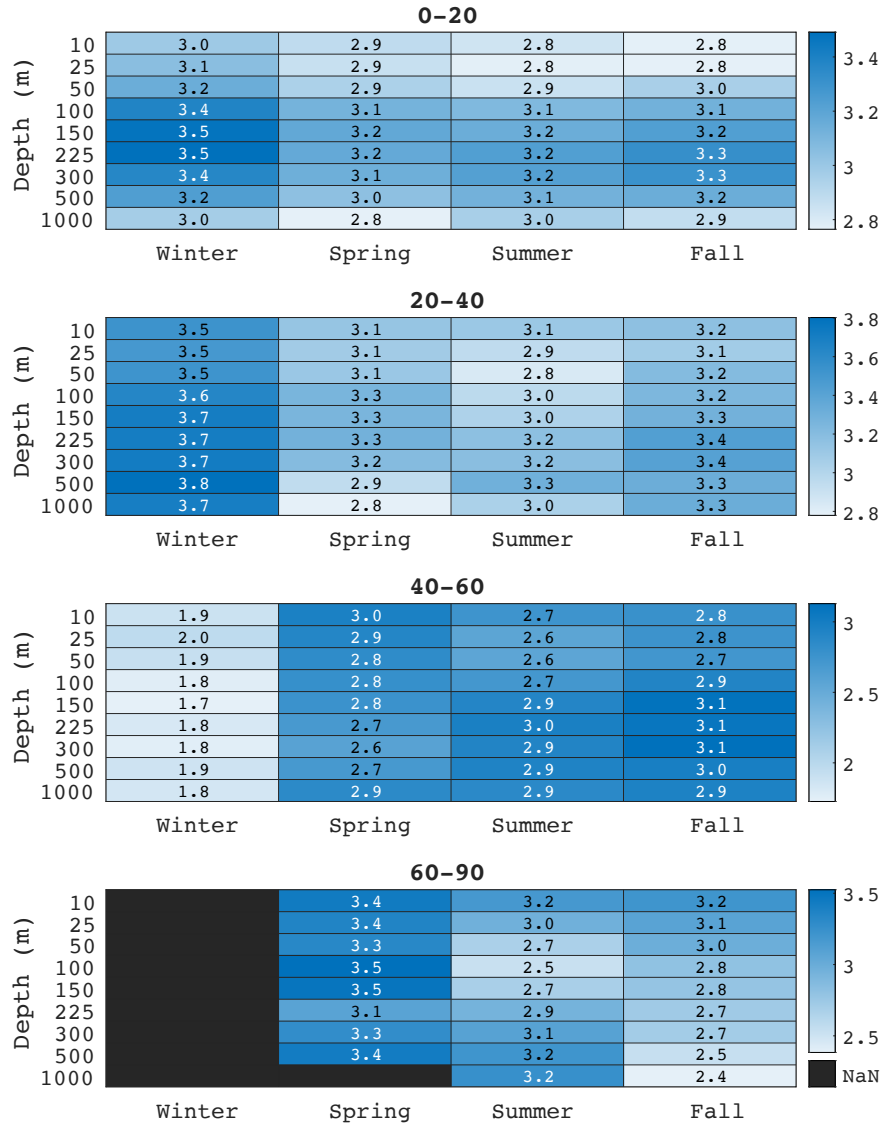
719

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

730

- 731 Kiko, R., Biastoch, A., Brandt, P., Cravatte, S., Hauss, H., Hummels, R., ... & Stemmann, L.  
732 (2017). Biological and physical influences on marine snowfall at the equator. *Nature*  
733 *Geoscience*, 10(11), 852-858.
- 734
- 735 Kriest, I. (2002). Different parameterizations of marine snow in a 1D-model and their influence  
736 on representation of marine snow, nitrogen budget and sedimentation. *Deep Sea Research Part I:*  
737 *Oceanographic Research Papers*, 49(12), 2133-2162.
- 738
- 739 Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., ... & Appeltans, W.  
740 (2019). Globally consistent quantitative observations of planktonic ecosystems. *Frontiers in*  
741 *Marine Science*, 6, 196.
- 742
- 743 Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). The BUGS Book: A  
744 Practical Introduction to Bayesian Analysis (CRC, Boca Raton, FL).
- 745
- 746 Michaels, A. F., & Silver, M. W. (1988). Primary production, sinking fluxes and the microbial  
747 food web. *Deep Sea Research Part A. Oceanographic Research Papers*, 35(4), 473-490.
- 748
- 749 Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., & Gorsky, G. (2010). The  
750 Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of  
751 particle size spectra and zooplankton. *Limnology and Oceanography: Methods*, 8(9), 462-473.
- 752
- 753 Reynolds, R. A., Stramski, D., Wright, V. M., & Woźniak, S. B. (2010). Measurements and  
754 characterization of particle size distributions in coastal waters. *Journal of Geophysical Research:*  
755 *Oceans*, 115(C8).
- 756
- 757 Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., & Boyd, P. W.  
758 (2014). Global assessment of ocean carbon export by combining satellite observations and food-  
759 web models. *Global Biogeochemical Cycles*, 28(3), 181-196.
- 760

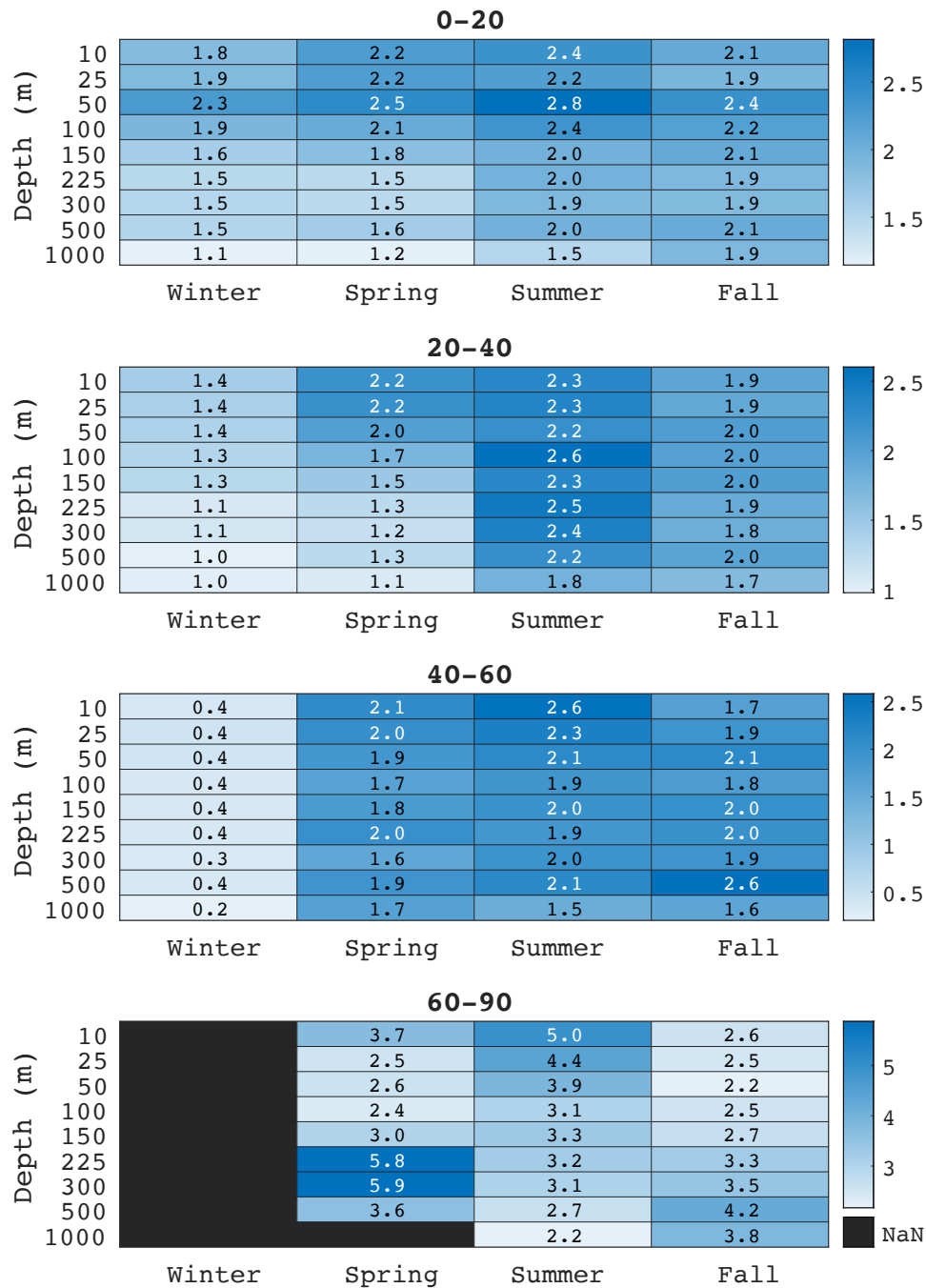
- 761 Sheldon, R. W., Prakash, A., & Sutcliffe Jr, W. (1972). The size distribution of particles in the  
762 Ocean 1. *Limnology and oceanography*, 17(3), 327-340.
- 763
- 764 Slade, W. H., & Boss, E. (2015). Spectral attenuation and backscattering as indicators of average  
765 particle size. *Applied optics*, 54(24), 7264-7277.
- 766
- 767 Stemmann, L., & Boss, E. (2012). Plankton and particle size and packaging: from determining  
768 optical properties to driving the biological pump. *Annual Review of Marine Science*, 4, 263-290.
- 769
- 770 Thomsen, S., Karstensen, J., Kiko, R., Krahnemann, G., Dengler, M., & Engel, A. (2019). Remote  
771 and local drivers of oxygen and nitrate variability in the shallow oxygen minimum zone off  
772 Mauritania in June 2014. *Biogeosciences*, 16(5), 979-998.
- 773
- 774 White, A. E., Letelier, R. M., Whitmire, A. L., Barone, B., Bidigare, R. R., Church, M. J., &  
775 Karl, D. M. (2015). Phenology of particle size distributions and primary productivity in the  
776 North Pacific subtropical gyre (Station ALOHA). *Journal of Geophysical Research: Oceans*,  
777 120(11), 7381-7399.
- 778
- 779
- 780
- 781



782

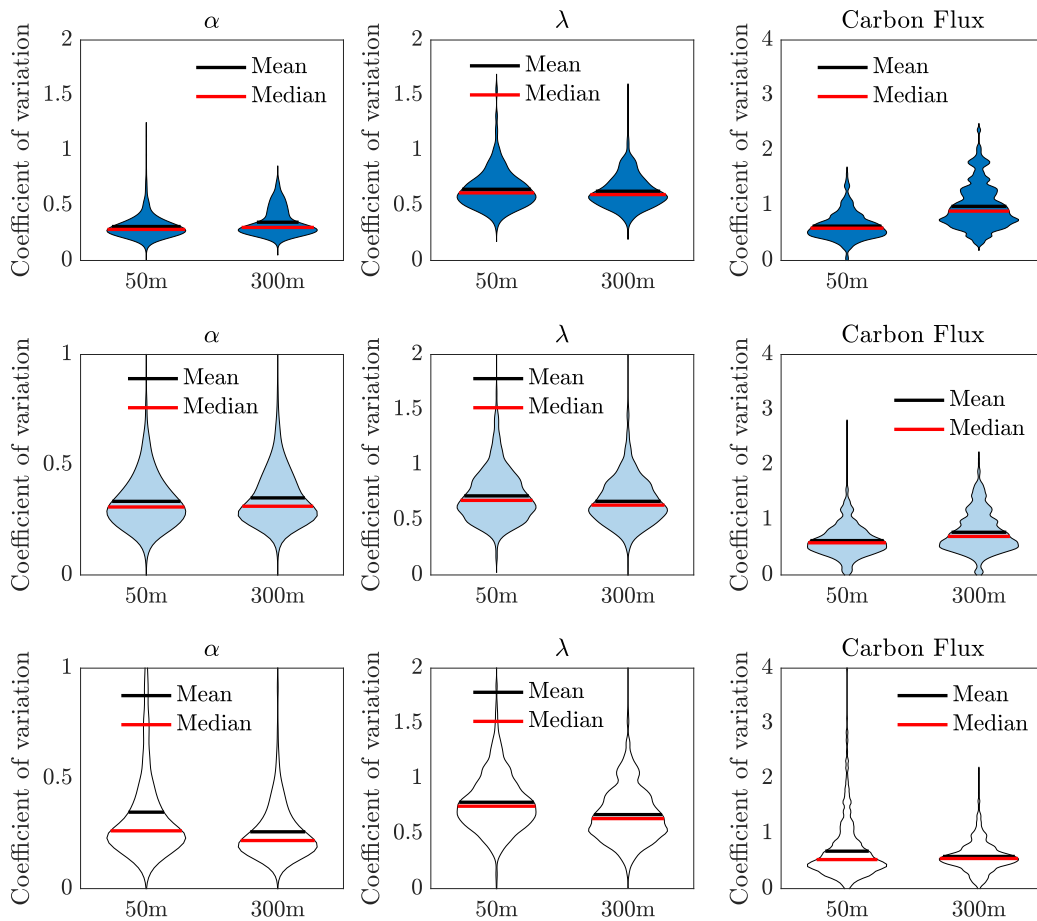
783 Supplementary Figure 1. Average values of alpha based on season, and at specific latitudinal

784 bands with depth. Note that 0-20 means -20 to 0 and 0 – 20.



785  
 786 Supplementary Figure 2. Average values of  $\lambda$  (normalized to  $\lambda_0 = 1$  mm) based on season, and at  
 787 specific latitudinal bands with depth. Note that 0-20 means 20S to 0 and 0 – 20N. Note average  
 788 values of lambda were performed on  $\log_{10}(\lambda)$  and are converted back to non  $\log_{10}$  transformed  
 789 values.  
 790  
 791

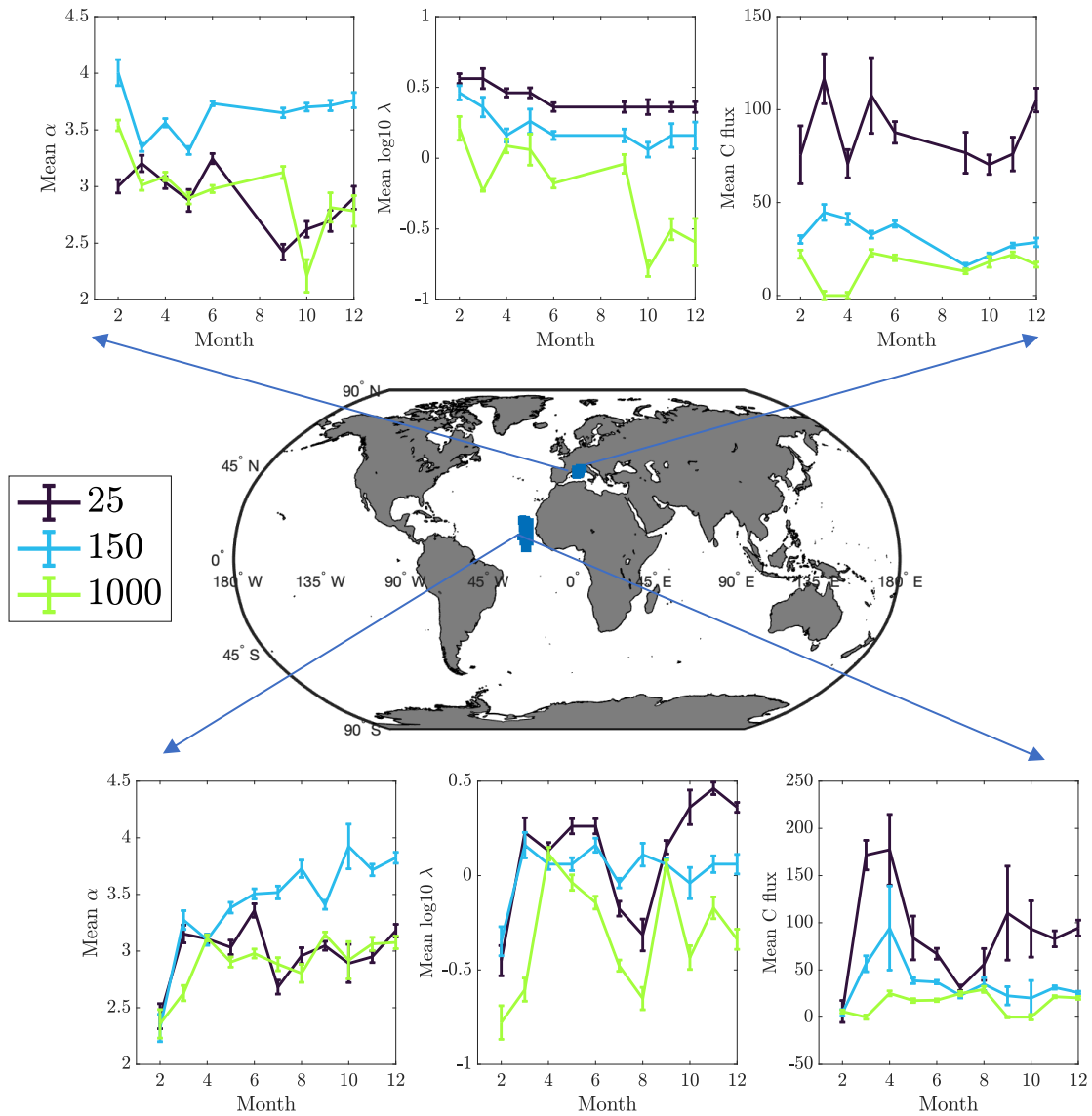
792  
793



794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806

Supplementary Figure 3. Violin plots for coefficient of variation in  $\alpha$  (left),  $\lambda$  (middle) and carbon flux based on either 50m or 300m for three locations. The top row (dark blue violins) is the P16 line in the Pacific. The middle row (light blue) is the Mediterranean. The third row (white) is the Arctic.

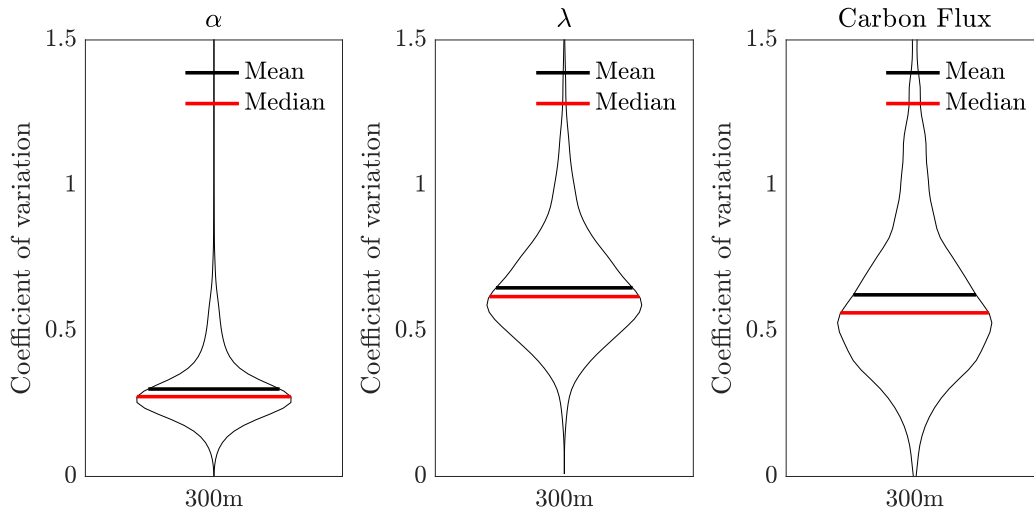
807  
808



809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822

Supplementary Figure 4. Seasonal transitions of alpha, lambda, and carbon flux (calculated with ‘Kiko’ parameters,  $\text{mg C m}^{-2} \text{d}^{-1}$ ) at the 2 locations worldwide that have sufficient observations over a 5x5 degree grid annually (i.e., must have at least 100 observations annually and at least 8 months of casts). Error bars represent the standard error for each month at different depths (25m in black, 150m in blue, and 1000m in green).

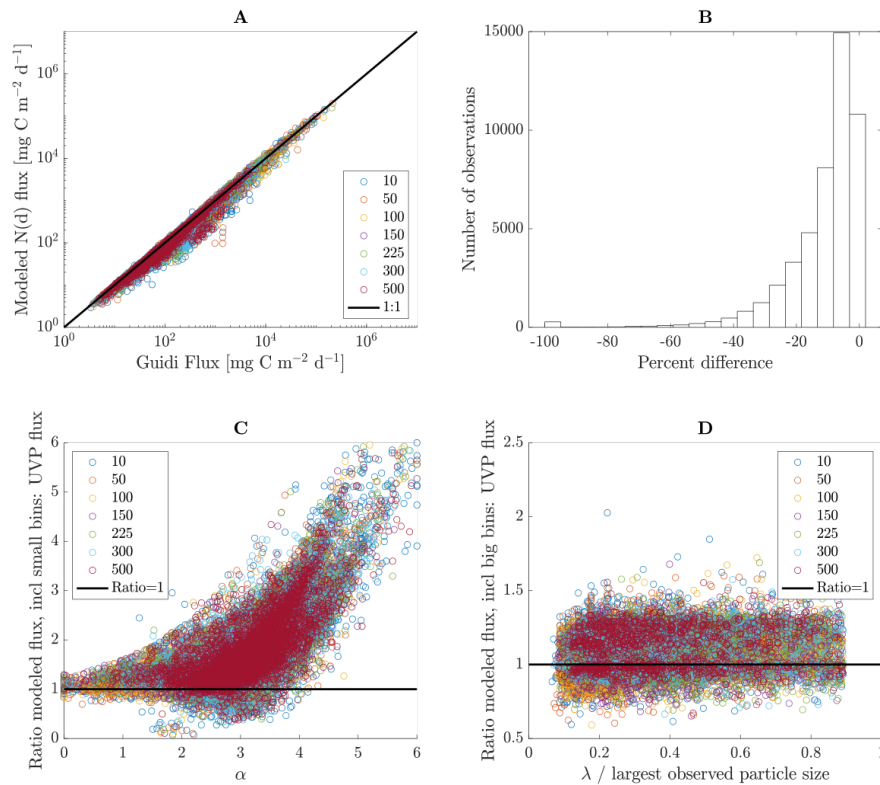
823



824

825

826 Supplementary Figure 5. Violin plots for coefficient of variation in  $\alpha$  (left),  $\lambda$  (middle) and  
 827 carbon flux based on 300m (resulting from the bootstrap procedure where sample volume was  
 828 doubled as a test case). The coefficient of variation reported in this figure is due only to the  
 829 sampling volume uncertainty and not natural spatiotemporal variations worldwide.



830

831 Supplementary Figure 6. A. Flux comparisons between modeled N(d) flux (y axis) and N(d) flux  
 832 from UVP observations ('Guidi flux') over the same particle size range. B. Histogram of percent  
 833 difference (relative to 'Guidi flux') between all flux determinations in A. C. The ratio of



834 modeled  $N(d)$  flux (including 2 size bins smaller than the observed size range) to observed  $N(d)$   
835 flux as a function of depth (colors dots) and alpha value. Black line represents equivalent fluxes.  
836 D. The ratio of modeled  $N(d)$  flux (including 2 size bins larger than the observed size range) to  
837 observed  $N(d)$  flux as a function of depth (colors dots) and lambda value relative to largest  
838 observed particle size. Black line represents equivalent fluxes.

839

840

841

842

843