MIT/WHOI

**Massachusetts Institute of Technology
Woods Hole Oceanographic Institution**

**Joint Program
in Oceanography/
Applied Ocean Science
and Engineering**

DOCTORAL DISSERTATION

Distributional models of ocean carbon export

by

Brendan Barry

February 2019

Distributional models of ocean carbon export

by

Brendan Barry

Submitted to the Joint Program in Physical Oceanography
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

February 2019

©2019 Brendan Barry.

All rights reserved.

The author hereby grants to MIT and WHOI permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document in whole or in
part in any medium now known or hereafter created.

Author

~~Joint Program in Physical Oceanography~~
~~Massachusetts Institute of Technology~~
~~& Woods Hole Oceanographic Institution~~
December 12, 2018

Certified by

Michael J. Follows
Professor of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Thesis Supervisor

Accepted by

Larry Pratt
Chairman, Joint Committee for Physical Oceanography
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution

Distributional models of ocean carbon export

by

Brendan Barry

Submitted to the Joint Program in Physical Oceanography
Massachusetts Institute of Technology
& Woods Hole Oceanographic Institution
on December 12, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Each year, surface ocean ecosystems export sinking particles containing gigatons of carbon into the ocean’s interior. This particle flux connects the entire ocean microbiome and constitutes a fundamental aspect of marine microbial ecology and biogeochemical cycles. Particle flux is also variable and intricately complex, impeding its mechanistic or quantitative description. In this thesis we pair compilations of available data with novel mathematical models to explore the relationships between particle flux and other key variables – temperature, net primary production, and depth. Particular use is made of (probability) distributional descriptions of quantities that are known to vary appreciably. First, using established thermodynamic dependencies for primary production and respiration, a simple mechanistic model is developed relating export efficiency (i.e. the fraction of primary production that is exported out of the surface ocean via particle flux) to temperature. The model accounts for the observed variability in export efficiency due to temperature without idealizing out the remaining variability that evinces particle flux’s complexity. This model is then used to estimate the metabolically-driven change in average export efficiency over the era of long-term global sea surface temperature records, and it is shown that the underlying mechanism may help explain glacial-interglacial atmospheric carbon dioxide drawdown. The relationship between particle flux and net primary production is then explored. Given that these are inextricable but highly variable and measured on different effective scales, it is hypothesized that a quantitative relationship emerges between collections of the two measurements – i.e. that they can be related not measurement-by-measurement but rather via their probability distributions. It is shown that on large spatial or temporal scales both are consistent with lognormal distributions, as expected if each is considered as the collective result of many subprocesses. A relationship is then derived between the log-moments of their distributions and agreement is found between independent estimates of this relationship, suggesting that upper ocean particle flux is predictable from net primary production on large spatiotemporal scales. Finally, the attenuation of particle flux with depth is explored. It is shown that while several particle flux-versus-depth models capture observations equivalently, these carry very different implications mechanistically and for magnitudes of export out of the surface ocean. A model is then proposed for this relationship that accounts for measurements of both the flux profile and of the settling velocity distribution of particulate matter, and is thus more consistent with and constrained by empirical knowledge. Possible future applications of these models are discussed, as well as how they could be tested and/or constrained observationally.

Thesis Supervisor: Michael J. Follows
Title: Professor of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology

Acknowledgments

Despite the premise that a thesis represents individual effort, to get here has taken a large and diverse village, each member of which I am grateful for and indebted to, for support, assistance, patience, encouragement, enthusiasm, challenge, trust, and all other phenomena that comprise life and inquiry. I thank my family, friends, labmates, classmates, and colleagues – those within WHOI and MIT, at the University of Chicago, and those that make up my larger scientific community – for everything. Any and all of you, if you are reading this: thank you for everything. I thank my thesis committee – Ken Buesseler, Raf Ferrari, Melissa Omand, and Dan Rothman – and my thesis chair, Andrew Babbín, for their perspectives, conversations, and precious time. I thank the many people who continue to make possible all of what goes on at MIT, WHOI, and in the Joint Program. I thank Geoff Lynn for designing Figure 1-1. I thank the Massachusetts Institute of Technology, the National Science Foundation, and the National Aeronautics and Space Administration for financial support. I thank Chris Follett for being a tremendous and generous font of wisdom, both scientific and otherwise, from whom I have learned so much. It would be remiss of me not to extend a particularly heartfelt thanks to Kelsey Bisson for every moment of our friendship and colleagueship, which has always been nurturing, inspiring, and joyous. Finally, it is a deep and true pleasure, which even as I write this elicits tears, to thank my advisor, mentor, and friend, Mick Follows, for putting up with me over these years – and for always being exactly what I needed you to be.

Contents

1	Introduction	19
2	Temperature and export efficiency	33
2.1	On the temperature dependence of oceanic export efficiency	33
2.1.1	Overview	33
2.1.2	Introduction	34
2.1.3	Temperature vs. export efficiency: Model and Data	35
2.1.4	Temperature shifts measurements' distribution	38
2.1.5	Discussion	42
2.2	How have recent temperature changes affected the efficiency of ocean biological carbon export?	43
2.2.1	Overview	43
2.2.2	Introduction	43
2.2.3	A metabolic model of export efficiency	44
2.2.4	Global estimates of multidecadal change in export efficiency	46
2.2.5	Conclusion	50
2.2.6	Sensitivity of atmospheric CO ₂ to export efficiency at equilibrium . . .	51
2.2.7	Statistical details and sensitivity	53
3	Primary production and export	57
3.1	Can Rates of Ocean Primary Production and Biological Carbon Export Be Related Through Their Probability Distributions?	57
3.1.1	Overview	57
3.1.2	Introduction	58
3.1.3	Theory: the lognormal distribution as a null model	62

3.1.4	Theory: relating export and production by their distributions' moments	64
3.1.5	Compilation of production measurements	65
3.1.6	Subregions of production measurements	66
3.1.7	Compilation of export measurements	68
3.1.8	Statistical methods	69
3.1.9	Production and export are log normally distributed	70
3.1.10	Log-Moments of subregions' production distributions sort predictably .	70
3.1.11	Export scales sublinearly with production	74
3.1.12	What Are the uses and requirements of a more complete theory? . . .	77
3.1.13	Can the probabilistic approach developed here be used to test models?	79
3.1.14	What about the spatiotemporal structure of the surface ocean?	80
3.1.15	Conclusion	81
3.1.16	Re-examining efficiency-production relationships	81
3.1.17	Convergence of the Central Limit Theorem	85
3.1.18	Derivation of log-moment relationships	86
3.1.19	Lower cutoff for production measurements	87
3.1.20	Kuiper's statistic and other cumulative distribution function statistics	89
4	Depth and particulate organic carbon flux	91
4.1	Particle flux parameterizations: quantitative and mechanistic similarities and differences	91
4.1.1	Overview	91
4.1.2	Introduction	92
4.1.3	Several mechanistic particle flux attenuation models	93
4.1.4	Large differences in export estimates	94
4.1.5	Quantitative indistinguishability	97
4.1.6	Conclusion	98
4.1.7	Derivations	99
4.1.8	Flux profiles	104
4.1.9	Euphotic and mixed layer depth climatologies	104
4.1.10	Normalizations	104
4.1.11	Statistical routine	106

4.2	Distributional model for particle flux attenuation	112
4.2.1	Overview	112
4.2.2	Particle flux complexity	113
4.2.3	Settling velocity distribution data	115
4.2.4	Distributional model	118
4.2.5	Fit to settling velocity distribution data	121
4.2.6	Fit to flux-depth data	122
4.2.7	Discussion	127
5	Summary and future directions	131

List of Figures

1-1	Illustration of the global carbon cycle.	21
1-2	Simplified illustration of the carbon cycle, taken from Figure S2 of [Cael et al., 2017].	22
1-3	Illustration of the pathways of the biological pump, taken from Figure 4 of [Siegel and Others, 2016].	26
2-1	Temperature sensitivity for phytoplankton and copepod growth rates, adapted from (a) Eppley [Eppley, 1972] and (b) Huntley and Lopez [Huntley and Lopez, 1992]. The temperature scaling prefactor is larger for the copepods, the key thermodynamic relation that drives this model. Another key difference is that copepod growth clusters along, while phytoplankton growth lies under, their respective curves; copepod growth appears to deviate significantly from this curve near $T = 24^{\circ}\text{C}$. These predictions from the above pictured original data sets have been additionally tested and verified in, e.g., Rose and Caron [Rose and Caron, 2007]. Note the x axes of each subplot cover different limits. The prefactor is the value which multiplies temperature in the expressions $\mu \propto e^{0.063T}$, $\lambda \propto e^{0.11T}$	36
2-2	Model prediction for maximum export efficiency as a function of temperature, compared with data from sources described in Methods. The four points enclosed by diamonds were those taken to fix parameter $\alpha = .24$. The model agrees well with data in its applicable temperature range – vertical dashed line indicates the cutoff at 24°C observable in Huntley and Lopez [Huntley and Lopez, 1992] and Rose and Caron [Rose and Caron, 2007]. Remaining dashed lines indicate possible extrapolations of the model prediction above 24°C	40

2-3	Mean and normalized variance of the data after being binned at width 1°C and rescaled by predicted maximum export efficiency. Binned values are best explained by a constant line; the best fit regression line computed with ordinary least squares has a negative adjusted r^2 ; thus, after the model's prediction for temperature dependence has been removed, no variation in the binned values can be explained by temperature.	41
2-4	Percent change in $\langle ef \rangle$ after a temperature change, as estimated by MM (see Eq. (2.9)), as a function of the initial and final temperatures (T_{old} and T_{new}). Contours are spaced at 5%.	47
2-5	Percent change in $\langle ef \rangle$ from 1982-2014, as estimated by MM, as a function of latitude, for different data products. Color corresponds to SST product and line type to $\mathcal{P}$ product. Black curve is the average across the SST- $\mathcal{P}$ pairs. . .	50
2-6	a-d) As in Figure 2-4, except for α or β are changed by $\pm 10\%$ in each case, as noted below each subfigure. The overall structure of ef changes is similar, but the magnitude of ef changes for a given temperature change increases with both α and β	54
3-1	Scatterplot of particulate organic carbon flux (f) at 100m and depth-integrated net primary production by phytoplankton ($\mathcal{P}$) data from [Maiti et al., 2013]. Notice the lack of a clear correlation between the two. Metrics reported are from Pearson's correlation: The fraction of variance of explained (R^2) and the p value of the correlation. Compare to Figure 3-10.	59
3-2	Schematic of approach: p and f data from the global ocean are grouped into probability distributions and the probability distributions are then analyzed. Production and flux are then related to each other (and to other quantities) via the moments of their probability distributions.	61

- 3-3 Top: Locations of p samples $> 0.5 \text{ } \mu\text{g C L}^{-1} \text{ day}^{-1}$ from the compilation by [Buitenhuis et al., 2013]; compare to their Figure 2a. The solid black boundary line denotes the separation of data into ocean basins (Section 3.1.6). The time series are indicated by crosses. Middle: Locations of f samples $> 3.6 \text{ mg C m}^{-2} \text{ day}^{-1}$ from the compilation. The time series are indicated by crosses. Bottom: Map of biomes into which p and f data are subregioned; compare to Figure 1 of [Banse, 1992]. Biomes are defined as B1: low-seasonality, nutrient-depleted (oligotrophic), B2: low-seasonality, nutrient-replete (eutrophic), B3: high-seasonality (seasonal), using a depth-integrated net primary production climatology from the Carbon-based Productivity Model [Westberry et al., 2008]. 67
- 3-4 Top: PDF of global p data versus lognormal fit. $n = 38,334$ refers to the number of samples included in the PDF, which excludes data from HOT, BATS, CARIACO, as well as data below the 10% of peak height threshold $< 0.5 \text{ } \mu\text{g C L}^{-1} \text{ day}^{-1}$. Green curve is the probability density function of the lognormal which minimizes the Kuiper statistic as compared to the data. Bottom: PDF of global f data versus lognormal fit. $n = 1,033$ refers to the number of samples included in the PDF, which excludes data from HOT, BATS, and CARIACO time series, as well as data below the 10% of peak height threshold of $3.6 \text{ mg C m}^{-2} \text{ day}^{-1}$. Green curve is the probability density function of the lognormal which minimizes the Kuiper statistic as compared to the data. HOT = Hawaii Ocean Time; BATS = Bermuda Atlantic Time series Study; CARIACO = CARbon Retention In A Colored Ocean time series; PDF = probability density function. 71
- 3-5 Probability density functions (PDFs) of p data for different subregions. In each case the green curves are the PDF of the lognormal which minimizes the Kuiper statistic as compared to the data. (μ, σ) is plotted in Figure 3-6. . . . 72
- 3-6 Log-moments for each of the p subregions. Error bars are 95% confidence intervals estimated from bootstrapping; contours are lines of constant mean for a lognormal distribution, equal to $\exp(\mu + \frac{1}{2}\sigma^2)$ 73
- 3-7 Histograms of time series' p and f data and biomes' f data. (μ, σ) is plotted in Figure 3-8; p data for the biomes are plotted in Figure 3-5. 76

3-8	μ_f versus μ_p for the three time series, the three biomes, and the global ocean. Dotted blue line is the estimated line from the regression on the biomes ($\mu_f = \alpha_b \mu_p + C_b$); dotted green line is the estimated line from the regression on the time series ($\mu_f = \alpha_t \mu_p + C_t$); parameter estimates from the regressions are reported in the top-left of the figure, along with their estimated standard errors. Solid black lines represent 95% bootstrap confidence intervals. The global distributions' log-means (yellow) are not included in the regressions that produce the dashed lines. Inset: σ_f plotted versus σ_p for the three time series, the three biomes, and the global ocean. Dashed line is the estimate $\sigma_f = \alpha_b \sigma_p$; grey shading corresponds to α_b 's standard error. None of the σ 's are included in the regressions that produce the dashed line.	78
3-9	Illustration of spurious relationships. a) scatterplot of x and y , two uncorrelated random variables sampled from a uniform distribution ($n = 200$). b) y/x plotted against x (dashed line is $1/x$; compare to Maiti et al. [Maiti et al., 2013]). c) X and Y are binned averages of x and y , into five bins; dashed line is two-parameter fit ($Y = 0.45; X^{-0.08}; r^2 = 0.99$); compare to Laws et al. [Laws et al., 2000a]. $1/8$ random sample sets will have a monotonic relationship such as the one seen here if binned into five bins ($2^{-4} = 1/16$ will be monotonically increasing and $1/16$ will be monotonically decreasing; this number increases rapidly even in response to very weak correlations, e.g. $\rho = 0.01$).	83
3-10	$\mathcal{P}$ and ef data from Laws et al. [Laws et al., 2000a], originally compiled by Dunne et al. [Dunne et al., 2005]. Metrics reported are from Pearson's correlation: the fraction of variance of explained (R^2) and the P-value of the correlation.	84
3-11	PDF of sums of triplets drawn from the unit uniform distribution, with normal distribution overlaid, illustrating the rapid convergence of the Central Limit Theorem.	85

3-12	Identification of the lower threshold for p samples considered. a) Bars are PDF of measurements of $p > 0$; dashed black line (for both panels) is the threshold identified at $0.5 \mu\text{g C L}^{-1} \text{ d}^{-1}$, or $\sim 10\%$ of the peak in the log-transformed distribution. Green line is a lognormal probability distribution function above the threshold. b) Bars are the relative residual error of the lognormal distribution shown in panel (a).	88
4-1	A: Seven models from the text fit to mesopelagic data from Figure 5 of Martin et al. [Martin et al., 1987b] (their ‘Open Ocean Composite’) and extended upwards to 50m depth. Goodness-of-fits are very similar, but the power-law model overestimates particulate organic carbon (POC) flux at 50m by 132% relative to the exponential model. Inset shows attenuation ($\frac{1}{f} \frac{df}{dz}$) differences are even more pronounced. Attenuation for the exponential model coincides with most others around $\sim 400\text{m}$. B: Maximum tolerable root-mean-square error (RMSE) to distinguish models based on data from (A) with 90% confidence, estimated using a bootstrap method (see text for details). Row corresponds to ‘true’ model (from which data are assumed to be generated) and column corresponds to ‘false’ model (to be rejected). RMSEs are small relative to those in (A), indicating differences in goodness-of-fits between the models are not statistically significant. C: Ratio of power-law-model- and exponential-model-estimated POC flux at euphotic layer and mixed layer depths from the profiles in the data compilation as a function of extrapolation distance. Grey shading shows nominal 25% measurement uncertainty. Note the y -axis is logarithmic. The power-law model systematically and substantially overestimates relative to the exponential model (as well as the other models – see Section 4.1.10). For 29 profiles, $f_p(\text{MLD})/f_e(\text{MLD}) > 10$ (not shown).)	95
4-2	A map of data locations with histograms of data properties. POC flux is reported in $\text{mg C m}^{-2} \text{ d}^{-1}$	103

4-3	Cumulative distribution function of fixed-parameter vs. fit-parameter normalizations. Horizontal lines indicate the fraction of profiles in each case where the fixed-parameter normalization underestimates relative to the fit-parameter normalization. That these are all close to 0.5 indicates that the approximations of 500 m for ℓ and 0.7 for b introduce negligible bias.	107
4-4	a) Illustration of the statistical routine. b) Percent bootstrap confidences in rejecting the ‘false’ (column) model in favor of the ‘true’ (row) one, from applying the statistical routine illustrated (a) to an artificial profile with observations at $z = [125, 250, 500, 1000, 2000, 4000]$ m and 20% measurement variability. Initials are first letter of each model, except I refers to ExpInt model.	110
4-5	Schematic of an IRSC sediment trap in settling velocity mode; figure adapted from [Peterson et al., 2005]. Settling particles are collected by the intended rotating sphere (1), which then rotates 180° to dump material into the skewed funnel (2), which is then separated by settling velocity via successive rotations of the sample carousel (3).	116
4-6	Empirical cumulative distribution functions (CDFs) for the inverse of the settling velocity (m or $1/w$) with respect to POC flux from deployments of modified indented rotating sphere carousel sediment traps at three oceanographic stations at different times and depths. Compare to Figure 4 from [Trull et al., 2008]. Note that the ALOHA CDF only has five values because collection cups had to be aggregated due to low total flux [Trull et al., 2008]. Viewed right-to-left these distributions are also complementary CDFs for w	117
4-7	Fits of exponential, power-law, and reciprocal distributions (corresponding to rational, power-law, and ExpInt models) to one CDF from Figure 4-6. v is Kuiper’s statistic.	124
4-8	a) Probability density functions (PDFs) of r^2 values for fits of the power-law, rational, exponential, and ExpInt models to the 722 profiles in the database described in Section 4.1. b,c) PDFs of F values for fits of the ballast (double) and ExpInt models to the profiles with $n \geq 4$ ($n \geq 5$) measurements – total of 277 (187) profiles. In each case, PDFs are estimated using the kernel method [Hill, 1985].	126

List of Tables

- 2.1 Percent change in $\langle\langle ef \rangle\rangle$, during the period 1982-2014, as estimated by MM, for different data products. All changes suggest a decrease in $\langle\langle ef \rangle\rangle$ and are statistically significant ($p < .001$; see Section 2.2.7). Mean and standard deviation of percent change across each SST- $\mathcal{P}$ pair are $-1.5 \pm 0.4\%$ 49

- 3.1 Summary of analyses of Dunne et al. [Dunne et al., 2005] data. First column is the model equations used to predict ef from $\mathcal{P}$ and/or T ; equations are those considered by Laws et al. [Laws et al., 2000a] or simplifications thereof. Second column indicates whether the parameters in the equation were fit to the data (free) or if their values were taken from Laws et al. [Laws et al., 2000a] (fixed). Third column indicates which variable(s) ef is a function of in each case. Fourth column indicates the percentage of variation in ef explained by the independent variables via each equation. 82

- 4.1 Fraction of profile fits for for which f_e significantly overestimates or f_p significantly underestimates export relative to other models, using a nominal measurement uncertainty of 25%. For instance, f_e is less than or within 25% of f_r for 97% of profiles, so only significantly overestimates f_r 3% of the time. The value reported is the larger value between normalizations to ELD and to MLD, e.g. $f_e(\text{MLD})$ significantly overestimates $f_r(\text{MLD})$ 2% of the time while $f_e(\text{ELD})$ significantly overestimates $f_r(\text{ELD})$ 3% of the time, so the ELD value is reported. We use 4/5 rather than 3/4 in the third column because 4/5 is the same *relative* distance from 1 as 5/4; if 3/4 was used instead the values in the table would be the same. 106

4.2	Number of profiles (out of a total of 722) for which the percent bootstrap confidence in rejecting the ‘false’ (column) model in favor of the ‘true’ (row) one is greater than 90%, assuming a measurement uncertainty of 20%. Initials are first letter of each model, except <i>i</i> refers to the ExpInt model.	111
4.3	Same as Figure S3b except $n = 20$	112
4.4	Same as Table 4.3 except $\max(z) = 1\text{km}$	112
4.5	Models of particle flux attenuation and their inverse Laplace transform, i.e. the distribution which when substituted into Eq. (4.30) yields that model. . .	120
4.6	Results of fitting the distributions corresponding to the power-law, rational, and ExpInt models to each CDF from Figure 4-6, using all three CDF statistics. Values in parentheses are the parameters for the best-fit distributions: b for the power-law model, ζ [m] for the rational model, the, and $\min(w)$ [m/d] for the ExpInt (free $\min(w)$) model.	123

Chapter 1

Introduction

Numerous elements play indispensable roles in living systems, their environments, and the interactions between these; arguably none plays a more vital role than carbon, the celebrated sixth element in the periodic table. The abundance of carbon in Earth’s crust and surface, along with its unique ability to form diverse compounds and polymers at the temperature ranges occurring there, underlie its function as the basis of living systems [Smith and Morowitz, 2016], constituting roughly half of all dry biomass [Schlesinger, 1991, Houghton, 2003] and making it the standard unit of measure for living material [Bar-On et al., 2018]. Carbon also features prominently in discussions of Earth’s climate; carbon is the central atom in carbon dioxide, methane, and other greenhouse gases (e.g. fluorocarbons) and therefore the partitioning of carbon between atmosphere, ocean, and land is of extreme importance to the state of Earth’s climate, both today and throughout Earth’s history [IPCC, 2014, Williams and Follows, 2011]. The myriad physical, chemical, and biological processes that transform and transport carbon through different states and places in the Earth system are collectively termed the carbon cycle. Furthermore, as living organisms mediate many of these transformations via their metabolisms, the carbon cycle is a principal conduit by which life influences and feeds back upon its environment [Falkowski et al., 2008]. Any comprehensive understanding of the biosphere or the climate thus requires an understanding of the carbon cycle, as is the broad motivation of this thesis. Here a particular oceanic branch of this cycle, strongly linked to both climate and microbial ecology, forms the object of study. The aim of the present chapter is to contextualize this branch within the global carbon cycle and the investigations described in subsequent chapters.

Figure 1-1 presents an illustration of the carbon cycle. Both its global nature and the interweaving of biotic and abiotic components are evident. While the carbon cycle is a consummate complex system, and therefore will be grossly oversimplified by any such representation, a few key features of interest for our present purposes can be gleaned. 1) The ocean – particularly the deep ocean – is the dominant reservoir of surficial carbon (i.e. carbon in the surface Earth) [Sarmiento, 2013, Williams and Follows, 2011, Schlesinger, 1991]. Even the organic carbon content of the deep ocean, a small fraction of the total carbon in the deep ocean, is comparable to the atmosphere’s total carbon content. The total oceanic carbon content is an order of magnitude larger than that of all other reservoirs combined. This underscores, at least on long timescales, the dominance of processes controlling the carbon transformations within the deep ocean and the exchanges of carbon between the deep ocean and other components of the system. 2) The surface ocean mediates the exchange of carbon between the deep ocean and the atmosphere (and thus the terrestrial components of the carbon cycle as well) [Sarmiento, 2013, Williams and Follows, 2011, Schlesinger, 1991]. One can also see that the net flux between the surface ocean and the atmosphere is small relative to the gross fluxes from either to the other, such that the surface ocean and atmosphere are nearly in equilibrium in terms of carbon storage. This further underscores the central importance of the vertical distribution of carbon in the ocean¹. 3) Despite comprising a relatively negligible reservoir of carbon, marine biota drive a carbon flux from the surface to the deep ocean on the order of the flux of carbon into the atmosphere due to fossil fuel burning and cement production [Sarmiento, 2013, Williams and Follows, 2011, Schlesinger, 1991]. (Note that this flux is largely balanced by a net upwards exchange of carbon from the deep ocean into the surface ocean, which results from higher carbon concentrations in deep ocean water.) This underscores the importance of this downward flux of carbon due to marine biota in controlling the vertical distribution of carbon in the ocean. Altogether, then, this illustration suggests that this flux has an appreciable influence on the global carbon cycle.

These observations are encapsulated in a further idealized depiction of the carbon cycle in Figure 1-2. One may neglect terrestrial fluxes and stores, and consider the partitioning of carbon between the atmosphere, surface ocean, and deep ocean. Carbon is exchanged across

¹While one might describe this more coarsely, but in better keeping with Figure 1-1, as the partitioning of carbon between the surface and deep ocean, the above description is less preferable because the distinction between ‘surface’ and ‘deep’ ocean is a useful but unavoidably inexact construction.

Figure 1-1: Illustration of the global carbon cycle.

the ocean-atmosphere interface and the surface ocean equilibrates with the atmosphere. The deep and surface ocean exchange carbon via ocean circulation and diffusion, and there is also a biological ‘export’ of carbon from the surface ocean into the deep [Cael et al., 2017, Ito and Follows, 2005, Williams and Follows, 2011, Goodwin et al., 2008]. The deep ocean carbon concentration is then the sum of a contribution from subducted surface waters and a biologically mediated accumulation. The biologically mediated pool is here assumed to represent the accumulation due to the same marine-biota-driven flux discussed above. Thus in this simple depiction, this flux maintains a vertical gradient of carbon in the ocean, thereby reducing the carbon content of all other reservoirs [Volk and Hoffert, 1985]. But what is this flux?

Though somewhat of a misnomer, for historical reasons the collection of processes that comprise this flux are referred to as the biological (carbon) pump. Volk and Hoffert [Volk and Hoffert, 1985] initially defined two² biotic ocean carbon pumps, i.e. processes that deplete the ocean surface of carbon relative to deep waters: biological fluxes of i) organic carbon and ii) CaCO_3 . Far more carbon is fluxed in organic form, though individual estimates of these pumps’ magnitudes vary and also depend greatly on the definition of the interface

²Volk and Hoffert [Volk and Hoffert, 1985] also defined a third, abiotic solubility pump resulting from increased CO_2 solubility in cold downwelling water, that accounts for about a third of the difference in surface vs. deep dissolved inorganic carbon.

Figure 1-2: Simplified illustration of the carbon cycle, taken from Figure S2 of [Cael et al., 2017].

between the surface and deep ocean [Siegel et al., 2014, Palevsky and Doney, 2018, Kwon et al., 2009]. The flux of organic carbon can in turn be decomposed further into contributing sub-processes, e.g. sinking of various classes of particles [Fowler and Knauer, 1986], vertical migration of zooplankton [Turner, 2015], or subduction of dissolved or suspended material [Omand et al., 2015]. Of these, particles are the largest vector of carbon transport to the deep sea overall [Buesseler and Boyd, 2009], though other processes can dominate in some situations. Whatever the level of resolution considered, in aggregate these biotic processes are the constituents of the biological pump, that flux which in large part maintains the ocean’s vertical gradient in carbon concentration.

The biological pump is of interest not only for its impact on this vertical gradient and therefore atmospheric carbon, but also as a fundamental feature of marine ecology and elemental cycling. The entry of carbon, energy, and other elements into the marine biosphere is concentrated in the thin lens that is the sunlit surface ocean, whereas the consumption of this mass and energy occurs throughout the entire ocean [Schlesinger, 1991, Sarmiento, 2013, Williams and Follows, 2011]. The downward flux therefore transports organic matter from source to sink, connects the global ocean microbiome, and supplies the metabolism of the largest habitat on Earth: the deep sea. The workings of the biological pump are therefore of great interest from several angles.

It comes then as no surprise that the biological pump is the subject of a substantial literature. A number of modeling studies explore the structure, function, and sensitivity of the biological pump; these studies also confirm its global significance, corroborating that the ocean exerts a dominant control on atmospheric carbon and that the biological pump is an essential feature of the marine carbon cycle and microbiome [Kwon et al., 2009, Ito and Follows, 2005, Sarmiento and Toggweiler, 1984, Siegenthaler and Wenk, 1984, Marinov et al., 2008, Cox et al., 2000, Queré and Others, 2007]. The biological pump has also been of sustained observational interest which continues today (e.g. the EXPORTS field campaign [Siegel and Others, 2016]), using as broad an array of measurement techniques and experimental designs as any problem in oceanography.

From these efforts a general portrait of the biological pump can be deduced. In the surface ocean, phytoplankton fix dissolved carbon into organic compounds via photosynthesis. Globally, this occurs at a rate of roughly ~ 60 GtC/yr [Buitenhuis et al., 2013] after subtracting autotrophic respiration, though the rate of net carbon fixation (termed

net primary production, NPP) both fluctuates and varies greatly over time and space [Cael et al., 2018]. Much of this carbon is respired and thus returned to its dissolved form by various heterotrophic microorganisms in the surface ocean such as bacteria and zooplankton, but a fraction leaves, or is ‘exported’ from, the surface ocean. By ‘a fraction’ it is meant something roughly an order of magnitude less than total NPP – the most commonly quoted range of estimates at present is 5-12 GtC/yr [Siegel et al., 2014], though this range depends on the surface through which the flux is defined and one can readily find estimates <5 or >12 GtC/yr – while the local fraction of NPP exported (often termed the ‘export efficiency’) varies substantially on local spatiotemporal scales, anywhere from close to 0 to close to 1 [Cael and Follows, 2016]. By ‘exported’ it is meant fluxed downward from the surface ocean either in the form of ‘marine snow’ aggregates of detrital material, as living or minimally degraded whole phytoplankton, as zooplankton fecal pellets, as physically subducted suspended material, as vertically migrating organisms, or as some other distinct pathway. By ‘the surface ocean’ it is often meant either the euphotic layer – the portion of the ocean receiving enough sunlit to sustain net photosynthetic activity, the mixed layer – the upper ocean boundary layer that is considered vertically homogenized by turbulence, or a fixed depth, commonly 100m; in practice all of the above are operationally defined to some extent [Palevsky and Doney, 2018]. The carbon exported is primarily particulate and organic (particulate organic carbon, POC), and is accompanied by particulate organic nitrogen, phosphorous, and various other materials such as biogenic silica or calcium carbonate [Schneider et al., 2003, Klaas and Archer, 2002]. The efficiency of the biological pump is closely related to the fraction of nutrients that are used to fuel NPP (and therefore associated with carbon) before being subducted; the more nutrients are ‘preformed’ (i.e. are subducted into the ocean interior without having fueled primary production in the surface ocean), the less efficient the biological pump [Ito and Follows, 2005]. The relatively larger nutrient utilization in high latitudes results in increased carbon storage and higher export efficiency, suggesting a disproportionate global role for high latitude oceans relative to their small areal extent [Sarmiento and Toggweiler, 1984]. Below the surface ocean, the vertical flux attenuates with depth, as microorganisms both attached to particles and within the water column consume, degrade, break apart, and in general solubilize or remineralize sinking particulate material. Depending on the depth at which sinking material is returned to the water column, it can remain sequestered from the surface ocean from months to millennia

[Kwon et al., 2009, Siegel and Others, 2016]; the climatic impact of the biological pump is thus very sensitive to the depths at which sinking material is remineralized. Various interdependencies between the atmospheric carbon concentration and the biological pump suggest in the net a complex climate-biosphere feedback. Many of these essential features are captured in Figure 1-3.

One can rightly say that the above portrait is rather uncertain and qualitative. Indeed, much of the fundamentals of the biological pump remain poorly understood; neither its global magnitude, nor its structure or variability over space and time, nor its attenuation with depth are well-characterized. Even less established are the relative importance of different pathways, how the dominant mechanisms controlling these pathways work, or how these are likely to change/have changed in the future/past. As an example: what governs how much POC makes it into the bathypelagic? Is this ‘refractory’ POC that is either not palatable or only very slowly degraded by microorganisms [Jiao and Others, 2010]? Or is this POC labile but physically protected from degradation by associated ‘ballast’ minerals [Armstrong et al., 2001]? Or does POC degradation simply proceed more slowly with increasing depth due to the much lower temperatures experienced by microbes in the deep sea [López-Urrutia et al., 2006]? Or is this POC transported by very fast-sinking particles that give microbes little time to act [Fowler and Knauer, 1986]? Likely all of these processes can and do influence how much POC is fluxed into the bathypelagic, more or less at different places and times, but in what combination, when and where, and what determines this?

The biological pump’s persistent mysteriousness can largely be ascribed to two factors: its complexity, and the difficulty of its measurement. To quote from Section 4.2:

In the ocean, particle flux is nothing if not complex and heterogeneous. Total fluxes comprise myriad sinking particles which are themselves the complex result of myriad biological, chemical, and physical factors [de la Rocha and Passow, 2007]. These particles’ physical properties vary widely, from their size and shape to their density and porosity [Kajihara, 1971, Alldredge and Gotschalk, 1988]. Individual particles themselves are comprised of a complex suite of organic and inorganic materials, which may represent a continuum ranging from highly labile to utterly refractory [Hedges et al., 2001, Dittmar, 2008, Jiao and Others, 2010]. Beyond this, the remineralization of particulate material is plausibly related to properties of the ambient fluid like temperature or oxygen concentration [Mooy

Figure 1-3: Illustration of the pathways of the biological pump, taken from Figure 4 of [Siegel and Others, 2016].

et al., 2002], may occur via consumption by different bacteria using different metabolic/chemical pathways even within a single particle [Bianchi et al., 2018], may occur at different rates due to the physical microstructure of the particle [Rothman and Forney, 2007], and may in turn alter the shape and other physical properties of the particle [Mayor et al., 2014]. Settling velocity is also related to the density of the ambient fluid, which of course changes with depth due to stratification [MacIntyre et al., 1995]. Individual particles may also interact with each other through myriad aggregation and disaggregation events [Burd and Jackson, 2009, Jackson and Burd, 1998, Alldredge et al., 1990], may be consumed or otherwise altered by free-swimming zooplankton [Steinberg et al., 2008, Kiørboe, 2000, Steinberg, 1995], etc.

Beyond this, neither zooplankton migration nor the submesoscale turbulence responsible for most physical subduction are any simpler – both are active areas of research and their general impacts are beyond robust quantification at present. One can readily see the difficulties in accounting for all of these processes rigorously, even in a highly controlled setting. Further complicating this is that POC flux is no simple thing to measure, often involving extended deployment of sediment traps that collect sinking material or inferred indirectly from chemical radiotracers that estimate POC flux on different effective timescales than the measurements of related processes and are subject to similarly large uncertainties as sediment traps [Buesseler, 1991, Buesseler and Others, 2000, Stanely et al., 2004, Buesseler and Others, 2007]. In other words, POC flux is a dynamic result of a large number of interactions between different organisms and with their physiochemical environment, which varies with space and time even on small scales and can only be measured somewhat accurately, usually by methods that are expensive and intensive. Where these measurements are made, the number of potentially important covariates is substantial, beyond the ability of most observational campaigns to capture – and the measurement of many of these covariates themselves present significant challenges. Thus the biological pump is not an easy thing to measure or to model.

Even while a complete description of the biological pump might require a great deal more data and detailed understanding of all of these factors’ interdependencies, insight is the product of simplicity, and for many applications an approximate quantitative description of the biological pump and what controls its variations is sufficient. Certain relationships

between POC flux and other phenomena have been proposed that make mechanistic sense and are empirically supported; in cases where sufficient data are available, simple and general relationships can be deduced that provide valuable intuition about the biological pump (though these may not always provide satisfactory predictive power). These relationships tend to be with more ‘basic’ variables – e.g. the temperature of the surrounding fluid, rather than e.g. the mean fractal dimension of marine snow aggregates. A formative paper by Eppley and Peterson [Eppley and Peterson, 1979], building on the work of Dugdale and Goering [Dugdale and Goering, 1967], argued from a steady-state perspective that POC export should be approximated by new production (NPP fueled by nutrients externally injected into the euphotic layer, primarily by the upwelling of nutrient-rich deep waters), and that new production is an increasingly large fraction of total production as total production increases – therefore that export efficiency increases with productivity. Suess [Suess, 1980] looked further at the relationship between POC flux and NPP to include depth, and found an approximately inverse relationship between flux and depth. Laws et al [Laws et al., 2000b], examining a detailed numerical model of the planktonic ecosystem, concluded that export efficiency should also be determined by mixed layer temperature; a small self-consistent set of observations corroborated this interpretation, with both suggesting a negative linear relationship between export efficiency and temperature. Relationships with other variables or processes, e.g. different metrics of community structure [Guidi et al., 2009], have been proposed – these relationships are however more challenging to assess, quantify, and measure than these basic oceanographic variables of temperature, depth, and productivity.

Subsequent literature has built upon these initially posited relationships, incorporating over time the additional constraints of many more observations, and exploring the validity of different, sometimes contradictory, models to relate POC flux to these other variables. The simple correspondences between these and POC flux are not clearly supported by these expanded datasets, which instead reflect the highly variable nature of POC flux, e.g. [Maiti et al., 2013, Dunne et al., 2005, Henson et al., 2011, Bisson et al., 2018]. Thus it is necessary to revisit these relationships in this context and to develop new models that capture the relationship between POC flux and these other variables while also acknowledging and being consistent with this variability. Improving our understanding of how these basic variables relate to POC flux can help provide a general quantitative characterization of the biological pump despite its complexity and variability. Export efficiency may not be a simple linear

function of temperature, but the fundamental insight underlying Laws et al. [Laws et al., 2000b] that temperature should influence export efficiency is no less plausible – perhaps temperature is one of many influences, and the mechanistic and quantitative nature of this influence can still be extracted. Even if individual measurements of NPP and POC flux seem to suggest differing or very weak relationships, the mechanistic link between the two is fundamental – all of the sinking material that comprises POC must previously be fixed via NPP – so it should be possible to identify some quantitative relationship between the two. And there is plenty of room for progress in understanding quantitatively and mechanistically the depth-attenuation of POC flux that determines the fate of this exported material. As the scientific community invests ever-greater resources into the study of the biological pump, there is a need to compile and make sense of the data that are out there.

To summarize: the biological pump is an integral component of the carbon cycle, marine ecology, and climate. While the biological pump has been the subject of much investigation, because of its complexity and variability, along with the extremely difficult nature of its measurement, in its characterization there is much room for improvement. The sinking flux of particulate organic carbon is thought to be the dominant vector of transport for the biological pump. Key relationships between this flux and temperature, productivity, or depth have been identified; specifying this flux in terms of these variables yields a necessarily oversimplistic yet still valuable description. These relationships are less clear in the light of additional data and analyses thereof, suggesting a need to revisit these relationships and to find ways of describing them that are consistent with the variability exhibited by these data. By pairing compilations of extant data with novel mathematical models, this thesis attempts to address this need.

The modeling approach taken here is to incorporate variability and/or heterogeneity in the biological pump by describing certain quantities as random variables sampled from probability distributions – what we will refer to as a distributional approach. Rather than considering export efficiency as a function of temperature, i.e. $ef = ef(T)$, we consider export efficiency as a random variable sampled from a probability distribution that is temperature-dependent, i.e. $ef \sim P_{ef} = P_{ef}(T)$. This is not to suggest that export efficiency is actually a random process, but rather that a number of factors, of which temperature is one, influence export efficiency, and in the absence of information about these other processes, we may for simplicity assume that they influence export efficiency in a way that appears to us as ran-

dom. Rather than considering POC flux as a function of NPP, i.e. $F = F(\text{NPP})$, we consider POC flux as a random variable sampled from a probability distribution that is a function of the probability distribution that NPP is sampled from, i.e. $F \sim P_F = P_F(P_{\text{NPP}})$. Rather than considering a single settling velocity w for sinking particles, we consider particulate material to sink heterogeneously at varying speeds, collectively described by a distribution P_w . Though only minimally more sophisticated than standard statistical models for these relationships, this type of distributional approach allows us to make appreciable progress in capturing the observed relationships, without idealizing out the variability that evinces the system’s complexity. We apply this distributional perspective to each of the relationships discussed above in turn.

Chapter 2 addresses the relationship between export efficiency and temperature. A simple mechanistic-distributional model for the relationship between export efficiency and temperature is derived from considering the dynamics of a phytoplankton population in the mixed layer. The underlying mechanism is the differential temperature sensitivity of autotrophic and heterotrophic processes; as temperatures increase, the maximum rate of autotrophy increases, but the typical rate of heterotrophic processes increases even more so, as reflected in growth-temperature relationships of phytoplankton and zooplankton. The model predicts that the distribution of export efficiency measurements is scaled, or ‘squeezed,’ by the temperature of the mixed layer; the model extracts the temperature dependency of a collection of observations >2 orders of magnitude larger than the original dataset considered by Laws et al. [Laws et al., 2000b]. The model is then applied to global sea surface temperature records to estimate a decline in global average export efficiency of 1-2% over the past three decades due to this metabolic mechanism. Finally, incorporation of this temperature sensitivity into a simple box model framework suggests that this mechanism accounts for a ~ 7 ppm K^{-1} decrease in atmospheric $p\text{CO}_2$, which along with solubility pump changes explains the atmospheric carbon drawdown during the Last Glacial Maximum.

Chapter 3 addresses the relationship between export and net primary production. As described above, NPP and export are inextricable by mass conservation, but the quantitative relationship between the two is elusive. We hypothesize that a qualitative relationship should emerge between *collections* of the two measurements, because they are measured on different effective spatiotemporal scales. We therefore describe the basis of a theory for interpreting measurements of the two fluxes in terms of their probability distributions. We take as

a null model that both should be lognormally distributed on large scales, and show that compilations of measurements of each are consistent with this hypothesis. The compilation of NPP measurements is extensive enough to subregion by biome, basin, depth, or season; these subsets are also well described by lognormals, whose log-moments sort predictably. Informed by this robust lognormality we infer a statistical scaling relationship between the two fluxes. Two independent estimates of the relationship between the distributions' log-moments agree, illustrating the utility of a distributional approach to biogeochemical fluxes and suggesting that POC flux is predictable from NPP on large scales.

Chapter 4 addresses the relationship between POC flux and depth. Several idealized models for this relationship are derived mechanistically, and the mechanistic differences between these models are highlighted. A large quantitative difference between these models is then demonstrated; from the same POC flux measurements, these models estimate widely different magnitudes of export (i.e. POC flux normalized to a reference depth such as the euphotic or mixed layer depth) and remineralization lengthscales at shallow depths. In contrast, it is shown that these models are quantitatively indistinguishable from available measurements, given the variability and uncertainty of POC flux measurements and the flexibility and similarity of these models over the depths at which most POC flux measurements are made, indicating the need to leverage additional measurements to produce multiple constraints. Next we propose a mechanistic, distributional model for the depth-attenuation of marine particle flux, that takes into account measurements of both the flux profile and the settling velocity distribution of particulate matter. Settling velocity distribution measurements are best captured by a reciprocal distribution. This distribution then implies a two-parameter model for the flux-depth relationship if it is assumed that the bulk flux profile is the superposition individual particles' fluxes and controlled by the initial variation in settling velocity. This model also fits measured flux profiles equivalently to or better than standard models. Though necessarily oversimplistic, it is therefore unique in capturing observations of both flux vs. depth and of the settling velocity distribution, and thus constitutes a particle flux parameterization that is more consistent with and constrained by empirical knowledge. Furthermore, the emergence of the reciprocal distribution underlying the model is predictable as a result of the large number of processes that produce sinking particulate material. The model also generates multiple hypotheses that are testable in situ. Lastly, Chapter 5 summarizes the main findings of the preceding chapters and discusses

several possible avenues of future inquiry. These avenues include both theoretical – e.g. applying the model developed in Chapter 2 to glacial-interglacial atmospheric $p\text{CO}_2$ changes – and observational – e.g. testing the model developed in Chapter 4 *in situ* – research, ranging from specific and targeted to broad and general. In the end, a great deal more work is required to come to a rigorous understanding of the biological pump and its relationship to other oceanic phenomena such as those examined herein.

Chapter 2

Temperature and export efficiency

The work in this chapter is based upon the following publications:

Cael, B. B., and M. J. Follows. 2016. On the temperature dependence of oceanic export efficiency. *Geophysical Research Letters*. [Cael and Follows, 2016]

Cael, B. B., K. Bisson, and M. J. Follows. 2017. How have recent temperature changes affected the efficiency of ocean biological carbon export? *Limnology and Oceanography: Letters*. [Cael et al., 2017]

2.1 On the temperature dependence of oceanic export efficiency

2.1.1 Overview

Quantifying the fraction of primary production exported from the euphotic layer (termed the export efficiency ef) is a complicated matter. Studies have suggested empirical relationships with temperature which offer attractive potential for parameterization. Here we develop what is arguably the simplest mechanistic model relating the two, using established thermodynamic dependencies for primary production and respiration. It results in a single-parameter curve that constrains the envelope of possible efficiencies, capturing the upper bounds of several ef - T data sets. The approach provides a useful theoretical constraint on this relationship and extracts the variability in ef due to temperature but does not idealize out the remaining variability which evinces the substantial complexity of the system in question.

2.1.2 Introduction

The export of organic carbon out of the upper ocean is an important component of the climate system, driving the “biological pump” which reduces the partial pressure of atmospheric carbon dioxide and fuels the ecosystems of the deep ocean and benthos [Archer et al., 2000]. The efficiency with which limiting resources (usually nutrients) are exported, relative to local recycling, is often termed the ef ratio, here defined as the ratio of the sinking flux of particulate organic carbon across a defined depth horizon and the integrated primary production P_p above that horizon, e.g., [Laws et al., 2000b]. Eppley and Peterson [Eppley and Peterson, 1979] identified a simple, empirical relationship between ef and integrated primary production, but it has been difficult to establish a clear theoretical basis for the controls on ef due to the myriad physical and biological processes at play [de la Rocha and Passow, 2007].

Laws et al. [Laws et al., 2000b] examined a relatively detailed numerical model of the planktonic ecosystem, which suggested that ef is shaped by P_p and mixed layer temperature T . A compilation of self-consistent observations of export efficiency ef along with local physical and biogeochemical factors, from the Joint Global Ocean Flux Study (JGOFS) Process Study data, supported this interpretation. Both model and data suggested an approximately linear, negative correlation between ef and mixed layer temperature, T , and that the temperature dependence of the ecosystem processes which shape export production provide a dominant control on ef . Indeed, the temperature variation explained far more of the variance in ef than P_p in that data set. A series of subsequent studies [Laws et al., 2000a, Henson et al., 2011, Dunne et al., 2005, Maiti et al., 2013] have explored the validity and possible forms of temperature- ef relationships, and there has been a significant increase in the empirical data constraints over the past 15 years. The simple correspondence between ef and T of Laws et al. [Laws et al., 2000b] is not clearly supported with a much expanded data set. Consequently, recent models and interpretations of these data sets have not lead to a consistent, simple relationship between ef and T .

However, Laws et al. [Laws et al., 2000a] revisited the ef - T relationship from an empirical perspective, showing that the upper bound of ef declines as temperature increases. Here we consider this upper bound from a mechanistic perspective. At the heart of the temperature dependence in the Laws et al. [Laws et al., 2000b] model is the differential

temperature sensitivity of phototrophic and heterotrophic metabolism [López-Urrutia et al., 2006, Huntley and Lopez, 1992, Eppley, 1972] (clearly characterized by Rose and Caron [Rose and Caron, 2007]). We develop a highly idealized framework which reflects this key element. We argue that while finding a simple relationship between ef and T will be confounded by many other factors in real systems, there is a predictable mechanistic relationship between the maximum export ratio, ef_{max} and T which reflects situations where all other limitations and constraints (e.g., nutrients) are relaxed, analogous to the interpretation of the Eppley Curve [Eppley, 1972] (see Figure 2-1). There the simple parameterization relates to the upper bound of growth rate: in any given circumstance other factors such as nutrient limitation might restrict growth and so only the maximum of the observed data points represents the effect of temperature clearly.

Hence, we seek to characterize an envelope bounding export efficiency based on thermodynamic constraints. To do so, we will write the simplest model which captures the essential dynamics (Section 2.1.3). Starting with an ordinary differential equation describing the time rate of change of autotrophic biomass, we derive a predicted curve for maximum export efficiency as a function of temperature. In Section 2.1.4 we show that this simple, mechanistic model captures the trends in a compilation of empirical data with global coverage [Laws et al., 2000b, Dunne et al., 2005, Henson et al., 2011, Buesseler and Boyd, 2009] and discuss the value and limitations of the framework.

2.1.3 Temperature vs. export efficiency: Model and Data

We first write a simple ordinary differential equation for the phytoplankton biomass density (p) in the upper ocean (the mixed layer, or euphotic zone, or above the thermocline):

$$\dot{p} = \mu p - \lambda p - \lambda' p - wp \quad (2.1)$$

where μ is a linear growth rate, λ represents loss due to grazing, and λ' represents losses from other processes including, e.g., viral lysis, senescence, and detrainment. w represents the rate of export as sinking particles; all coefficients have dimensions of inverse time. Assuming steady state and dividing through by μp , Eq. (2.1) becomes

$$0 = 1 - \left(\frac{\lambda + \lambda' + w}{\mu} \right) \quad (2.2)$$

Figure 2-1: Temperature sensitivity for phytoplankton and copepod growth rates, adapted from (a) Eppley [Eppley, 1972] and (b) Huntley and Lopez [Huntley and Lopez, 1992]. The temperature scaling prefactor is larger for the copepods, the key thermodynamic relation that drives this model. Another key difference is that copepod growth clusters along, while phytoplankton growth lies under, their respective curves; copepod growth appears to deviate significantly from this curve near $T = 24^{\circ}\text{C}$. These predictions from the above pictured original data sets have been additionally tested and verified in, e.g., Rose and Caron [Rose and Caron, 2007]. Note the x axes of each subplot cover different limits. The prefactor is the value which multiplies temperature in the expressions $\mu \propto e^{0.063T}$, $\lambda \propto e^{0.11T}$.

Export efficiency is defined as the ratio of export production and primary production: $ef := wp/\mu p$ [Laws et al., 2000b], so Eq. (2.2) becomes

$$ef = 1 - \frac{\lambda + \lambda'}{\mu} \quad (2.3)$$

The key element of the model will be the differential temperature dependence of autotrophic and heterotrophic metabolism [Laws et al., 2000b]. We draw upon empirically established temperature (T [°C]) dependencies for photosynthetic growth rates μ [Eppley, 1972, Rose and Caron, 2007] and zooplankton grazing rates λ [Huntley and Lopez, 1992, Rose and Caron, 2007] which are illustrated in Figure 2-1:

$$\lambda \propto e^{0.11T} \quad \mu \propto e^{0.063T} \quad (2.4)$$

Such temperature dependencies result from the different activation energies for photosynthesis and respiration, thus constrain growth thermodynamically [López-Urrutia et al., 2006]. Loss processes represented by λ' may have their own temperature dependencies, e.g. [Danovaro et al., 2011]; we test for effects of possible λ' temperature dependence in Section 2.1.4. After substituting these expressions for λ and μ into Eq. (2.3) the export ratio can now be expressed as

$$ef = 1 - \alpha e^{0.047T} - \lambda'/\mu \quad (2.5)$$

where α is a constant and λ'/μ represents reductions in ef due to loss processes other than grazing. The above expression highlights an important asymmetry; while there is only one way for biomass to increase, there are multiple pathways by which biomass p is lost. Thus, setting $\lambda' \equiv 0$ in Eq. (2.5) results in an expression for the upper bound on the export efficiency ratio when loss is comprised only of sinking and processes related to grazing. It is an upper bound because the copepod growth rates cluster around a compact curve in temperature space, and phytoplankton growth rates are spread beneath a maximum growth rate versus temperature [Huntley and Lopez, 1992, Eppley, 1972, Rose and Caron, 2007]. We note that the temperature scaling for autotrophic processes is valid up to 40°C, but the temperature dependency for zooplankton growth rate (thus grazing rate) is only an reasonable fit for up to 24°C, after which the temperature scaling breaks down and appears

to flatten out, possibly indicative of a more complicated predator physiological response to temperature. This cutoff near 24°C is observable both in the original Huntley and Lopez [Huntley and Lopez, 1992] paper and in the significantly expanded data of Rose and Caron [Rose and Caron, 2007]: one possible hypothesis for this cutoff is that the decreased solubility of oxygen at higher temperatures begins to constrain heterotrophic processes, e.g., [Deutsch et al., 2015].

There is no clear theoretical constraint with which to set α , so we may leave it as a parameter. To set α , we use one point from each data set described below, so that the remaining data can be used to test the prediction. We let α set the intercept, because $ef_{max}(T = 0) = 1\alpha$. This is done by setting α so $ef_{max}(T = 0)$ corresponds to the weighted average of the maxima of the data sets analyzed, leaving $\alpha = 0.24$. The formula used to compute α is

$$\alpha := 1 - \frac{1}{N} \sum_{i=1}^4 \frac{\max(ef)_i}{n_i} \quad (2.6)$$

where i is an index of each dataset used, N is the total number of data points, and n_i is the number of data points in each data set. The slope/curvature of the prediction itself is set by the temperature dependencies. In general, a larger α would correspond to a lower possible $ef_{max}(T)$ for all values of T , with larger differences at lower temperatures. However, changing α by $\pm 10\%$ yielded no effect on the results described in the following section.

2.1.4 Temperature shifts measurements' distribution

In Figure 2-2 we compare the predicted upper bound on ef from Eq. (2.5) with four empirical data sets compiled from recent publications: Laws et al. [Laws et al., 2000b], Dunne et al. [Dunne et al., 2005], Henson et al. [Henson et al., 2011], and additional data available from the Biological and Chemical Oceanography Data Management Office database [Buesseler, 2004]. The data sets are described in more detail in the cited sources; they include either sea surface or mixed layer temperature data, and export efficiency is evaluated using a variety of approaches, including estimates of export production based on carbon, oxygen, nitrate and other nutrients, and estimates of particle export based on sediment trap sinking fluxes or thorium isotopes. While the eclectic nature of the compilation introduces some comparative uncertainty, the number of data sources provides a large time window and global spatial coverage. Given the challenging nature of measuring and quantifying export, only recently

has a sufficiently large body of data emerged with which to test such theories.

Over the temperature range for which the model is applicable, 0–24°C, the predicted upper bound and the maximum observed ef show very good agreement. There is only one significant outlier at 23°C, and the data points span the entirety of the domain beneath as expected. The spread reflects the influence of other loss processes (λ') that reduce ef , as well as constraints other than temperature on phytoplankton growth rates – i.e., the spread that underpins the Eppley μ - T relationship (see Figure 2-1 and Eq. (2.3)). Above 24°C the exponential extrapolation of the prediction does not capture the maximum extent of the empirical data. We attribute this to a breakdown of the exponential relationship between zooplankton growth rates and temperature above this threshold (Figure 2-1). However, the mechanisms underlying that breakdown are not clear and may be the result of either metabolic or environmental constraints.

The variability in the distribution of the data in Figure 2-2 is indicative of the natural variability and complexity of carbon export and is likely due to a number of compounding effects, e.g., sampling difficulties [Buesseler et al., 1994], growth intermittency, eddy subduction [Omand et al., 2015], and the number of different remineralization pathways available to the community (i.e., in the λ' term) which this model ignores. Indeed, this complexity indicates the intractability of a simple mechanistic description of export, which suggests the utility of mechanistic approaches for establishing *constraints* rather than parameterizations.

Here our central hypothesis is that the difference in temperature dependencies of autotrophic and heterotrophic growth controls the envelope of variations in ef . This implies that the temperature dependence of the unresolved loss processes, λ' , is relatively weak. This assumption can be tested by rescaling all ef data points by the predicted ef_{max} curve and asking if the distribution changes as temperature increases. If the mean $\bar{ef}$ or normalized variance $\sigma^2/\bar{ef}$ varies with temperature, this suggests λ' has some temperature dependence. In turn, this would demand to be incorporated into a relationship between ef and T . However, when all data is normalized by the $ef_{max} = 1\alpha e^{0.047T}$ curve, i.e., the predicted influence of temperature is removed, we see no change in $\bar{ef}$ or $\sigma^2/\bar{ef}$ with temperature. This demonstrates that the model captures all systematic temperature dependence in the data (Figure 2-3) and that the remaining variability is attributable to other factors.

Figure 2-2: Model prediction for maximum export efficiency as a function of temperature, compared with data from sources described in Methods. The four points enclosed by diamonds were those taken to fix parameter $\alpha = .24$. The model agrees well with data in its applicable temperature range – vertical dashed line indicates the cutoff at 24°C observable in Huntley and Lopez [Huntley and Lopez, 1992] and Rose and Caron [Rose and Caron, 2007]. Remaining dashed lines indicate possible extrapolations of the model prediction above 24°C.

Figure 2-3: Mean and normalized variance of the data after being binned at width 1°C and rescaled by predicted maximum export efficiency. Binned values are best explained by a constant line; the best fit regression line computed with ordinary least squares has a negative adjusted r^2 ; thus, after the model's prediction for temperature dependence has been removed, no variation in the binned values can be explained by temperature.

2.1.5 Discussion

The model prediction described here shows good agreement with data, yielding an upper bound on potential export efficiency guided by thermodynamic constraints. This envelope, in the spirit of the Eppley curve, hinges on the asymmetry between the number of pathways for primary production (one) and loss (several). Despite the variety of methods used to estimate export efficiency from field data [Henson et al., 2011] all data sets approach and fall within the upper bound suggested by the model.

Limitations of the model include the strong assumption of steady state. The implications of this could be examined in time-dependent models. The assumption is more applicable in regions where upper ocean phytoplankton biomass density is slowly varying and least applicable in bloom forming regions. Collecting further metabolic data for grazers growth rates above 24°C, and an exploration of the mechanisms controlling metabolism at warmer temperatures could improve the foundation of the model.

Using this model to extract the temperature dependence of ef provides a foundation for investigating dependencies of ef on other variables including primary productivity [Laws et al., 2000a, Maiti et al., 2013], opening the door for further development of mechanistic models of ocean carbon export [Siegel et al., 2014]. There is a longstanding history of relating primary production and export [Eppley and Peterson, 1979, Laws et al., 2000b], though the relationship may be complex and regionally varying [Maiti et al., 2013]. Clearly, understanding and mechanistically modeling the relationship to primary production would be a valuable extension to this model (see [Laws et al., 2000a]). An interesting avenue for extension of this model would be to incorporate an allometric component since cell size has been empirically related to export rates [Guidi et al., 2009] and the theoretical foundations have been explored [Laws, 1975].

We show that this model describes all systematic temperature dependence of the ef data (Figure 2-3) and that this is due to the difference in temperature dependencies of autotrophic and heterotrophic metabolisms. A significant inference of the model is that independent of other factors, a warming ocean should lead to a weakened export efficiency and biological carbon pump, by a quantifiable extent. Of course this may not be the dominant factor since other, indirect effects of warming such as changes in stratification and primary productivity will also be at play.

In summary, we present a model that provides a mechanistic constraint on the range of potential export efficiency values, based on empirically established temperature dependencies and differential equation approaches fundamental to marine ecology. The prediction is quite general, reinforced by a collection of data sets and the simplicity of its derivation. While this approach moves away from predictions or parameterizations which expect a tight relationship between export efficiency and temperature, it provides a potentially useful framework by linking metabolic theory and the biological carbon pump and acknowledges the substantial complexity of the flow of carbon in the upper ocean.

2.2 How have recent temperature changes affected the efficiency of ocean biological carbon export?

2.2.1 Overview

The ocean’s large, microbially mediated reservoirs of carbon are intimately connected with atmospheric CO₂ and climate, yet quantifying the feedbacks between them remains an unresolved challenge. Through an idealized mechanistic model, we consider the impact of documented climate change during the past few decades on the efficiency of biological carbon export out of the surface ocean. This model is grounded in universal metabolic phenomena, describing export efficiency’s temperature dependence in terms of the differential temperature sensitivity of phototrophic and heterotrophic metabolism. Temperature changes are suggested to have caused a statistically significant decrease in export efficiency of $1.5\% \pm 0.4\%$ over the past 33 yr. Larger changes are suggested in the midlatitudes and Arctic. This interpretation is robust across multiple sea surface temperature and net primary production data products. The same metabolic mechanism may have resulted in much larger changes e.g., in response to the large temperature shifts between glacial and interglacial time periods.

2.2.2 Introduction

The ocean’s “biological pumps” sustain large reservoirs of carbon, mediated by microbial activity, with significant leverage on atmospheric CO₂ and climate [Volk and Hoffert, 1985]. While the general significance of the biological pumps for the carbon cycle and climate is clearly demonstrated [Volk and Hoffert, 1985, Cox et al., 2000] the details of their relationship

remain elusive [Boyd, 2015]. Understanding the current functioning of the biological pumps is limited because in situ data are sparse [Boyd and Trull, 2007], data collection is difficult and expensive, and the system is extremely complex and variable [Buesseler and Boyd, 2009].

Even so, there exists a clear and quantifiable imprint of metabolic sensitivity to temperature in the ocean system [Eppley, 1972], which can be exploited to understand global changes in carbon export with climate change. In particular, differential sensitivities of phototrophic and heterotrophic metabolisms to environmental temperature are documented [Eppley, 1972, Huntley and Lopez, 1992], and a model of their effect on export efficiency (ef , the ratio of the flux of organic matter exported across the base of the euphotic zone to the integrated primary production within that layer) explains the observed dependence of ef on temperature (Section 2.1) [Cael and Follows, 2016]. Here we ask: how has the documented trend in global ocean temperatures over the past few decades impacted the efficiency of this export flux? We use temperature records with the above model to infer temperature’s contribution to global change in ef through time. We focus on multidecadal changes (after [Henson et al., 2010]) over the past 33 yr (the duration over which suitable data products are available) using multiple data products to examine the sensitivity of calculated changes to inputs used.

2.2.3 A metabolic model of export efficiency

Export efficiency is a combination of growth, respiration, sinking, remineralization, and other processes. It has primarily been considered as a function of temperature (T), primary production ($\mathcal{P}$), and community structure (i.e. the size distribution of plankton, who is eating whom, and so forth) [Laws et al., 2000b, Eppley, 1989, Michaels and Silver, 1988]. While all three are important, community structure variables and their influence on ef are challenging to assess, quantify, and measure, making an estimation of how recent climatic shifts have produced global shifts in ef via community structure challenging. Previous studies disagree substantially on both global trends in $\mathcal{P}$ [Behrenfeld et al., 2016, Siegel and Others, 2013] (also references therein), and the relationship between ef and $\mathcal{P}$ [Buesseler, 1998, Laws et al., 2000a, Maiti et al., 2013], making an estimation of how climatic shifts have affected ef via $\mathcal{P}$ similarly intractable.

In contrast, global trends in sea surface temperature (SST) over the past few decades are well-characterized [IPCC, 2014]. SST is commonly used as a proxy for upper-ocean

temperature, is anticorrelated with ef [Laws et al., 2000b, Henson et al., 2011], and is also one of the only variables for which long-term, global observational records exist [Ishii et al., 2005, Reynolds et al., 2007, Dee and Others, 2011]. Do SST observations, suggest a shift in ef ?

In Section 2.1, a simple model was proposed ([Cael and Follows, 2016]; herein the model will be referred to as MM, for ‘Metabolic Model’) to explain the ef - T relationship seen in observations. Heterotrophic and phototrophic growth rates increase with temperature, but the former increase more so [Huntley and Lopez, 1992, Eppley, 1972, Rose and Caron, 2007]; metabolic ecological theory relates this to the different activation energies of respiration and photosynthesis [López-Urrutia et al., 2006]. As originally posited in [Laws et al., 2000b], this differential dependence suggests that increasing temperatures should increase community respiration relative to production and therefore decrease ef . Rather than absorbing these dependencies into a numerical food-web model as in [Laws et al., 2000b], MM considers ef as a random variable scaled by temperature according to these dependencies, and this description is shown to be consistent with observations. We refer the reader to section 2.1 for a full description and discussion of MM, but describe it briefly below in terms of the present context.

Within a basic differential equation for plankton biomass p in the euphotic layer,

$$\dot{p} = \mu p - \lambda p - \lambda' p - wp \quad (2.7)$$

where μ is the growth rate, λ is the grazing rate, λ' is the loss rate due to factors other than grazing, and w is the sinking rate, in steady state $ef = \frac{wp}{\mu p}$ can be written as

$$ef = 1 - \frac{\lambda + \lambda'}{\mu} \quad (2.8)$$

Then one can find the maximum efficiency by neglecting λ' and incorporating the temperature dependencies of phototrophy and heterotrophy as $\mu \propto e^{0.063T}$, $\lambda \propto e^{0.11T}$. This yields a curve of maximum export efficiency as a function of temperature:

$$ef_{max}(T) = 1 - \alpha e^{\beta T} \quad (2.9)$$

where the parameter¹ α is one minus the maximum efficiency at $T = 0$, which is estimated empirically to be $\alpha = 0.24$, and the parameter $\beta = 0.11 - 0.063 = 0.047$. ef values can be rescaled by $ef_{max}(T)$ to extract this temperature dependence; that is, $ef/ef_{max}(T)$ is variable but independent of temperature, suggesting that this rescaling captures all of the temperature dependency of ef [Cael and Follows, 2016].

Here we then use this temperature scaling to derive an average $\langle ef \rangle$ as a function of temperature²:

$$\langle ef \rangle (T) = \langle \widetilde{ef} \rangle \times ef_{max}(T). \quad (2.10)$$

While empirical models have been proposed to relate temperature to ef [Laws et al., 2000b, Laws et al., 2000a, Dunne et al., 2005, Henson et al., 2011], we focus on MM because it isolates a single, understood metabolic mechanism - the hypothesized and observed differential response of temperature on phototrophic and heterotrophic metabolisms. We emphasize that MM does not seek to be a complete model for or explain all the variability in ef ; it isolates the variability in ef due to the differential temperature effect on metabolism. Because the differential temperature response is assumed to arise from chemical kinetics, namely the activation energies of respiration and photosynthesis, it is assumed to be constant over time.

Figure 2-4 shows the estimated percent change of $\langle ef \rangle$ resulting from a temperature change from T_{old} to T_{new} as predicted by MM. Percent changes are a function of both, and as Eq. (2.10) varies by a factor of three from low to high temperatures, percent changes in response to large temperature differences can be very large. Note that temperature differences shown in Figure 2-4 are larger than those that have been observed over the past 33 years, which are $<1^\circ\text{C}$ [IPCC, 2014].

2.2.4 Global estimates of multidecadal change in export efficiency

MM can be used to calculate $\langle ef \rangle$ from SST³, so it can be used to infer trends in globally averaged ef (for which we will use the symbol⁴ $\langle\langle ef \rangle\rangle$) from SST reanalyses. Does MM

¹Figure 2-6 (see Section 2.2.7) shows the sensitivity of MM to changes in the parameters α and β .

² $\langle \widetilde{ef} \rangle = 0.37$ for the observations used in [Cael and Follows, 2016] so we use that value here nominally, though we note the value of this factor is largely irrelevant for the analyses of this chapter because we focus on percent changes.

³as a proxy for upper-ocean temperature

⁴all export efficiency notation: $ef :=$ export flux divided by primary production. $\mathcal{P} :=$ primary production. $T :=$ temperature. $ef_{max}(T) :=$ theoretical maximum ef for a given temperature. $\alpha = 0.24$: one minus the maximum efficiency at $T = 0$, $\beta = 0.047$: the differential temperature sensitivity of phototrophy and heterotrophy. $\widetilde{ef} := ef/ef_{max}(T)$. $\langle ef \rangle :=$ mean ef averaged over a spatial region, e.g. a 1° box or a

Figure 2-4: Percent change in $\langle ef \rangle$ after a temperature change, as estimated by MM (see Eq. (2.9)), as a function of the initial and final temperatures (T_{old} and T_{new}). Contours are spaced at 5%.

indicate a trend? Is this trend consistent between data products?

To test these questions, we generate three time series of $\langle ef \rangle$ from different SST reanalyses and Eq. (2.10). We use the ECMWF ERA-Interim SST [Dee and Others, 2011], the NOAA OISST [Reynolds et al., 2007], and the ESRL COBE-SST [Ishii et al., 2005] products. Description and discussion of how these reanalyses are developed can be found in the above references. For consistent comparison, we use a common

- spatial resolution of 1° , the resolution of the coarsest product
- temporal resolution of 1 month, i.e. monthly averages available for each product
- start time of January 1982, the January of the earliest year common to all three products
- end time of December 2014, the December of the latest year common to all three products

From each SST product we compute global time series of $\langle ef \rangle$ for each 1° bin.

Because we define $\langle\langle ef \rangle\rangle$ as the ratio of globally integrated export flux to globally integrated production, to compute it $\langle ef \rangle$ for each 1° bin must be weighted by both area and $\mathcal{P}$. We use climatologies⁵ from the two most common algorithms to estimate $\mathcal{P}$: the Carbon-based Productivity Model (CbPMv2) [Westberry et al., 2008], and the Vertically Generalized Productivity Model (VGPM) [Behrenfeld and Falkowski, 1997]. In total, the time series $\langle\langle ef \rangle\rangle(t)$ is calculated by

$$\langle\langle ef \rangle\rangle(t) := \frac{\sum_{x,y} A(x,y) \mathcal{P}(x,y,t \bmod 12) \langle ef \rangle(x,y,t)}{\sum_{x,y} A(x,y) \mathcal{P}(x,y,t \bmod 12)} \quad (2.11)$$

where (x,y) are latitude and longitude, $A(x,y)$ is the area of the $1^\circ \times 1^\circ$ box at (x,y) , and $t \bmod 12$ is the month of the $\mathcal{P}$ -climatology.

The above three SST and two $\mathcal{P}$ products yield six time series of $\langle\langle ef \rangle\rangle(t)$. We regress each against time using the simplest statistical model that resolves a seasonal cycle and a linear trend:

$$\langle\langle ef \rangle\rangle(t) = mt + b(\text{month}) + \text{error} \quad (2.12)$$

latitudinal band. $\langle\langle ef \rangle\rangle :=$ globally averaged ef .

⁵both available at <http://www.science.oregonstate.edu/ocean.productivity/standard.product.php>. We use climatologies rather than time series because of the lack of $\mathcal{P}$ time series over the duration of the SST time series. This is justifiable in light of the disagreement on global trends in $\mathcal{P}$ [Behrenfeld et al., 2016, Siegel and Others, 2013].

Table 2.1: Percent change in $\langle\langle ef \rangle\rangle$, during the period 1982-2014, as estimated by MM, for different data products. All changes suggest a decrease in $\langle\langle ef \rangle\rangle$ and are statistically significant ($p < .001$; see Section 2.2.7). Mean and standard deviation of percent change across each SST- $\mathcal{P}$ pair are $-1.5 \pm 0.4\%$.

SST	$\mathcal{P}$	% Change
ERA	CbPMv2	-0.83
ERA	VGPM	-1.12
COBE	CbPMv2	-1.55
COBE	VGPM	-1.98
OISST	CbPMv2	-1.52
OISST	VGPM	-1.91

regressing $\langle\langle ef \rangle\rangle(t)$ against a time variable that runs $t = 1, 2, \dots$ along with an indicator variable for each month (see Section 2.2.7). To estimate statistical significance of the trends, we use both the standard method and two resampling methods, one of which accounts for autocorrelation in the time series (see Section 2.2.7). The estimated rates of change m can be multiplied by the duration of the time series to estimate a percent change in $\langle\langle ef \rangle\rangle(t)$ from 1982-2014 for each SST- $\mathcal{P}$ pair; see Table 2.1.

Independent of SST and $\mathcal{P}$ product, a global decline in $\langle\langle ef \rangle\rangle$ is observed of $1.5 \pm 0.4\%$, where the uncertainty is the standard deviation across SST- $\mathcal{P}$ pairs. All declines are found to be significant ($p < .001$) by all three significance estimation procedures. Over the entire timeseries, MM predicts a $\langle\langle ef \rangle\rangle$ value of 0.128 ± 0.016 .

How consequential is a $1.5 \pm 0.4\%$ decrease in $\langle\langle ef \rangle\rangle$? A simple box model of the carbon cycle [Ito and Follows, 2005, Williams and Follows, 2011] (see Section 2.2.6) suggests a $1.5 \pm 0.4\%$ decline in $\langle\langle ef \rangle\rangle$ would result in a $1.2 \pm 0.3\%$ increase in the mixing ratio of atmospheric CO_2 on millennial timescales (when the solubility pump has equilibrated but prior to carbonate compensation). Generalizing this result, the sensitivity of the soft-tissue carbon pump to changes in mean sea surface temperature is predicted to be $\sim 7 \text{ ppm K}^{-1}$, comparable to the $\sim 10 \text{ ppm K}^{-1}$ sensitivity of the solubility pump to global mean ocean temperature suggested by theory and models [Williams and Follows, 2011] (their Fig. 13.9b). Interestingly, the combined sensitivity of solubility and metabolic effects predicts a $\sim 70 \text{ ppm}$ drawdown of atmospheric CO_2 for a 4K global cooling of the ocean associated with the last Glacial Maximum [Adkins et al., 2002].

We note that substantial latitudinal variation in SST trends have occurred during the

Figure 2-5: Percent change in $\langle ef \rangle$ from 1982-2014, as estimated by MM, as a function of latitude, for different data products. Color corresponds to SST product and line type to $\mathcal{P}$ product. Black curve is the average across the SST- $\mathcal{P}$ pairs.

record of these SST reanalyses. Thus, we employ the same procedure as above, but only averaging zonally, yielding an estimated percent change in $\langle ef \rangle$ (here the ratio of longitudinally integrated export flux to longitudinally integrated production) at each degree of latitude from 1982-2014. See Figure 2-5; substantial latitudinal variation exists, with percent changes ranging between +2% to -8%. $\langle ef \rangle$ in the Southern Ocean increases slightly, corresponding to cooling, while $\langle ef \rangle$ changes little in the tropics where SST is high. Thus, it appears the global decrease of 1-2% is driven by decreases in the Arctic and at midlatitudes, where $\langle ef \rangle$ decreases on the order of 5%. While variation exists between each of the six SST- $\mathcal{P}$ pairs, their latitudinal dependence is similar.

2.2.5 Conclusion

Neither the existence, magnitude, nor driving mechanisms of a change in the biological pump over the past few decades of recent climate change can be established definitively. This is an unavoidable result of searching for small shifts in a system exhibiting substantial variability on all time scales that is challenging to measure adequately. A simple metabolic perspective accounting for the differential temperature dependencies of autotrophy and heterotrophy underpins a model of export efficiency, which explains its observed dependence on temperature. Driving that model with observed changes in SST leads to a suggestion that global export efficiency has decreased 1.5 ± 0.4 % over the past few decades, with larger decreases in midlatitudes and Arctic. This decrease is robust across SST and primary production data

products. Larger temperature differences than those observed over the past few decades are predicted to cause larger changes in global export efficiency.

2.2.6 Sensitivity of atmospheric CO₂ to export efficiency at equilibrium

We are motivated to understand the controls on the export ratio, ef (NB: for this section, as we are only interested in the global export ratio, i.e. the fraction of global primary production that is exported from the surface ocean, we drop the double brackets $\langle\langle\cdot\rangle\rangle$ on ef) because of its potential significance for the so-called "biological pump", i.e. the biologically mediated storage of inorganic carbon in the ocean at the expense of atmospheric CO₂. To provide context, here we use an idealized, globally integrated perspective to estimate the sensitivity of the mixing ratio of atmospheric CO₂.

Consider a simple two box model of the ocean carbon cycle (see Figure 1-2). A surface layer has a mean concentration of dissolved inorganic carbon, $\overline{C^{\text{surf}}}$ (mol kg⁻¹). CO₂ is exchanged across the sea surface and $\overline{C^{\text{surf}}}$ is at equilibrium with an atmospheric reservoir of mixing ratio χ_{CO_2} (mol CO₂ mol air⁻¹). Thus, by definition, $\overline{C^{\text{surf}}} = \overline{C^{\text{sat}}}$, where $\overline{C^{\text{sat}}}$ is the saturation concentration of dissolved inorganic carbon; a known function of temperature, pressure, total alkalinity (which we will assume to be constant) and χ_{CO_2} . The deep ocean carbon concentration, $\overline{C^{\text{deep}}}$ is the sum of a contribution from subducted surface waters and a biologically mediated accumulation: $\overline{C^{\text{deep}}} = \overline{C^{\text{sat}}} + \overline{C^{\text{reg}}}$. The biologically mediated pool, $\overline{C^{\text{reg}}}$, is here assumed to represent the so-called "soft-tissue pump" [Volk and Hoffert, 1985]; the result of the remineralization of sinking or subducted organic matter which was formed in the sunlit surface layer. The downward flux of organic carbon leaving the surface layer is assumed to be completely remineralized in the deep layer, leading to an average source over the deep ocean of E (mol C kg⁻¹ s⁻¹), which can be related to globally integrated marine primary production, P (mol C s⁻¹) by

$$E = \frac{efP}{V\rho} \quad (2.13)$$

where ef is the global export ratio, V is the volume of the ocean (m³), and ρ is the mean density of seawater (kg m⁻³). The exchange of surface and deep water properties is represented by a two-way volume flux, ψ (m³ s⁻¹) and the rate of change of $\overline{C^{\text{deep}}}$ can be

described by

$$\frac{\partial \overline{C^{\text{deep}}}}{\partial t} = \frac{\psi}{V} (\overline{C^{\text{surf}}} - \overline{C^{\text{deep}}}) + E \quad (2.14)$$

which, at steady state, leads to an expression relating the strength of the biological pump, $\overline{C^{\text{reg}}}$, to the export ratio, ef:

$$\overline{C^{\text{reg}}} = ef \frac{P}{\psi \rho} \quad (2.15)$$

If we assume that global primary production is constant and set by the availability of light or nutrients, then $\overline{C^{\text{reg}}}$ is a simple, linear function of ef .

Following [Ito and Follows, 2005, Goodwin et al., 2008, Williams and Follows, 2011], it can be shown that variations in the ocean's biologically mediated carbon reservoir and atmospheric CO₂ can be expressed as

$$\frac{\delta \chi_{\text{CO}_2}}{\chi_{\text{CO}_2}} \sim -\frac{V \rho}{I_B} \delta \overline{C^{\text{reg}}} \quad (2.16)$$

where I_B is the buffered carbon inventory (mol C), effectively constant for the purposes of this analysis. Combining Eq. (2.15) and Eq. (2.16) we can relate the steady state response of the mixing ratio of atmospheric CO₂ to changes in export efficiency, ef (ignoring carbonate compensation, which would act to reduce the impact on even longer timescales).

$$\frac{\delta \chi_{\text{CO}_2}}{\chi_{\text{CO}_2}} \sim -\frac{efPV}{I_B \psi} \frac{\delta ef}{ef} \quad (2.17)$$

We evaluate this sensitivity with numerical values of $ef = 0.13$, $P = 1.3 \times 10^8 \text{ mol C s}^{-1}$ ($= 50 \text{ PgC yr}^{-1}$), $V = 1.4 \times 10^{18} \text{ m}^3$, $I_B = 2.6 \times 10^{17} \text{ mol C}$, and $\psi = 1.2 \times 10^8 \text{ m}^3 \text{ s}^{-1}$ (derived by assuming an Ekman pumping rate of 20 m yr^{-1} over half the ocean's surface and the equal and opposite suction rate over the other half). This suggests that

$$\frac{\delta \chi_{\text{CO}_2}}{\chi_{\text{CO}_2}} \sim -0.8 \frac{\delta ef}{ef} \quad (2.18)$$

so decreasing ef by $1.5 \pm 0.4\%$ would lead to an increase in steady state χ_{CO_2} of about $1.2 \pm 0.3\%$, or about $3.4 \pm 0.8 \text{ ppm}$ relative to a preindustrial baseline of 280 ppm . As i) Eq. (2.18) is linear, ii) MM predicts that percent changes in ef are approximately linear with temperature changes (see Figure 2-4 and Eq. (2.10)), and iii) global mean sea surface temperatures increased approximately 0.5°C from 1982-2014 (IPCC, 2014), this implies

an approximate ~ 7 ppm K^{-1} sensitivity of atmospheric CO_2 to the metabolic mechanism considered in this chapter on millennial timescales.

2.2.7 Statistical details and sensitivity

Sensitivity of MM to α and β :

We include Figure 2-6 to show that the exact percent changes in $\langle ef \rangle$ are sensitive to the exact values of the MM model parameters α and β , but that the overall pattern of $\langle ef \rangle$ changes, resulting from temperature changes from T_{old} to T_{new} , is not substantially different. Thus the predicted 1-2% decrease in $\langle ef \rangle$ is robust to some uncertainty in these parameters, as the uncertainty across data products is larger than that due to the uncertainty in α and β .

Implementation of the Regression:

Numerically, the regression of the statistical model is implemented by regressing $\langle\langle ef \rangle\rangle(t)$ against a 13×396 matrix T_{ij} , where the first row $T_{1j} = 1, \dots, 396$ and the remaining rows consist of repeating 12×12 identity matrices. This is the simplest model that allows for both a seasonal cycle (without assuming a particular shape to that cycle) and a linear trend.

p -value Estimation:

We use three procedures to estimate the statistical significance of the linear trends reported.

The first is the standard method, which assumes that residuals from the regression are normally, identically, and independently distributed. This method is described in many statistics textbooks so we will not describe it further herein.

The other two are bootstrapping (i.e. resampling with replacement) procedures. In each case, we generate 10^5 bootstrapped time series $b_i(t)$ from each time series of $\langle\langle ef \rangle\rangle(t)$, then apply the regression to each time series. The p -value is then the proportion of slopes estimated from each $b_i(t)$ which have an absolute value larger than the slope estimated from the $\langle\langle ef \rangle\rangle(t)$ from which the $b_i(t)$'s were generated (that is, the p -value is a two-sided p -value, because slopes can be positive or negative).

For the first bootstrapping procedure, if t is the index referring to the month in the time series ($t = 1, \dots, 396$), for each $b_i(t)$ we randomly sample from all of the data with the same month as t .

This procedure allows us to estimate the slopes resulting from random variability of data

Figure 2-6: a-d) As in Figure 2-4, except for α or β are changed by $\pm 10\%$ in each case, as noted below each subfigure. The overall structure of ef changes is similar, but the magnitude of ef changes for a given temperature change increases with both α and β .

generated by a distribution similar to that of the sampled data while preserving the annual structure. However, it does not account for the fact that the residuals of the regressions of $\langle\langle ef \rangle\rangle(t)$ time series are substantially autocorrelated (i.e. an anomalously large January value is likely followed by an anomalously large February value), thus it underestimates random interannual variability.

The second bootstrapping procedure corrects for this by resampling year-blocks rather than individual month values. The procedure is similar to above except that for each $b_i(t)$, if $b_i(t+1)$ has the same year in the time series, it does in the bootstrapped time series as well. That is, for a given $b_i(t)$, the first twelve measurements are all taken from the same (randomly chosen) year, then the next twelve measurements are all taken from the same year, and so forth.

This procedure allows us to estimate in a straightforward way the slopes generated by random variability from a distribution similar to the sampled data while preserving the annual structure and the autocorrelation structure; comparison of the autocorrelation structures of many year-bootstrapped time series and the time series from which they were derived reveals that their autocorrelation structures are nearly identical.

We find that for all six time series of $\langle\langle ef \rangle\rangle(t)$, one for each SST- $\mathcal{P}$ pair, all three p -value estimation procedures find $p < .001$. Thus the estimated trends are unequivocally statistically significant.

Chapter 3

Primary production and export

The work in this chapter is based upon the following publication:

Cael, B. B., K. Bisson, and C. L. Follett. 2018. Can Rates of Ocean Primary Production and Biological Carbon Export Be Related Through Their Probability Distributions? *Global Biogeochemical Cycles*. [Cael et al., 2018]

3.1 Can Rates of Ocean Primary Production and Biological Carbon Export Be Related Through Their Probability Distributions?

3.1.1 Overview

We describe the basis of a theory for interpreting measurements of two key biogeochemical fluxes – primary production by phytoplankton (p , $\mu\text{g C L}^{-1} \text{ d}^{-1}$) and biological carbon export from the surface ocean by sinking particles (f , $\text{mg C m}^{-2} \text{ d}^{-1}$) – in terms of their probability distributions. Given that p and f are mechanistically linked but variable and effectively measured on different scales, we hypothesize that a quantitative relationship emerges between *collections* of the two measurements. Motivated by the many sub-processes driving production and export, we take as a null model that large-scale distributions of p and f are lognormal. We then show that compilations of p and f measurements are consistent with this hypothesis. The compilation of p measurements is extensive enough to subregion by biome, basin, depth, or season; these subsets are also well-described by lognormals, whose log-moments sort predictably. Informed by the lognormality of both p

and f we infer a statistical scaling relationship between the two quantities and derive a linear relationship between the log-moments of their distributions. We find agreement between two independent estimates of the slope and intercept of this line and show that the distribution of f measurements is consistent with predictions made from the moments of the p distribution. These results illustrate the utility of a distributional approach to biogeochemical fluxes. We close by describing potential uses and challenges for the further development of such an approach.

3.1.2 Introduction

Two of the most important processes in the carbon cycle are net primary production by phytoplankton (NPP; here we use p [$\mu\text{g C L}^{-1} \text{ d}^{-1}$]) and the export flux of sinking biogenic particles out of the upper ocean (f [$\text{mg C m}^{-2} \text{ d}^{-1}$]) [Williams and Follows, 2011]. p is typically defined either as the amount of photosynthetically fixed carbon available to first-level heterotrophs in an ecosystem [Osmond, 1989] or as the difference between autotrophic photosynthesis and respiration [Lindeman, 1942]. f is preferably defined as the downward flux of particulate organic carbon (POC) at the base of the euphotic layer z_{eu} [Buesseler and Boyd, 2009], but operationally is often defined as the flux through a particular depth, e.g. 100m, because most measurements of POC flux historically have not been taken at z_{eu} .

Much remains unknown about the variability of p and f . Both p and f are influenced by a myriad of biological, physical, and chemical processes. Measurements of each range over several orders of magnitude [Buitenhuis et al., 2013, Buesseler and Boyd, 2009], and exhibit substantial variability. The interpretation of the most common method of measuring p , based on radiocarbon uptake, is debated [Marra, 2009], and f measurements are notorious for having large uncertainties [Buesseler, 1991]. However, in order to gain a thorough understanding of production and flux we must characterize their variability.

The relationship between f and p is also of interest; mass conservation suggests an inextricable dependence of f on p as all of the sinking material that comprises f must previously be fixed via p . However, because many other fluxes, e.g. heterotrophic respiration, affect the organic carbon balance of the surface ocean, their quantitative relationship is not immediate [Williams and Follows, 2011]. Multiple relationships have been proposed specifying f as a function of local depth-integrated production $\mathcal{P}$ [$\text{mg C m}^{-2} \text{ d}^{-1}$] [Eppley and Peterson, 1979, Dunne et al., 2005, Maiti et al., 2013, Laws et al., 2000a], often a scaling

Figure 3-1: Scatterplot of particulate organic carbon flux (f) at 100m and depth-integrated net primary production by phytoplankton ($\mathcal{P}$) data from [Maiti et al., 2013]. Notice the lack of a clear correlation between the two. Metrics reported are from Pearson’s correlation: The fraction of variance of explained (R^2) and the p value of the correlation. Compare to Figure 3-10.

relationship with an error term which accounts for processes the model does not resolve. However, direct comparison of these two measurements is challenging; these relationships have indicated $\mathcal{P}$ has little predictive power for f . Figure 3-1 shows Southern Ocean f and $\mathcal{P}$ data compiled by Maiti et al. [Maiti et al., 2013], whose correlation is non-significant at the 0.05 level [Figure 3-1], illustrating the difficulty in understanding the relationship between f and p – though other datasets have suggested stronger relationships [Dunne et al., 2005] (see Section 3.1.16)

This lack of predictive power is in no small part because p and f can become locally decoupled, i.e. not consistently comparable pointwise [Buesseler, 1998]. p is measured on the scales of liters and days; f is effectively measured on the spatial scale of kilometers and/or the timescale of weeks (which due to mixing also implies far larger spatial scales –

also see Sections 3.1.5 and 3.1.7) [Siegel and Deuser, 1997, Buesseler, 1991, Marra, 2009]. The variable nature of both p and f therefore implies that a given $\mathcal{P}$ measurement can be a poor measure of the productivity occurring in the waters sampled by a co-located f measurement.

Given the mechanistic relationship between p and f , it is nonetheless reasonable to expect that a quantitative relationship between measurements of the two should emerge at larger scales. In other words, even if p and f can become locally decoupled, it is still plausible that the p characteristics of a region could be predictive of the f characteristics of that region. Here we adapt previous approaches quantitatively relating p and f to larger scales by analyzing collections of measurements – describing these in terms of their probability distributions and then inferring a relationship between the two quantities via their distributions’ moments.

One way to understand variable processes such as p and f is through their probability distributions. Rather than trying to predict the value of every measurement, can the underlying distribution that the values are ‘drawn’ from be understood? How does the distribution (or its parameters) differ for different environmental conditions? Can two quantities be related not by a pointwise function but via their probability distributions? Here, we develop the basics of such an approach for p and f [Figure 3-2]. We hope our efforts can provide direction towards a more complete theory.

We argue that the lognormal distribution is a natural null model for both p and f on large scales. We then show that an extensive compilation of ^{14}C measurements for p [Buitenhuis et al., 2013], and a large compilation of sediment trap and ^{234}Th measurements for f (Section 3.1.7) are well-described by lognormal distributions. Subregions of the p compilation are also well-described by lognormals, whose moments sort according to oceanographic intuition, indicating that the lognormal is a general and robust feature of large-scale p variability. In agreement with recent models for the relationship between f and $\mathcal{P}$ [Laws et al., 2000a, Maiti et al., 2013, Britten et al., 2017], we posit a statistical scaling relationship between f and p , noting that it is the only relationship mapping one lognormal into another [Campbell, 1995]; we then demonstrate that the moments of the two distributions should be linearly related via this scaling relationship. We then find agreement between two independent approaches to estimate the scaling relationship’s parameters – one an out-of-sample test utilizing the log-means of p and f from three open ocean time series stations, and the other subregioning

Figure 3-2: Schematic of approach: p and f data from the global ocean are grouped into probability distributions and the probability distributions are then analyzed. Production and flux are then related to each other (and to other quantities) via the moments of their probability distributions.

the p and f data into three biomes, following [Banse, 1992]. We show that the prediction made by this scaling relationship agrees with the log-moments of the global distributions p and f , and with the log-standard deviations of the time series' and biomes' distributions. We close by discussing advantages, potential uses, and limitations of a distributional approach.

3.1.3 Theory: the lognormal distribution as a null model

What distribution might be expected to underlie p or f ? The processes which affect the distributions are both structured and stochastic. Structured processes are governed by large-scale geophysical fluid dynamics and climate, while stochastic processes are inherently variable. For instance, the average nutrient supply to the euphotic layer over a region should directly shift the mean of the distribution for p or f , while turbulence can drive local fluctuations in nutrient supply [Falkowski and Ziemann, 1991]. Here, we explore what is expected when the distributions for p and f are driven by a combination of many processes, both structured and stochastic.

We begin with the observation that both p and f are complex processes, occurring as the result of a number of sub-processes. In the case of p , a single carbon fixation event (i.e. carbon being fixed by phytoplankton at a particular place and time) requires the occurrence of a number of sub-events: the presence of a phytoplankton, the absence of a predator that would consume the phytoplankton, the presence of light and various nutrients at suitable concentrations, and the presence of appropriate temperature and physical conditions [Eppley, 1972, Marra, 1978, Sverdrup, 1953, Durham et al., 2013].

In the case of f , a single export event (in the microscopic sense, i.e. a particle sinking out of the upper ocean) requires that a particle be generated by p (and thus all of the above sub-processes), that it not be consumed or remineralized by heterotrophic processes [Steinberg et al., 2008], that it not be deflected by flow [Siegel and Deuser, 1997], that it not be disintegrated by shear processes [Alldredge et al., 1990] and that it be of sufficient density and size to overcome the viscosity of water and sink [Smayda, 1970, Alldredge and Silver, 1988, Jackson and Burd, 1998]. Thus, we can consider the completion of a p or f event as dependent on the completion of many sub-events. Given the general description of p and f as being composed of many sub-processes, the natural candidate for the probability distribution of both is the lognormal distribution. By the Central Limit Theorem [Shockley, 1957, Montroll and Shlesinger, 1982], the probability of completing any task that relies on

the successful completion of many sub-tasks is lognormal. To illustrate, we adapt an example from Shockley [Shockley, 1957, Montroll and Shlesinger, 1982], who used a similar model to explain the observed lognormal distribution of papers published by researchers. Let the fixation of carbon require the occurrence of n sub-processes (in a general sense, including e.g. the absence of predation), each of which occur at a given place and time with probabilities $q_1, q_2, \dots, q_n$. The probability Q that a phytoplankton fixes carbon at that place and time is then their product: $Q = q_1 q_2 \dots q_n$. Therefore, taking the log of both sides,

$$\ln Q = \ln q_1 + \ln q_2 + \dots + \ln q_n \quad (3.1)$$

so by the Central Limit Theorem, the distribution function of $\ln Q$ (and therefore the sum, i.e. $\ln p$), is the normal distribution, irrespective of the individual sub-processes. Hence, given a few weak conditions [Montroll and Shlesinger, 1982] (see Section 3.1.17), the distribution function for p should be lognormal.

Equivalently, lognormal distributions emerge from a large number of variable, multiplicative factors [Limpert et al., 2001], so the lognormal may be thought of as resulting as the product of variable environmental and compositional effects. The lognormal distribution, which has the form

$$\mathcal{N}(\ln x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad (3.2)$$

is agnostic of the underlying sub-processes, and reduces all of the complexity into two parameters – the log-mean μ and the log-standard deviation σ – providing a simple way to describe the quantity in question¹.

Thus we use the lognormal as a null hypothesis that p and f occur as the result of many distinct requirements, or equivalently that they are influenced by many variable compositional and environmental factors. We note that f involving additional factors does not necessarily imply the σ of f 's distribution will be larger because of possible correlations between these factors; for instance, one might expect a higher abundance of detritivores (who consume material that could otherwise contribute to f) in areas where sinking particles are more abundant.

The lognormal distribution has enjoyed some utility in ocean ecology and carbon cycling.

¹Note μ and σ have the log- and log-square-units of the quantity x ; these are omitted below.

Some optical variables like chlorophyll appear to be lognormally distributed, and it has been recognized that these distributions could be useful for estimating primary productivity from satellite data [Campbell, 1995]. In a terrestrial system Forney and Rothman [Forney and Rothman, 2012] showed decomposition rates of organic carbon followed a lognormal distribution. In the surface ocean, ecological abundance data is also lognormal [Luo and Others, 2012]. When biology is involved, one often encounters the lognormal distribution [Koch, 2005].

3.1.4 Theory: relating export and production by their distributions' moments

Relating f to p via their probability distributions can circumvent the decoupling problem discussed in Section 3.1.2, because measurements are no longer related pointwise; the question becomes how the distribution of f values is related to the distribution of p values. If we hypothesize that p and f are lognormally distributed, can we describe the parameters of f 's distribution as a function of the parameters of p 's distribution?

As previous studies have posited a scaling relationship between export efficiency ef and depth-integrated production $\mathcal{P}$ [Laws et al., 2000a, Maiti et al., 2013], and a scaling relationship is consistent with p and f being lognormally distributed (see Section 3.1.18), we hypothesize a scaling relationship between f and p , i.e.

$$f = e^C p^\alpha \quad (3.3)$$

where C and α are constants², e is the base of the natural logarithm, and the equality holds in a statistical sense. If p is lognormally distributed according to

$$\ln p \sim \mathcal{N}(\mu_p, \sigma_p) \quad (3.4)$$

and f scales with p according to Eq. (3.3), this implies f is lognormally distributed according to

$$\ln f \sim \mathcal{N}(\alpha\mu_p + C, \alpha\sigma_p) \quad (3.5)$$

²As a conversion factor, C has units of log-km.

which therefore implies the log-moments of f and p are related according to

$$\mu_f = \alpha\mu_p + C, \quad \sigma_f = \alpha\sigma_p \quad (3.6a,b)$$

A full derivation of the above is given in Section 3.1.18. Estimates of C and α can thus be used to predict the log-moments of f (and hence the distribution of f) from those of p . Because μ is a lower-order moment than σ , it can be estimated more accurately [Flannery et al., 1992]; therefore in Section 3.1.11 we estimate C and α via Eq. (3.6a).

3.1.5 Compilation of production measurements

The most common *in situ* measurements of p are those taken by the ‘ ^{14}C method’, which form much of the basis of our understanding of p . The method is discussed comprehensively elsewhere [Peterson, 1980, Marra, 2009, Pei and Laws, 2013]; in brief, p is inferred from differences in light and dark bottles incubated with isotopically labeled (^{14}C) carbon dioxide. The principal issues discussed in the literature are the interpretation of metabolic terms and the accuracy with which *in vitro* measurements reflect *in situ* conditions. Individual sample errors are less concerning when considering logarithmic probability distributions. Additive errors would yield significant measurement artifacts for small p samples, but increasingly negligible ones for larger p values; multiplicative errors would affect estimates of σ but not the shape of the probability distribution itself. In any case, as long as natural variability is larger than measurement error, it is a plausible assumption that the statistics of ^{14}C measurements reflect those of p itself.

Recently, Buitenhuis et al. [Buitenhuis et al., 2013] compiled an extensive database of 50,050 measurements of p via the ^{14}C method, spanning the ocean in time, depth, and latitude-longitude. This provides an excellent database with which to test the lognormal null model. See [Buitenhuis et al., 2013] for a complete description of the compilation. We augment this compilation with data from the CARIACO time series (available at <http://www.imars.usf.edu/cariaco>; accessed 7 June 2017) in order to compare with the f data from that time series in Section 3.1.11. To minimize the effect of additive errors, we compare the lognormal distribution to the empirical distribution of p measurements above a small threshold of $0.5 \mu\text{g C L}^{-1} \text{ d}^{-1}$, or 10% of the distribution’s peak height in log-space, which we determine by a relative residual error analysis (see Section 3.1.19). This operation

excludes 16% of the observations where $p > 0$, but the excluded observations account for $< 0.01\%$ of the total primary production in the database. For consistency we use the same threshold for all analyses of p data throughout the chapter. Figure 3-3 shows the locations³ of these p measurements. Our results are not sensitive to factor of two changes in this threshold.

3.1.6 Subregions of production measurements

This total set of p measurements is large enough to subregion in time, depth, and latitude/longitude. We test four ways of coarsely subregioning the p data, noting that many other, more sophisticated subregioning schemes are possible [Longhurst, 1998]. We do so to test for robustness of the lognormal and to address the issue of non-random sampling. If the lognormality of the total p distribution is a coincident result of the spatiotemporal structure of p measurements, subregions' p distributions should not be lognormal. We also do this to test whether the the log-moments of the subregions' p distributions sort predictably (see Section 3.1.10).

1. Biome. Following Banse [Banse, 1992], who defined three ocean ‘Domains’ – respectively low-seasonality, nutrient depleted (oligotrophic), low-seasonality, nutrient-replete (eutrophic), and high-seasonality (seasonal) – we generate latitude-longitude subregions by objectively identifying three biomes using a $\mathcal{P}$ climatology from the Carbon-based Productivity Model [Westberry et al., 2008]⁴ The ‘Seasonal (B3)’ biome includes all locations where the log-variance of climatological $\mathcal{P}$ is $>0.3 \text{ mg C m}^{-2} \text{ d}^{-1}$; the ‘Eutrophic (B2)’ biome includes all remaining locations where the log-mean of climatological $\mathcal{P}$ is $>6.5 \text{ mg C m}^{-2} \text{ d}^{-1}$; the ‘Oligotrophic (B1)’ biome includes all other locations. Threshold values were chosen to correspond with the ‘Domains’ drawn in [Banse, 1992]; biomes are shown in Figure 3-3 (cf. Figure 1 of [Banse, 1992]). Our results are not sensitive to the values of these thresholds.
2. Basin. We generate three latitude-longitude subregions by splitting the data by basin into Atlantic, Southern Ocean, and Indo-Pacific [Figure 3-3].

³Note the data in Figure 3-3 are fairly evenly distributed seasonally.

⁴available at <http://www.science.oregonstate.edu/ocean.productivity/standard.product.php>.

Figure 3-3: Top: Locations of p samples $> 0.5 \text{ } \mu\text{g C L}^{-1} \text{ day}^{-1}$ from the compilation by [Buitenhuis et al., 2013]; compare to their Figure 2a. The solid black boundary line denotes the separation of data into ocean basins (Section 3.1.6). The time series are indicated by crosses. Middle: Locations of f samples $> 3.6 \text{ mg C m}^{-2} \text{ day}^{-1}$ from the compilation. The time series are indicated by crosses. Bottom: Map of biomes into which p and f data are sub-regioned; compare to Figure 1 of [Banse, 1992]. Biomes are defined as B1: low-seasonality, nutrient-depleted (oligotrophic), B2: low-seasonality, nutrient-replete (eutrophic), B3: high-seasonality (seasonal), using a depth-integrated net primary production climatology from the Carbon-based Productivity Model [Westberry et al., 2008].

3. Depth. We generate three depth subregions by sorting p values by the depth at which they were sampled and splitting the data evenly into $<6\text{m}$, $6\text{--}20.5\text{m}$, and $>20.5\text{m}$ (i.e. 6m is the 33.3rd percentile of the sampling depths, and 20.5m is the 66.7th percentile).
4. Time. We generate four seasonal subregions by taking all data from the Northern Hemisphere and splitting them into boreal ‘winter’ (DJF) ‘spring’ (MAM), ‘summer’ (JJA), and ‘autumn’ (SON), where e.g. DJF refers to December, January, and February.

3.1.7 Compilation of export measurements

The most common measurements of f are those using sediment traps [Honjo et al., 2008], which collect sinking material, and those using the ‘ ^{234}Th method’ [Buesseler and Others, 2006], which measures disequilibrium of the particle reactive ^{234}Th isotope as a chemical signature of particulate flux.

We compiled a database of 1,770 shallow ($\leq 200\text{m}^5$) POC export flux measurements by trap and ^{234}Th spanning the global ocean; see Figure 3-3. The database comprises 49% ^{234}Th measurements. This database is the largest available compilation of shallow POC export flux measurements of which we are aware. For accessibility, it is organized according to the principles of ‘tidy data’ [Wickham, 2014] and is available online [Cael et al., 2018]. Note that the f compilation is smaller than even any of the subregions of the p compilation because f measurements are far less abundant. Because resolving a distribution with large dynamic range requires a large sample size, we include as many measurements as possible, across two methods and multiple depths⁶; as with p , we assume that the distribution of measurement values is reflective of the distribution of the process being measured. To ensure data fidelity and to maximize the number of available observations no adjustments to the data were conducted (e.g. normalizing to a particular depth, averaging by location or time, choosing cutoffs for trap duration or trap type, and so forth). About 1/3 of our f data comes from three time series – HOT (the Hawaii Ocean Time series) [Karl and Lukas,

⁵200m depth and shallower, a depth sufficient to capture instances of deep euphotic and/or mixed layers – most measurements were at 100m or 150m

⁶Comparison of the probability distributions of data from 100m vs. 150m, and of the thorium data vs. the trap data, indicates that conflating measurements across sampling depths or methods did not affect our results. Thorium- vs. trap-derived measurements often disagree at individual sites, but differences may average out for groups of measurements, or systematic biases may be small relative to the measurement variability and sample size here.

1996], BATS (the Bermuda Atlantic Time series Study) [Michaels and Knap, 1996], and CARIACO (CARbon Retention In A COlored Ocean) [Thurnell, 2016] – and a large number of p samples (> 1000) are also available from each of these time series; we therefore exclude these from the global distribution in order to make an out-of-sample estimate for C and α in Section 3.1.11.

To minimize the effect of additive errors and to analyze both p and f data the same way, we compare the lognormal distribution to the empirical distribution of f measurements above a noise threshold of 10% of peak height in log-space, or $3.6 \text{ mg C m}^{-2} \text{ d}^{-1}$. As with the p data, for consistency we use the same noise threshold for all analyses of f data throughout the chapter, and our results were not sensitive to factor of two changes in this threshold.

Figure 3-3 indicates that many of the f samples are taken at distinct locations from the p samples. As we are comparing collections of measurements, our approach requires the assumption that the variability sampled by a collection of measurements is large compared to the bias generated by those measurements’ spatiotemporal structure. This assumption is supported for large-scale sets of measurements by the robustness of p ’s lognormality (see Section 3.1.9), but means that a lognormal is not expected to emerge when this assumption breaks down, e.g. when looking at measurements from a single location, or when sampling across boundary currents and other areas with large physical gradients.

3.1.8 Statistical methods

In total we have fourteen sets of p measurements and one set of f measurements. We fit a lognormal distribution to each of these, estimating the log-moments by minimizing Kuiper’s statistic, which measures deviations between the hypothesized lognormal distribution and the empirical cumulative distribution function (CDF) of the measurements. Kuiper’s statistic is preferred in many applications because it balances ease of interpretation with sensitivity to tails [Flannery et al., 1992], though our results are not sensitive to this choice of statistic as compared to the Kolmogorov-Smirnov and Anderson-Darling statistics (see Section 3.1.20).

We estimate uncertainties in the estimated moments by bootstrap sampling the p and f subsets 10,000 times and repeating the procedure in each case [Efron, 1979]. Note that this procedure underestimates uncertainty as it only accounts for statistical errors. Systematic

uncertainty including measurement errors, interpretation errors, and non-randomness of the sampling are not accounted for. To be conservative we use a uniform uncertainty of 0.04 for μ_p and 0.03 for σ_p corresponding to the largest 95% bootstrap confidence interval across all subsets, that of the winter data.

3.1.9 Production and export are log normally distributed

We find that the lognormal null model is a good description of the distribution for p globally; Figure 3-4 shows good correspondence between the data and hypothesized distribution. We find moments of $\mu_p = 2.23 \pm 0.04$ and $\sigma_p = 1.63 \pm 0.03$. We also find that the lognormal null model is a good description of the distribution for f globally; see Figure 3-4. We find moments of $\mu_f = 3.96 \pm 0.07$ and $\sigma_f = 1.20 \pm 0.05$. While the distribution's fit to the f data is less visually compelling, this is more than accounted for by sample size (see Section 3.1.20).

Additionally, we find that for the subregions' p distributions, the lognormal model is a good description of the distributions; see Figure 3-5. We find μ_p ranges from 1.13 to 2.86 (± 0.04), and σ_p ranges from 1.20 to 1.69 (± 0.03). The robustness of the lognormal across the various subregions strongly supports its applicability to modeling p from a large-scale perspective, and indicates that the lognormality of p is robust to the spatiotemporal structure of measurements.

3.1.10 Log-Moments of subregions' production distributions sort predictably

Figure 3-6 shows the μ_p and σ_p of the lognormal distributions shown in Figure 3-5. Because of the additional uncertainties mentioned at the end of Section 3.1.8, we restrict ourselves to the sorting of moments across subregions. Variation in σ_p is 7% of the variation in μ_p across subregions, and σ_p is a higher-order moment and therefore more difficult to estimate accurately, so we focus on variation in μ_p .

The ordering of μ_p across subregions is sensible in terms of light and nutrient availability. μ_p is lower in the oligotrophic biome (B1) than in the eutrophic and seasonal biomes (B2 and B3), whose μ_p are not significantly different. μ_p is lower in the Indo-Pacific, the basin having the largest proportion of low-productivity waters [Williams and Follows, 2011], than in the Atlantic or Southern Ocean. The large μ_p for the Southern Ocean reflects that the

Figure 3-4: Top: PDF of global p data versus lognormal fit. $n = 38,334$ refers to the number of samples included in the PDF, which excludes data from HOT, BATS, CARIACO, as well as data below the 10% of peak height threshold $< 0.5 \mu\text{g C L}^{-1} \text{ day}^{-1}$. Green curve is the probability density function of the lognormal which minimizes the Kuiper statistic as compared to the data. Bottom: PDF of global f data versus lognormal fit. $n = 1,033$ refers to the number of samples included in the PDF, which excludes data from HOT, BATS, and CARIACO time series, as well as data below the 10% of peak height threshold of $3.6 \text{ mg C m}^{-2} \text{ day}^{-1}$. Green curve is the probability density function of the lognormal which minimizes the Kuiper statistic as compared to the data. HOT = Hawaii Ocean Time; BATS = Bermuda Atlantic Time series Study; CARIACO = Carbon Retention In A Colored Ocean time series; PDF = probability density function.

Figure 3-5: Probability density functions (PDFs) of p data for different subregions. In each case the green curves are the PDF of the lognormal which minimizes the Kuiper statistic as compared to the data. (μ, σ) is plotted in Figure 3-6.

Figure 3-6: Log-moments for each of the p subregions. Error bars are 95% confidence intervals estimated from bootstrapping; contours are lines of constant mean for a lognormal distribution, equal to $\exp(\mu + \frac{1}{2}\sigma^2)$.

Southern Ocean p data are dominantly sampled from, and therefore only representative of, high-productivity times of year in that basin. μ_p is highest in the shallowest depth subregion, where light is highest, and is lowest in the deepest subregion, where light is lowest. μ_p is higher in spring and summer, when the Northern Hemisphere experiences more sunlight.

The lognormality of p has implications for the interpretation of variability in p measurements, because mean and variance are controlled both by σ and μ . Note the contours in Figure 3-6; the mean (n.b. not the log-mean) of a lognormal distribution is given by $\exp(\mu + \frac{1}{2}\sigma^2)$, so the two moments can compensate to achieve the same mean, e.g. the mean for the Atlantic (μ_p, σ_p) is closer to the mean for the Southern Ocean (μ_p, σ_p) than the individual moments are; the same goes for summer and autumn. Variance (n.b. not the log-variance) of a lognormal is given by $(\exp(\sigma^2) - 1)(\exp[2\mu + \sigma^2])$, so two distributions with the same σ can have different variances; e.g. $\sigma_p = 1.50 \pm 0.03$ for both the Indo-Pacific and the Southern Ocean, but because μ_p is larger for the Southern Ocean, its variance is also higher, the difference driven by μ .

The sorting of σ_p is also consistent with intuition. σ_p is largest in the seasonal biome (Biome 3) and lowest in the oligotrophic biome (Biome 2). σ_p is largest in the Atlantic, the basin with the most pronounced seasonality [Silsbe et al., 2016]. σ_p is larger in boreal autumn and spring, when broadly speaking the Northern Hemisphere's oceans experience larger changes, than in boreal winter and summer. However, because understanding variability in log-space is no trivial matter, and these variations are small compared to those in μ_p , we caution against making too much of the sorting of σ_p barring a more quantitative theory for these moments (see Section 3.1.12).

3.1.11 Export scales sublinearly with production

That f and p are well-described by lognormal distributions implies the statistical scaling relationship $f = e^C p^\alpha$ (Section 3.1.4). This suggests that estimates of C and α should allow one to predict μ_f from μ_p (via Eq. (3.6a)) and σ_f from σ_p (via Eq. (3.6b)). We use two procedures to estimate C and α based on Eq. (3.6a) and the μ 's of independent data collections – the biomes and the time series – then compare these estimates and evaluate i) their ability to predict μ_f from μ_p globally and ii) whether the estimates for α are consistent with those for σ_p and σ_f .

First, we subregion the global f data into the three biomes described in Section 3.1.6.

These data have the advantage of being parsed into and sampled across objective biogeographic regions, but are not co-localized with the p data from the same biomes and are not independent of the global p and f data. We take the (μ_p, σ_p) values from Figure 3-6, and then estimate (μ_f, σ_f) by the log-mean and log-standard deviation of the f data (above the 10% peak height threshold) rather than by fitting a lognormal distribution as for Figures 3-4 and 3-5, though our results are not sensitive to this choice. We then estimate uncertainty in each log-moment for each biome by bootstrap sampling as in Section 3.1.8, repeating the procedure 10,000 times and using the 95% bootstrap confidence interval as an uncertainty estimate. We then estimate C and α by fitting Eq. (3.6a) to the three μ_f, μ_p pairs from the three biomes. We use type II Least Squares Cubic regression [York, 1966] of μ_f on μ_p because each moment estimate has a different, non-negligible uncertainty. We use the subscripts C_b and α_b to indicate the estimates derived from the biomes' data.

Second, we make an out-of-sample estimate with the three time series' sediment trap data excluded from the global distributions, which provide an alternative set of μ_f, μ_p pairs to fit C and α . These data have the advantage of being co-localized, and independent of the global p and f data so as not to confound the prediction, but are sampled only across three total locations. As above, we estimate $(\mu_p, \sigma_p, \mu_f, \sigma_f)$ by the log-moments of the time series' data, estimate uncertainty via bootstrap sampling, and then make estimates C_t and α_t from the time series' μ 's and type II Least Squares Cubic Regression in Eq. (3.6a).

Figure 3-7 shows the histograms of the time series' p and f data and the biomes' f data. We note that the time series' p distributions are visibly (as well as quantitatively) not log-normal, consistent with the idea that the lognormal holds only for large-scale variability, i.e. as one considers larger and/or more variable spatiotemporal regions and therefore samples over more variability in the processes influencing p and f (see Section 3.1.14). The f data in Figure 3-7 are consistent with a lognormal distribution, though the statistical power to compare each collection of data with a hypothesized distribution is limited due to sample size (see Section 3.1.20).

Nonetheless, as each dataset in Figure 3-7 is unimodally distributed in log-space and we focus here on the first moment, it is sensible to approximate each distribution with a lognormal, estimating μ and σ for each by the data's log-mean and log-standard deviation. Probability distributions are defined by the series of their moments, and discrepancies from this approximation will arise only for higher-order moments. That this estimation procedure

Figure 3-7: Histograms of time series' p and f data and biomes' f data. (μ, σ) is plotted in Figure 3-8; p data for the biomes are plotted in Figure 3-5.

ultimately results in a good prediction, despite the time series' distributions p data not being lognormally distributed, evinces the utility and robustness of a distributional approach.

Figure 3-8 shows the results of the regressions, which find $\alpha_t = 0.64 \pm 0.13$, $\alpha_b = 0.65 \pm 0.14$, $C_t = 2.47 \pm 0.17$, and $C_b = 2.59 \pm 0.30$ (where here the $\pm$ is the standard error as estimated by the regression, not a 95% confidence interval). Thus, the parameter estimates are in good agreement. The biomes' parameter estimates result in a prediction for the global $\mu_f = C_b + \alpha_b \mu_p = 4.04 \pm 0.43$, and the time series' parameter estimates result in a prediction of $\mu_f = C_t + \alpha_t \mu_p = 3.90 \pm 0.34$ (where the $\pm$ is the standard error from propagating uncertainty in α and C). Both of these compare well with the estimated value for the global distributions' $\mu_f = 3.96 \pm 0.07$.

The inset of Figure 3-8 shows the results of the second prediction – whether $\sigma_f = \alpha \sigma_p$ for the time series, the biomes, and globally. For the global distributions, the eutrophic and seasonal biomes (B2 and B3), BATS, and CARIACO, σ_f/σ_p is within one standard error of both α_b and α_t ; σ_f/σ_p for HOT is less than α_b and α_t by more than one standard error, and σ_f/σ_p for the oligotrophic biome (B1) is more than α_b by more than one standard error. As the σ values are not used to estimate the α 's, and the σ_f values estimated from the biomes' and the time series' data have larger uncertainty, this correspondence is quite good, corroborating that $f \sim p^\alpha$ and therefore that μ_f and σ_f are predictable from μ_p and σ_p . Variation in σ_f/σ_p could be due to biogeochemical differences between time series, to measurement uncertainties, or to other factors.

3.1.12 What Are the uses and requirements of a more complete theory?

As stated previously, the objective herein is only to develop the foundations of a distributional theory, as a more complete theory involves a good deal more sophistication and would be benefitted by more measurements than are available at present, particularly for f . How to subdivide the ocean is a key question. We have only done so coarsely here as proof of concept and to test the robustness of the lognormal to subregioning. The refinement of this theoretical framework requires first developing a more considered subregioning of the ocean in space and time, i.e. determining an answer to the ubiquitous question in oceanography of what spatiotemporal region a given set of observations represents.

A predictive theory for subregions' μ_p and σ_p must also be developed. We have only attempted to demonstrate that sorting of the μ and σ of subregions' p distributions is

Figure 3-8: μ_f versus μ_p for the three time series, the three biomes, and the global ocean. Dotted blue line is the estimated line from the regression on the biomes ($\mu_f = \alpha_b \mu_p + C_b$); dotted green line is the estimated line from the regression on the time series ($\mu_f = \alpha_t \mu_p + C_t$); parameter estimates from the regressions are reported in the top-left of the figure, along with their estimated standard errors. Solid black lines represent 95% bootstrap confidence intervals. The global distributions' log-means (yellow) are not included in the regressions that produce the dashed lines. Inset: σ_f plotted versus σ_p for the three time series, the three biomes, and the global ocean. Dashed line is the estimate $\sigma_f = \alpha_b \sigma_p$; grey shading corresponds to α_b 's standard error. None of the σ 's are included in the regressions that produce the dashed line.

consistent with oceanographic intuition; we have not developed a quantitative theory for either. p being lognormally distributed implies $\exp(\mu_p) = \text{median}(p)$, so it may be possible to describe μ_p in terms of that which controls median productivity; σ_p may possibly be described in terms of the variances of the sub-processes distributions. As the lognormal distribution is controlled by these two parameters, a complete distributional theory for p requires the ability to predict these two parameters, i.e. how the distribution shifts in response to the environment. Such a refinement could constitute a highly simplified yet accurate description of p globally.

Due to its relative sparsity, subregioning f data substantially reduces the statistical power one has to test f 's lognormality. While an f database of comparable size and methodological consistency to that for p compiled by Buitenhuis et al. [Buitenhuis et al., 2013] is far from reach, a large number of additional measurements are expected to be available in the near future from EXPORTS [Siegel and Others, 2016] and other programs, which may be leveraged to evaluate the robustness of f 's lognormality to subregioning and the consistency of the scaling relationship inferred here. A distinct limitation of the distributional approach presented here is that it requires a large number of measurements.

Finally, given that the two procedures for estimating C and α result in very similar values despite using distinct data, this raises the interesting question of what physical, chemical, and biological processes might control these parameters.

3.1.13 Can the probabilistic approach developed here be used to test models?

Satellite ocean color models and numerical biogeochemical models of p and f do not contain spatiotemporal information explicitly – if they use such information, it is to determine some other variable, e.g. solar irradiance. In principle, the probability distribution of a variable contains all of the non-spatiotemporal information about that variable. Thus, the strictest possible non-spatiotemporal criterion by which one can evaluate numerical biogeochemical models of p or f is by evaluating the probability distributions they generate. Rather than comparing model values and data measurements pointwise, one can sample a model at the same places and times as a dataset and compare the resulting probability distributions. For models of p , the null model described herein can be utilized as a strict test. From model output, i) objectively generate regions based on biogeochemically relevant

parameters, ii) determine the probability distributions of production (and export) in these regions, and iii) analyze the forms of the distributions and the behavior of their log-moments. Do spatiotemporal subregions of the model produce lognormal distributions, and do the moments of these subregions sort predictably? Such approaches are underutilized in Earth sciences, and may be especially useful in oceanic contexts because slight misfits in the locations of fronts and other features would not be over-penalized. In this chapter we have not analyzed measurements of depth-integrated production, the quantity of most interest for satellite ocean color algorithms [Behrenfeld and Falkowski, 1997, Westberry et al., 2008, Silsbe et al., 2016]; depending on how a given algorithm incorporates depth integration, comparison may require augmenting the lognormal null model to account for integration in depth. Nonetheless, it appears that a distributional approach could be useful not only for the analysis of p and f measurements but also as a more rigorous methodology of model-data comparison.

3.1.14 What about the spatiotemporal structure of the surface ocean?

Although the processes underlying production and flux suggest a lognormal distribution, the spatiotemporal structure of the surface ocean suggests otherwise. If p and f were not at all stochastic quantities, and were entirely determined by the mean state when and where they were measured with negligible variation, or if p and f measurements were strongly correlated in space or time, variability would not necessarily be lognormal; the observed distribution of p and f would then instead be wholly determined by the places and times at which those measurements were made. Even if p and f values in the ocean were truly lognormal, available measurements of p and f are far from a truly random sample [Figure 3-3], which could affect their distribution similarly. However, given sufficient underlying complexity and variability in the processes that result in p or f , a lognormal distribution will emerge; μ and σ of that distribution will then be reflective of the sample locations and times, e.g. p data from higher productivity areas will have a larger μ . We therefore consider the lognormal distribution to characterize *large-scale* variability in p and f , to hold when the spatiotemporal region being sampled contains enough variability in the factors p and f that a lognormal distribution emerges. The robustness of the lognormal distribution across subsets of p measurements, despite large spatial patterns in p across the ocean and the highly nonrandom nature of global p sampling, suggests that substantial multiplicative

variability is a fundamental characteristic of p .

3.1.15 Conclusion

Understanding the and relationships between biogeochemical fluxes like primary production and export is difficult because there are many distinct sub-processes which play a role. Under some circumstances systems become complex enough that they can become simple again if viewed through their probability distributions rather than sets of co-localized measurements. We suggest that this is the case for primary production and export. We demonstrate that measurements of these two quantities are lognormally distributed and that the moments of the distributions are a linearly related to one another. With additional data and a clearer understanding of how to define biogeochemical provinces we believe this approach will help us build a mechanistic understanding of these, and other biogeochemical fluxes in the sea.

3.1.16 Re-examining efficiency-production relationships

We re-examine two posited relationships for the relationship between ef , the fraction of primary production exported out of the surface ocean, and depth-integrated production ($\mathcal{P}$, $\text{mg C m}^{-2} \text{ d}^{-1}$) [Laws et al., 2000a, Maiti et al., 2013]. In brief, both are based on analyses of compilations of coincident f and $\mathcal{P}$ measurements; ef is defined as $ef = f/\mathcal{P}$, and the relationship between ef and $\mathcal{P}$ is analyzed; Maiti et al. [Maiti et al., 2013] find an inverse relationship, $ef \sim \mathcal{P}^{-1}$, for $(F, \mathcal{P})$ data from the Southern Ocean, while Laws et al. [Laws et al., 2000a] find a sublinear relationship $ef \sim \mathcal{P}^{0.3}$ for $(f, \mathcal{P})$ data from the global ocean. Note that the relationship between f and p we describe is not directly comparable with the two relationships considered here (we do not consider $\mathcal{P}$); a comparison would require an analysis of the relationship between p and $\mathcal{P}$.

We examine the Maiti et al. [Maiti et al., 2013] relationship first. Maiti et al. [Maiti et al., 2013] find an inverse relationship between ef and $\mathcal{P}$ data from the Southern Ocean. Substituting the definition of ef as $f/\mathcal{P}$, this relationship becomes $f/\mathcal{P} \sim 1/\mathcal{P}$. It is known that ratios of uncorrelated variables can produce spurious correlations; in particular, for two uncorrelated variables A and B , it has been demonstrated in the aquatic sciences literature that comparing their ratio AB^{-1} with B can produce spurious inverse relationships and correlations [Berges, 1997]. This suggests that the measurements of f and $\mathcal{P}$ in their data compilation could be unrelated. We find this to be the case; see Figure 3-1. The Pearson

Equation	Parameters	Ind. vbls.	r^2
$ef = \frac{(a-bT)\mathcal{P}}{c+\mathcal{P}}$	fixed	$T, \mathcal{P}$	0.16
$ef = \frac{(a-bT)\mathcal{P}}{c+\mathcal{P}}$	free	$T, \mathcal{P}$	0.47
$ef = \frac{a\mathcal{P}}{b+\mathcal{P}}$	free	$\mathcal{P}$	0.07
$ef = (a - bT)\mathcal{P}^c$	fixed	$T, \mathcal{P}$	-0.59
$ef = (a - bT)\mathcal{P}^c$	free	$T, \mathcal{P}$	0.47
$ef = a\mathcal{P}^b$	free	$\mathcal{P}$	0.07
$ef = a - bT$	free	T	0.45

Table 3.1: Summary of analyses of Dunne et al. [Dunne et al., 2005] data. First column is the model equations used to predict ef from $\mathcal{P}$ and/or T ; equations are those considered by Laws et al. [Laws et al., 2000a] or simplifications thereof. Second column indicates whether the parameters in the equation were fit to the data (free) or if their values were taken from Laws et al. [Laws et al., 2000a] (fixed). Third column indicates which variable(s) ef is a function of in each case. Fourth column indicates the percentage of variation in ef explained by the independent variables via each equation.

correlation of these data is not significant at the 95% confidence level.

We then turn to the relationships posited by Laws et al. [Laws et al., 2000a]. Laws et al. [Laws et al., 2000a] analyzed a dataset of coincident f , $\mathcal{P}$, and temperature T ($^{\circ}\text{C}$) measurements compiled from locations across the global ocean by Dunne et al. [Dunne et al., 2005]. Table 3.1 summarizes our findings, which are that in the Dunne et al. [Dunne et al., 2005] data, $\mathcal{P}$ is too weakly related to ef to extract a meaningful relationship, and that the relationships posited by Laws et al. [Laws et al., 2000a] appear artificially strong in their analyses either because they only consider a subset of the data or because they fit binned data, which is known to spuriously amplify relationships [Figure 3-9] [Wainer et al., 2006].

Laws et al. [Laws et al., 2000a] posited two relationships for ef as a function of $\mathcal{P}$ and T (their equations 2 and 3). However, one relationship was only tested on a subset of the dataset analyzed, and the other was only tested on binned data; thus these analyses do not address the question of how well the ef data are explained by the $\mathcal{P}$ and T data.

We fit the equations reported by Laws et al. [Laws et al., 2000a], with the parameters reported in Laws et al. [Laws et al., 2000a], to the data analyzed in Laws et al. [Laws et al., 2000a] (compiled by Dunne et al. [Dunne et al., 2005]). This yielded $r^2 = 0.16$ and $r^2 = -0.59$, respectively (a negative r^2 indicates a worse fit than assuming a constant ef). When

Figure 3-9: Illustration of spurious relationships. a) scatterplot of x and y , two uncorrelated random variables sampled from a uniform distribution ($n = 200$). b) y/x plotted against x (dashed line is $1/x$; compare to Maiti et al. [Maiti et al., 2013]). c) X and Y are binned averages of x and y , into five bins; dashed line is two-parameter fit ($Y = 0.45; X^{-0.08}; r^2 = 0.99$); compare to Laws et al. [Laws et al., 2000a]. $1/8$ random sample sets will have a monotonic relationship such as the one seen here if binned into five bins ($2^{-4} = 1/16$ will be monotonically increasing and $1/16$ will be monotonically decreasing; this number increases rapidly even in response to very weak correlations, e.g. $\rho = 0.01$).

the same equations are fit without fixing the parameters, both have an $r^2 = 0.47$. However, both include a linear dependency on temperature which alone (i.e. neglecting dependency on $\mathcal{P}$; see Table 3.1) yields an $r^2 = 0.45$, while neglecting the temperature dependency in each equation (see Table 3.1) yields $r^2 = 0.07$. Indeed, Figure 3-10 demonstrates this straightforwardly, indicating that $\mathcal{P}$ can explain little of the variation in ef in these data.

Therefore it appears that in the data analyzed by Laws et al. [Laws et al., 2000a], there is no useful relationship between ef and $\mathcal{P}$, and that instead the correspondence between independent and dependent variables for their equations 2 and 3 is due to a combination of i) the dependence of ef on T , ii) considering only a subset of the data, and/or iii) analyzing binned data. Note that we do not take issue with their choice to analyze a subset of the data (see their text for details); we merely intend to point out that their equation 2 should not be taken as a global relationship for ef and $\mathcal{P}$ given the above considerations. The lack of a clear relationship between ef and $\mathcal{P}$ could be due to similar reasons as in the Maiti et al. [Maiti et al., 2013] data, or for other reasons. Our intention herein is not to explain the relationships (or lack thereof) between these variables, but rather to demonstrate that the existing literature on the relationship between ef and $\mathcal{P}$ does not provide satisfactory relationships between production and efficiency (and therefore production and export flux). Figure 3-9 shows how spurious relationships can emerge from

Figure 3-10: $\mathcal{P}$ and ef data from Laws et al. [Laws et al., 2000a], originally compiled by Dunne et al. [Dunne et al., 2005]. Metrics reported are from Pearson's correlation: the fraction of variance of explained (R^2) and the P-value of the correlation.

Figure 3-11: PDF of sums of triplets drawn from the unit uniform distribution, with normal distribution overlaid, illustrating the rapid convergence of the Central Limit Theorem.

comparing ratios of uncorrelated variables or from binning data.

3.1.17 Convergence of the Central Limit Theorem

We briefly give an example of how the convergence of the Central Limit Theorem [Pitman, 1999] can be surprisingly rapid.

We take a $100,000 \times 3$ matrix where each value is sampled uniformly from the unit interval $[0, 1]$, then sum each row, thereby generating 100,000 samples of x_i , where each x_i is the sum of three draws from a uniform distribution (a distribution quite unlike the normal distribution!). Figure 3-11 shows the PDF of x_i , which looks strikingly like a normal even for the very small value of $n = 3$. Of course, the exact value of n where convergence occurs will depend on many things, i.e. the distributions of the numbers being summed over, correlations in those numbers, the criterion whereby the distribution is considered converged, & so forth, but this gives a demonstration of the nonintuitive rapidity of the Central Limit Theorem's convergence.

3.1.18 Derivation of log-moment relationships

Let $f = e^C p^\alpha$, which after taking the logarithm of both sides can be written

$$\ln f = C + \alpha \ln p \quad (3.6)$$

where the equality is taken in the statistical sense. Because $\ln f$ can be expressed as a continuous monotonic function of $\ln p$, their probability distributions are related through their differentials,

$$P_f(\ln f) d \ln f = P_p(\ln p) d \ln p \quad (3.7)$$

Dividing both sides by $d \ln p$,

$$P_f(\ln f) \frac{d \ln f}{d \ln p} = P_p(\ln p) \quad (3.8)$$

which implies

$$\alpha P_f(\ln f) = P_p(\ln p) \quad (3.9)$$

Dividing by α and substituting for $\ln p$ yields

$$P_f(\ln f) = \frac{1}{\alpha} P_p\left(\frac{\ln f - C}{\alpha}\right) \quad (3.10)$$

Then, substituting in that $P_p(\ln p)$ is a normal distribution with parameters μ_p and σ_p ,

$$P_f(\ln f) = \frac{1}{\alpha \sqrt{2\pi} \sigma_p^2} \exp\left(\frac{\left(\frac{1}{\alpha} \ln f - \frac{1}{\alpha} C - \mu_p\right)^2}{2\sigma_p^2}\right) \quad (3.11)$$

which can be rearranged as

$$P_f(\ln f) = \frac{1}{\sqrt{2\pi} (\alpha \sigma_p)^2} \exp\left(\frac{(\ln f - C - \alpha \mu_p)^2}{2\alpha^2 \sigma_p^2}\right) \quad (3.12)$$

which is a normal distribution with parameters $(\alpha \mu_p + C)$ and $(\alpha \sigma_p)$. Thus the probability density function of f is a lognormal distribution, and its parameters μ_f and σ_f are related to those of p by

$$\mu_f = \alpha \mu_p + C, \quad \sigma_f = \alpha \sigma_p. \quad (3.13)$$

3.1.19 Lower cutoff for production measurements

When analyzing the complete dataset of p measurements, we first must discard the 1,173 measurements where $p \leq 0$. Here we are only interested in positive measurements; note that much of the ocean (in principle, all of the ocean below the euphotic depth) has $p \approx 0$. Then, for small values of p , even very small errors or truncation of significant digits can strongly affect the log-transformed p values; to illustrate, a $+0.1$ error on a measurement of 10.1 causes a difference in log-space of < 0.01 , but on a measurement of 0.1 causes a difference in log-space of > 0.69 . Therefore we must impose a lower threshold, below which we neglect all samples. Samples below the threshold may not conform to the lognormal for various reasons, including error introduced by rounding, measurement errors, measurement sensitivity, because the null model is a poor one at low p values, or because the lognormal overlaps with a distribution of measurements for samples where $p = 0$ (which because of measurement error will not be a delta-function at $p = 0$).

The choice of threshold is not well-constrained; we use a threshold of 10% of the empirical p distribution's peak height (in log-space), which because the peak occurs at $\sim 5 \mu\text{g C L}^{-1} \text{ d}^{-1}$ means the threshold is set at $0.5 \mu\text{g C L}^{-1} \text{ d}^{-1}$. That this threshold separates measurements which neatly follow a lognormal distribution from those which do not can be seen easily in Figure S4a. By computing relative residual error of a lognormal fit as a function of p value, we confirm this cutoff value to be sensible; see Figure 3-12. To compute relative residual error, we sort the samples with $p > 0$ into logarithmically spaced bins, and then fit a lognormal distribution to the data above the threshold (see Section 3.1.9). For each bin i , the lognormal predicts ℓ_i samples to be in that bin, while there are actually m_i samples in that bin. The relative residual error is then given by $(m_i - \ell_i)/m_i$. Figure S4b shows that the lognormal underpredicts the number of samples by $\geq 40\%$ for all bins below the threshold, while the average over- or under-prediction for bins above the threshold is 5% and no bin exceeds 12%. This strongly indicates that the samples below the threshold do not follow the same distribution as those above the threshold, i.e. there is a distributional shift at the threshold (given that the samples above the threshold are well-described as lognormally distributed; see Section 3.1.9).

Thus, in the chapter we only discuss measurements of $p > 0.5 \mu\text{g C L}^{-1} \text{ d}^{-1}$. As noted above, neglecting measurements below the threshold is reasonable because they constitute

Figure 3-12: Identification of the lower threshold for p samples considered. a) Bars are PDF of measurements of $p > 0$; dashed black line (for both panels) is the threshold identified at $0.5 \mu\text{g C L}^{-1} \text{ d}^{-1}$, or $\sim 10\%$ of the peak in the log-transformed distribution. Green line is a lognormal probability distribution function above the threshold. b) Bars are the relative residual error of the lognormal distribution shown in panel (a).

<0.1% of all production in the data compilation, and our results are not sensitive to the value of the threshold as we repeated the analyses with the threshold varied by a factor of 2 (i.e. at 0.25 and 1 $\mu\text{g C L}^{-1} \text{ d}^{-1}$) and found that this had no impact on the results. Given a complete description of measurement error introduced by the ^{14}C method, one could incorporate samples below the threshold into the following analysis by convolving the lognormal distribution with the measurement error distribution for ^{14}C , as is done in image processing [Wiener, 1949]. For consistency, we treat the f data the same way, only analyzing data greater than 10% of peak height (3.6 $\text{mg C m}^{-2} \text{ d}^{-1}$), and also find that factor of two changes in the threshold value used made no difference to our results.

3.1.20 Kuiper’s statistic and other cumulative distribution function statistics

The standard way to compare a distribution to empirical data is to compare their cumulative distribution functions (CDF) [Flannery et al., 1992]. A variety of metrics to do so exist, though the presence of a lower threshold makes some unavailable, and simulations suggest that all have comparable statistical power for lognormal distributions of large sample size [Stephens, 1974]. We use Kuiper’s statistic:

$$v = \max \left(C(p) - c(p) \right) + \max \left(c(p) - C(p) \right)$$

where $C(p)$ is the empirical CDF of p and $c(p)$ is the lognormal fit, i.e. v is the sum of the maximum distance of c above and below C . This modification of the original CDF statistic (the Kolmogorov-Smirnov statistic; see below) is preferred in many applications because it balances ease of interpretation with sensitivity to the tails of the distribution [Flannery et al., 1992]. We note that, like the threshold, our results are not sensitive to this choice of metric; we repeated the subsequent analyses with two other metrics [Flannery et al., 1992], the original Kolmogorov-Smirnov statistic (D):

$$D = \max \left| C(p) - c(p) \right|$$

and the Anderson-Darling statistic (D^*):

$$D^* = \max \frac{|C(p) - c(p)|}{\sqrt{c(p)(1 - c(p))}}$$

and found that this had no impact on the results.

The standard techniques for calculating a p-value from these statistics are not appropriate here. This is due to many factors; the p and f data we test are not randomly distributed in time, depth, or latitude-longitude; measurement error exists and is not fully constrained. Also, sample sizes are large (i.e. n in the thousands or tens of thousands, and CDF metrics are known to be biased against large sample sizes [Field, 2009]). Instead, the Kuiper test must be used heuristically to estimate moments of lognormal fits to the p data and to quantify the goodness-of-fits of these lognormal distributions - higher values meaning worse fits. Visual comparison of the PDF of the p datasets with their lognormal fits suffices to demonstrate that in all cases p is well-described as lognormally-distributed, especially given the non-randomness of the p measurements. All of the above also applies to f . We find that for the global p and f distributions and the subregions' p distributions, $v \leq 0.0400$ in all cases, i.e. total deviations above and below between the empirical and hypothesized CDF's sum to $\leq 4\%$, corroborating the visual comparisons. As $v \sim n^{-1/2}$ [Stephens, 1974] (n.b. this also holds for D and D^*), the value of $v = 0.0400$ for the global distribution of f ($n = 1033$) is comparable to the value $v = 0.0107$ for that for the global distribution of p ($n = 38334$) when accounting for differences in sample size. Using the tests described in Stephens [Stephens, 1974], v , D , and D^* indicate that the time series' p distributions are inconsistent with a lognormal distribution, and that the time series' and biomes' f distributions are consistent with a lognormal distribution, though the smaller sample sizes for these latter datasets ($n \in [115, 447]$) afford less statistical power to reject a lognormal distribution.

Chapter 4

Depth and particulate organic carbon flux

The work in this chapter is based in part upon the following publication:

Cael, B. B., and K. Bisson. 2018. Particle flux parameterizations: quantitative and mechanistic similarities and differences *Frontiers in Marine Science*. [Cael and Bisson, 2018]

4.1 Particle flux parameterizations: quantitative and mechanistic similarities and differences

4.1.1 Overview

The depth-attenuation of sinking particulate organic carbon (POC) is of particular importance for the ocean’s role in the global carbon cycle. Numerous idealized flux-versus-depth relationships are available to parameterize this process in Earth System Models. Here we show that these relationships are statistically indistinguishable from available POC flux profile data. Despite their quantitative similarity, we also show these relationships have very different implications for the flux leaving the upper ocean, as well as for the mechanisms governing POC flux. We discuss how this tension might be addressed both observationally and in modeling studies.

4.1.2 Introduction

The sinking flux of particulate organic carbon (POC) and other elements into and through the ocean’s interior is a key flow in global elemental cycles [Williams and Follows, 2011]. As the complex mechanisms determining this material’s vertical redistribution remain poorly understood, Earth System Models parameterize particle flux as an attenuating function of depth [Gloege et al., 2016]. The idealized models used for this parameterization thereby influence the characterization of global biogeochemistry. Several such models have been proposed, and have generally been argued for based on their ability to fit measurement profiles [Martin et al., 1987b, Armstrong et al., 2001]. Particle fluxes are however notoriously variable and challenging to measure [Buesseler, 1991], while different functions’ goodness-of-fit are often strikingly similar [Figure 4-1]. Indeed, Gloege et al. [Gloege et al., 2016] showed the power and ballast models capture observations equally, as does an exponential in the upper 1km. Such similarity is not uncommon when modeling variable environmental data, but is of particular significance here for two reasons:

- All these models have plausible and distinct (if oversimplistic) mechanistic interpretations – e.g. that some material is labile and some is refractory, that remineralization systematically declines with time/depth, or that remineralization is a second-order kinetic process. One may also argue for other mechanistic models on similar grounds (see below). Thus, these quantitatively similar models have differing implications for how particle flux works¹.
- The ‘export’ of carbon through an upper reference depth horizon – usually the euphotic² or mixed layer depth (ELD or MLD) – is of substantial interest [Buesseler and Boyd, 2009, Palevsky and Doney, 2018]. However, most particle flux measurements have been made below the relevant horizon(s), necessitating depth-correction when comparing models with data and/or making global export estimates. These models diverge at shallow depths, meaning they also have widely differing implications for the magnitude of export through these depth horizon(s).

Our purpose herein is to rigorously demonstrate the tension between these models’ i) different mechanistic interpretations and ii) their often sizable quantitative differences at

¹To leading order and on average over large spatiotemporal scales.

²Defined in terms of 1% of sea-surface photosynthetically available radiation.

shallow depth horizons, versus iii) their quantitative similarity over the depths at which most particle flux measurements have been made. We derive several models mechanistically, compile a large database of measurement profiles, and perform a suite of statistical analyses. It is generally understood that these models behave similarly relative to measurement variability, and that extrapolation exaggerates differences. However, it is our opinion that the extent of these issues when modeling particle flux – the variety of plausible mechanistic models, the substantial differences that can result from even small upwards extrapolations, and the virtual impossibility of parsing between models by fitting them to measurements – is underappreciated. We discuss (i-iii) in turn, and how this tension might be addressed.

4.1.3 Several mechanistic particle flux attenuation models

We describe seven models [Figure 4-1] relating POC³ flux to depth, and provide mechanistic derivations for each in the Section 4.1.7. We note that these models are neither exhaustive nor uniquely derivable, and numerous neglected processes, from particle (dis)aggregation to zooplankton behavior to ambient environmental conditions, also play important roles [Mooy et al., 2002, Turner, 2015, Burd and Jackson, 2009].

The basic model is an exponential [Banse, 1990]:

$$f_e(z) := Ce^{-z/\ell} \quad (4.1)$$

which considers POC flux f [$\text{g m}^{-2} \text{d}^{-1}$] as attenuating with a characteristic lengthscale $\ell := w/k$ [m] determined by the settling velocity w [m/d] and remineralization rate k [1/d]. Here and below, C [$\text{g m}^{-2} \text{d}^{-1}$] is the scale parameter for POC flux, e.g. $C = f_e(0\text{m})$. The most common model is a power-law or ‘Martin curve’ [Martin et al., 1987b]:

$$f_p(z) := Cz^{-b} \quad (4.2)$$

which results from a systematic depth/time decrease of k (or increase of w – see Section 4.1.7), e.g. due to consumption of progressively more refractory material. The other major model is a ‘ballast’ model [Armstrong et al., 2001]:

$$f_b(z) := Ce^{-z/\ell} + c \quad (4.3)$$

³These equivalently apply to other elements or total matter.

which augments Eq. (4.1) with a refractory flux c . The original depth-attenuation model is a rational [Suess, 1980]:

$$f_r(z) := \frac{C}{z + a} \quad (4.4)$$

which considers POC remineralization to function as a second-order kinetic process. A double exponential [Lutz et al., 2002]:

$$f_d(z) := C_1 e^{-z/\ell_1} + C_2 e^{-z/\ell_2} \quad (4.5)$$

is used to capture both rapidly- and slowly-attenuating fluxes. A stretched exponential:

$$f_s(z) := C e^{-z^{(1-s)}/\ell} \quad (4.6)$$

where $s \in [0, 1)$, results similarly to Eq. (4.2) from a systematic depth/time decrease of k . Lastly, what we term a ExpInt model

$$f_g(z) := C E_1(z/\ell_{max}) \quad (4.7)$$

results from labile material within particles being protected from bacteria by ballast minerals [Rothman and Forney, 2007], where $E_1(\cdot, \cdot)$ is the exponential integral and the lengthscale ℓ_{max} depends partially on particles' microbial density.

Thus we have a suite of different models, each rooted in the balance between gravitational settling and bacterial remineralization, but assuming different basic characteristics for this balance. Given the enormous complexity and variability of the processes involved, all of the above are equally plausible when treated as idealized mechanistic models for particle flux attenuation.

4.1.4 Large differences in export estimates

As e.g. $f_p(z \rightarrow 0) = \infty$ while $f_e(z \rightarrow 0) = C$, the differences between these models becomes arbitrarily large for increasingly shallower depths. While these models are of course not applicable all the way to $z \rightarrow 0$, they are generally used to model f below a chosen depth horizon that is often appreciably shallower than most measurements (e.g. 50m versus 150m). Thus the different models may imply very different exports through a depth horizon.

Figure 4-1: A: Seven models from the text fit to mesopelagic data from Figure 5 of Martin et al. [Martin et al., 1987b] (their ‘Open Ocean Composite’) and extended upwards to 50m depth. Goodness-of-fits are very similar, but the power-law model overestimates particulate organic carbon (POC) flux at 50m by 132% relative to the exponential model. Inset shows attenuation ($\frac{1}{f} \frac{df}{dz}$) differences are even more pronounced. Attenuation for the exponential model coincides with most others around ~ 400 m. B: Maximum tolerable root-mean-square error (RMSE) to distinguish models based on data from (A) with 90% confidence, estimated using a bootstrap method (see text for details). Row corresponds to ‘true’ model (from which data are assumed to be generated) and column corresponds to ‘false’ model (to be rejected). RMSEs are small relative to those in (A), indicating differences in goodness-of-fits between the models are not statistically significant. C: Ratio of power-law-model- and exponential-model-estimated POC flux at euphotic layer and mixed layer depths from the profiles in the data compilation as a function of extrapolation distance. Grey shading shows nominal 25% measurement uncertainty. Note the y -axis is logarithmic. The power-law model systematically and substantially overestimates relative to the exponential model (as well as the other models – see Section 4.1.10). For 29 profiles, $f_p(\text{MLD})/f_e(\text{MLD}) > 10$ (not shown).)

This is evident in Figure 4-1; the functions are similar through the measurements' depth range (100–1100m), but $f_p(50\text{m})$ exceeds $f_e(50\text{m})$ by 132%, with the other models yielding intermediate values. While the existence of such differences is unsurprising, these have not yet been thoroughly quantified; upon doing so, we found their magnitudes to be surprisingly large.

To investigate how model choice affected export estimates, we compiled a database of f measurement profiles. The compilation (see Section 4.1.8 for description – data are available online [Cael and Bisson, 2018]) a) contains 722 profiles ($n \geq 3$ depth points per profile) measured with sediment traps or the ^{234}Th technique across the global ocean. We used profiles rather than individual measurements to derive the models' parameter values from the data being extrapolated. We then fit the above models using nonlinear least-squares regression and minimizing relative⁴ error, and extrapolated these fits to climatological estimates of ELD and MLD (see Section 4.1.9).

We found f_p systematically overestimated and f_e systematically underestimated export relative to other models, and the range $f_p:f_e$ was often substantially larger than the nominal variability between replicate measurements of 20-30% [Buesseler and Others, 2000, Buesseler and Others, 2007, Stanely et al., 2004]. For ELD, $f_p:f_e > 1.25/1.5/2$ for 67/45/25% of profiles, respectively. For MLD, which was typically shallower than ELD, $f_p:f_e > 1.25/1.5/2$ for 75/62/49% of profiles, respectively. Figure 4-1 shows $f_p:f_e$ for all profiles as a function of extrapolation distance (i.e. the distance between ELD or MLD and the shallowest measurement); $f_p:f_e$ increases with extrapolation distance as expected, but large differences are not uncommon even for small extrapolation distances. These analyses clearly indicate that the above models are often quantitatively disparate at shallow depths. As these estimates are all based on the same $f(z)$ data from below the ELD and MLD, they imply even larger differences in attenuation ($\frac{1}{f} \frac{df}{dz}$ [m^{-1}]) just below these depth horizons [Figure 4-1].

Fortunately, having consistent end-members (f_e, f_p) simplifies the situation. This suggests that when estimating export by extrapolating a measurement, one can use (f_e, f_p) as lower/upper bounds. We determined unbiased parameter values of $\ell \approx 500\text{m}$ and $b \approx 0.7$ (similar to Primeau [Primeau, 2006]) for doing so by comparing profile fits to fixed-parameter extrapolations of only the shallowest measurement in each profile (see Section 4.1.10). Using these parameters, f_p still systematically overestimates with respect

⁴Minimizing absolute error yielded similar results.

to f_e ; $f_p(\text{ELD}) : f_e(\text{ELD}) \leq 3.38$ (median = 1.50), and $f_p(\text{MLD}) : f_e(\text{MLD}) \leq 3.84$ (median = 2.31). These parameter estimates are however influenced by spatiotemporal biases in sampling.

4.1.5 Quantitative indistinguishability

In contrast with the above, these models are known to behave similarly over the depths at which most particle flux measurements have been made – in the mesopelagic and bathypelagic⁵ – especially relative to these measurements’ substantial variability and uncertainty [Figure 4-1]. Are the differences in these models’ ability to fit measurements sufficient to parse between them?

To address this question we applied a statistical routine to the profile database. The routine estimates for a set of measurements what level of variability is tolerable to reject one model versus another at a chosen confidence level, or inversely the percent confidence with which those measurements can reject one model versus another. Detailed description of the routine and examples of its application are provided in Section 4.1.11. In brief, data are simulated from a ‘true’ model, which is fit by a ‘false’ model to be rejected. The root-mean-square error (RMSE) of the simulated data relative to each model provides a metric of goodness-of-fit. Bootstrapping yields a percent confidence in the true model versus the false one. Repeating this process yields a true-model-versus-false-model table either of tolerable measurement variabilities [Figure 4-1] or confidence levels of false model rejection (see Section 4.1.11).

Applying this routine to the data in Figure 4-1 indicates these data are too variable for the differences in fit to be significant, and that the RMSE would have to be $< \frac{1}{2}$ its true value even to reject an exponential in favor of other models. The stretched and rational models are so similar that they cannot be distinguished even at 1% measurement variability from data with this sample size and depth spacing.

Assuming a conservative characteristic measurement variability of 20%, no one model can be rejected with 90% confidence in favor of any other for more than a handful of profiles in the dataset. Increasing depth range or sample size (either for composites of profiles [Figure 4-1], or for high-resolution individual profiles) tends to increase the maximum tolerable RMSE. However, the statistical non-significance of models’ differences in fit appears extremely robust

⁵Excepting the exponential, which performs worse in the bathypelagic [Gloege et al., 2016].

for realistic measurement variability [Section 4.1.11]. Thus existing measurement profiles appear insufficient to distinguish between models. The exception to this is the exponential when bathypelagic data are included, as this is the only model for which attenuation does not change with depth. If for instance the bathypelagic data from Figure 5 of Martin et al. [Martin et al., 1987b] are included in the above analysis (11 data points from 1400–2000m), an exponential could be rejected with 90% confidence relative to a power-law or rational and vice versa if RMSE were ≤ 0.24 (an appreciable increase from the values seen in the table in Figure 4-1, though still below the data’s true value), but the power-law and rational can still only be rejected with respect to each other when $\text{RMSE} \leq 0.04$.

The root issue that complicates parsing between these models from $f(z)$ data⁶ is that the models have multiple free parameters and therefore can take a wide range of shapes. Treating these models as mechanistic and using additional measurements to constrain their parameters may be useful in this regard. For instance, by measuring average $\bar{k}$ and $\bar{w}$ independently of $f(z)$ data, one could fit the far-less-flexible exponential model $f_e(z) = C \exp[-z/(\bar{w}/\bar{k})]$. Being mechanistic, this approach can also provide quantitative insight about the dominant processes controlling particle flux. Of course, there are also other means by which to parse between different representations of $f(z)$, such as suites of numerical simulations [Gloege et al., 2016], chemical composition [Armstrong et al., 2001], or mass/energy budgets [Burd and Others, 2010]; comprehensive understanding is likely only possible from combining multiple approaches to generate multiple constraints.

4.1.6 Conclusion

Several particle flux models capture observations equivalently, but carry different implications mechanistically and for magnitudes of export. Here we have attempted to illustrate this tension. When depth-correction is necessary, functional extrapolations are better used as lower (exponential) and upper (power-law) bounds. Vertically extensive high-resolution data⁷, paired with measurements of related quantities, may potentially elucidate how best to represent these processes mechanistically in Earth System Models.

⁶Besides the data’s sparsity and variability, neither of which can be easily ameliorated.

⁷i.e. tens of measurements spanning the full water column and including ELD and MLD.

4.1.7 Derivations

The full three-dimensional time-varying conservation equation for POC [p , g m⁻³] can be written:

$$\partial_t p + \vec{u} \cdot \nabla p - \kappa \nabla^2 p = \mathcal{S} \quad (4.8)$$

where $\vec{u} = (u_x, u_y, u_z)$ is the incompressible velocity field, κ is the molecular diffusivity, and $\mathcal{S}$ represents the net effect of sources and sinks. We are interested in large horizontal and temporal scales, so we consider the system as integrated over large enough spatial and temporal scales for time-variation and lateral advection and diffusion to be negligible, or equivalently we assume that the system is in steady state and one-dimensional. The equation then becomes

$$u_z \partial_z p - \kappa \partial_{zz} p = \mathcal{S} \quad (4.9)$$

We then assume that the two terms on the left hand side are small, because both molecular diffusivity and ocean vertical velocities on large horizontal and temporal scales are small relative to POC fluxes. We are left with the expression that sources must balance sinks (at least to leading order), i.e. $\mathcal{S} \approx 0$; note we have now assumed away all of the effects of tracer advection and diffusion. We then assume that the divergence⁸ of POC flux dominates the source component of $\mathcal{S}$ and is balanced by the sink component of $\mathcal{S}$, which is commonly referred to as remineralization but includes all processes that remove material from the POC pool (e.g. fragmentation).

Splitting $\mathcal{S}$ into these parts, then, we have the basic equation from which these models arise, where w is apparent settling velocity, and k is the apparent first-order remineralization rate:

$$-\partial_z(wp) - kp = 0 \quad (4.10)$$

If we then assume $k \approx \text{const.}$ and $w \approx \text{const.}$, the solution to this first-order linear differential equation for p is

$$p(z) = Ae^{-\frac{k}{w}z} \quad (4.11)$$

where A is a constant dependent on initial conditions, i.e. on p at a shallow reference depth.

⁸i.e. the difference between the POC flux entering a infinitesimal depth-layer $z + \Delta z$ and the POC flux leaving that depth-layer, divided by the thickness of that layer.

Multiplying both sides by w to get POC flux f yields

$$f(z) = wp(z) = wAe^{-\frac{k}{w}z} \quad (4.12)$$

so if we define $\ell := w/k$ and $C := wA$, we find

$$f_e(z) = Ce^{-z/\ell} \quad (4.13)$$

i.e. the exponential model. The ballast⁹ and double exponential models can be trivially derived as extensions of this. If we consider two distinct pools of POC, i.e. $p = p_1 + p_2$, each with their own associated (w_1, w_2, k_1, k_2) , then each has an exponential solution for f_1 and f_2 as above, and the total flux is $f = f_1 + f_2$ which is the sum of these exponentials, i.e. the double exponential model. If $k_2 = 0$, e.g. because p_2 is protected from remineralization by an associated mineral matrix, then $f_2 = c$, where $c = \text{const.}$, i.e. the ballast model.

If instead we assume that k is inversely related to time (evidence indicates this relationship to hold for the decomposition of organic carbon in sediments across many orders of magnitude [Middelburg, 1989]), i.e. $k = \hat{k}/t$, which assuming $w \approx \text{const.}$ implies¹⁰ $k = \hat{k}/t = \hat{k}/(z/w) = \hat{k}w/z$, Eq. (4.10) becomes

$$-\partial_z(wp) - \left(\frac{\hat{k}w}{z}\right)p = 0 \quad (4.14)$$

The solution to this first-order linear differential equation for p is

$$p(z) = Az^{-\hat{k}} \quad (4.15)$$

where again A is a constant dependent on initial conditions; multiplying both sides by w and defining $b := \hat{k}$ and $C := wA$ yields

$$f_p(z) = Cz^{-b} \quad (4.16)$$

i.e. the power-law model. Note this means that the interpretation of b is that it is the

⁹n.b. these models are discussed in a different order here than above. The order above we found to be most natural in the context of the literature on these models; the order here we found to be most natural in terms of deriving these models sequentially.

¹⁰n.b. $\hat{k}$ is dimensionless. Throughout, hats are given to parameters whose units are specified similarly, e.g. the units of $\hat{w}$ are 1/d.

remineralization rate of p after one day¹¹ of decomposition and sinking. We can similarly derive f_p from assuming $w = \hat{w}z$; we then have instead of Eq. (4.14)

$$-\partial_z(\hat{w}zp) - kp = 0 \quad (4.17)$$

which similarly leads to a power-law solution for f_p but with a different $b := (k + \hat{w})/\hat{w}$. Though there is evidence for increasing settling velocities with depth [Berelson, 2001], for a sinking particle z is the time-integral of w , making this definition of w in terms of z somewhat ad hoc. Furthermore, $b > 1$ as defined above, unlike in most fits to $f(z)$ data.

If we assume instead of an inverse relationship between k and t a weaker scaling with time such that $k = \hat{k}t^{-s}$, where $s < 1$, we have $k = \hat{k}w^s z^{-s}$ for a depth-dependence of k . This is more consistent with the scaling law found by Middelburg [Middelburg, 1989], who found $k \sim t^{-.95}$. Eq. (4.14) then becomes instead

$$-\partial_z(wp) - (\hat{k}w^s z^{-s})p = 0 \quad (4.18)$$

which is a first-order linear differential equation for p , the solution to which is

$$p(z) = Ae^{(\frac{\hat{k}w^{s-1}}{s-1})z^{1-s}} \quad (4.19)$$

where again A is a constant dependent on initial conditions, and multiplying both sides by w yields

$$f(z) = wAe^{(\frac{\hat{k}w^{s-1}}{s-1})z^{1-s}} \quad (4.20)$$

so if we define $\ell := (s - 1)/(\hat{k}w^{s-1})$ and $C := wA$, we find

$$f_s(z) = Ce^{z^{(1-s)}/\ell} \quad (4.21)$$

i.e. the stretched exponential model. Interestingly, though s can take a range of values $s \in [0, 1)$, typically in the fits to profiles described below we find s close to the 0.95 value from Middelburg [Middelburg, 1989]; for instance, in Figure 4-1, $s = 0.90$.

Rather than assuming a decreasing k with time, we can assume that remineralization of p occurs as a second-order kinetic process, i.e. $\frac{dp}{dt} = -\hat{k}p^2$, rather than a first-order process

¹¹ $t = 0$ can be thought of as the time at which particles sink through the reference depth horizon.

as above. We then have instead of Eq. (4.10)

$$-\partial_z(wp) - \hat{k}p^2 = 0 \quad (4.22)$$

which is a first-order nonlinear differential equation for p , the solution to which is

$$p(z) = \frac{w/\hat{k}}{z + Aw/\hat{k}} \quad (4.23)$$

where again A is a constant dependent on initial conditions, and multiplying both sides by w and defining $a := Aw/\hat{k}$ and $C := w^2/\hat{k}$ yields

$$f_r(z) = \frac{C}{z + a} \quad (4.24)$$

i.e. the rational model.

The last model above, what we refer to as the ExpInt model, can be derived from the application of the model described by Rothman & Forney [Rothman and Forney, 2007] to particles with a constant settling velocity w . The full derivation of their model is involved, so we begin with their main finding for $p(t)$ (in our notation) and refer the reader to their paper for a derivation from first principles. They assume organic carbon is randomly distributed on the mineral surfaces of a porous medium populated by randomly distributed heterotrophic bacteria. A sinking particle certainly fits this description: a porous medium with organic carbon, minerals, and bacteria distributed throughout. They then assume that the decay of a particular parcel of organic carbon is limited by its accessibility to hydrolytic enzymes. Although all of the organic carbon in a given particle is assumed to be equally labile, bacterial remineralization occurs at an increasingly slower rate as bacteria must consume material that becomes increasingly more difficult to access within its heterogeneous microenvironment. They derive as a solution (their equation 2) for the POC concentration on a particle with time to be

$$p(t) \propto \int_{k_{min}}^{\infty} \frac{1}{k} e^{-kt} dk \quad (4.25)$$

which yields a flux if we assume a constant w by multiplying by w and replacing $t = z/w$:

$$f(z) \propto \int_{k_{min}}^{\infty} \frac{1}{k} e^{-\frac{k}{w}z} dk = E_1(z/\ell_{max}) \quad (4.26)$$

Figure 4-2: A map of data locations with histograms of data properties. POC flux is reported in $\text{mg C m}^{-2} \text{ d}^{-1}$.

where $\ell_{max} := w/k_{min}$, and $E_1(\cdot)$ is the exponential integral

$$E(y) := \int_y^\infty v^{-1} e^{-v} dv \quad (4.27)$$

k_{min} itself is associated with the typical distance between microbes, which is a function of microbial density assuming microbes are randomly distribution within particles. If we subsume all of the prefactors (including w) into a constant C , we are left with

$$f(z) = C E_1(z/\ell_{max}) \quad (4.28)$$

i.e. the ExpInt model.

4.1.8 Flux profiles

We compiled 722 POC flux profiles measured with sediment traps or the ^{234}Th technique, including studies from 1978 to 2014 at depths up to 4680m. The spatial bias in the profile locations reflects that of available data; many of the profiles come from three major time series sites: Bermuda Atlantic Time Series (BATS), Hawaii Ocean Time Series (HOT), and Carbon Retention In A Colored Ocean (CARIACO). Profiles have between 3 and 14 observations, with a mean of 3.9. The duration of deployment for the trap profiles ranges from 1 to 425 days; no corrections were made to account for potential POC remineralization within sample cups [Antia, 2005]. All fluxes are converted to $\text{mg C m}^{-2} \text{ d}^{-1}$ using a molar mass of 12.01 for carbon. Flux observations range from 0 to $1238 \text{ mg C m}^{-2} \text{ d}^{-1}$, with a mean of 40 mg C m^{-2} . Data are available online [Cael and Bisson, 2018].

4.1.9 Euphotic and mixed layer depth climatologies

From each flux-depth model and measurement profile we estimated flux at the ELD and MLD by normalizing to a climatological depth. To generate an ELD climatology, we first downloaded level-3 Chl values from the NASA ocean color website <<https://oceancolor.gsfc.nasa.gov/cgi/l3>> and constructed monthly climatologies from 1997-2008 observations, the interval over which there were no major sampling issues in the SeaWiFS mission and during which most of the profiles were measured. From this we calculated ELD, defined as the depth of 1% photosynthetically available radiation, according to relationships between Chl and ELD described in Morel et al. [Morel et al., 2007] at 9 km spatial resolution. The minimum ELD is 14m, the mean ELD is 62m, and the maximum ELD is 97m. For MLD we used the MLD-DT02 climatology [de Boyer Montégut et al., 2004] which defines MLD as the depth at which temperature varies greater than 0.2°C from its value at 10m and has a spatial resolution of 2° . The minimum MLD is 14m, the mean MLD is 43m, and the maximum MLD is 156m.

4.1.10 Normalizations

We normalize the POC flux profiles to each depth climatology using the seven flux-depth models. For each profile we fit each of the seven functions using non-linear least squares regression and minimizing relative error, resulting in optimum parameter values specific to

each profile and function. We then estimate export at both depth horizons for each profile at that profile’s latitude, longitude, and month, and using these optimum parameter values. When fitting functions with $n > 2$ free parameters, e.g. the ballast model, we restricted fits to profiles with $> n$ measurements, to retain at least one degree of freedom in the fit. To choose unbiased parameter values for normalization by a power law or exponential, we compare the fits of profiles described above to normalizations of only the shallowest observation in each profile with a fixed b or ℓ , respectively. We systematically vary both b and ℓ (to two significant digits) and choose the parameter value as that results in an equal amount of over- and underestimations of export relative to the estimate using profile-specific parameters. In all cases we only normalize measurements that are below the depth of interest. Our results were not sensitive to restricting to subsets of the total set of profiles (e.g. only profiles with a measurement $> 1\text{km}$ deep, or with $n \geq 4$). As the results for ELD and MLD were qualitatively similar, we do not expect they are sensitive to the choice of depth horizon. We do not expect the general findings to be sensitive to the choice of depth horizon as the results for ELD and MLD are qualitatively similar, though certainly the magnitude of the disparity between models is magnified at shallower depths.

Power-law- and exponential-normalized export estimates at both the euphotic and mixed layer depth tended to bound estimates from other models, e.g. $f_e < f_r < f_p$. For any of the five models other than f_p and f_e , f_p yielded a larger export estimate for and f_e yielded a smaller export estimate for 76-100% of fits, for both ELD and MLD. Table 4.1 shows the fraction of profiles for which f_e significantly overestimates or f_p significantly underestimates export relative to the five other models for a nominal measurement uncertainty of 25%. Occasionally f_p significantly underestimates relative to f_b or f_g (11% and 14% of the time, respectively), but in general significant over/underestimations are infrequent. We view this as justification for the use of (f_e, f_p) as lower and upper bounds of an estimated export range when depth-normalizing particle flux measurements.

For unbiased parameter estimates, we find $b \approx 0.7$ and $\ell \approx 500\text{m}$ (exact values: $b = 0.69$ for $f_p(\text{ELD})$, $b = 0.68$ for $f_p(\text{MLD})$, $\ell = 530\text{m}$ for $f_e(\text{ELD})$, and $\ell = 500\text{m}$ for $f_e(\text{MLD})$; the differences between these exact values and the approximate values reported above are negligible) – see Figure 4-3. Across all profiles, $f_{0.7}$ (power-law-normalized export using the shallowest measurement in a profile and $b = 0.7$) always overestimates relative to $f_{500\text{m}}$ (exponential-normalized export using the shallowest measurement in a profile and

model	# fits	% $< \frac{4}{5}f_e$	% $> \frac{5}{4}f_p$
ballast	277	1	11
rational	722	3	0
double	187	0	0
stretched	277	1	0
ExpInt	722	0	14

Table 4.1: Fraction of profile fits for which f_e significantly overestimates or f_p significantly underestimates export relative to other models, using a nominal measurement uncertainty of 25%. For instance, f_e is less than or within 25% of f_r for 97% of profiles, so only significantly overestimates f_r 3% of the time. The value reported is the larger value between normalizations to ELD and to MLD, e.g. $f_e(\text{MLD})$ significantly overestimates $f_r(\text{MLD})$ 2% of the time while $f_e(\text{ELD})$ significantly overestimates $f_r(\text{ELD})$ 3% of the time, so the ELD value is reported. We use $4/5$ rather than $3/4$ in the third column because $4/5$ is the same *relative* distance from 1 as $5/4$; if $3/4$ was used instead the values in the table would be the same.

$\ell = 500\text{m}$).

Minimizing absolute error rather than relative error in the fitting procedure yielded similar quantitative results and did not affect our main conclusions. For instance, $f_p(\text{MLD}):f_e(\text{MLD}) > 1.25/1.5/2$ for 71/57/39% of profiles, respectively; $f_p(\text{ELD}):f_e(\text{ELD}) > 1.25/1.5/2$ for 62/35/14% of profiles, respectively. These values are similar but slightly smaller than those for minimizing relative error (75/62/49% for MLD and 67/45/25% for ELD, from above). Because larger variability is seen in shallower measurements [Buesseler, 1991] and uncertainty and variability are typically reported as percentages for these measurements, minimizing relative error is more justifiable in the absence of more comprehensive error information for these measurements.

4.1.11 Statistical routine

The statistical routine addresses whether the difference in two models' quality of fit to data has the capacity to be statistically significant at those data's sample number and depth spacing, relative to the uncertainty in the models' quality of fit. For a set of $f(z)$ measurements, data are simulated at the measurements' sample depths from a 'true' model f_T thought to be generating the data; the true model is also fit by a 'false' model f_F to be rejected. The parameters for the true model are taken from fitting the $f(z)$ data as above, and measurement 'noise' with characteristic magnitude ν is superimposed on the simulated data. The root-mean-square error (RMSE) of the simulated data relative to each

Figure 4-3: Cumulative distribution function of fixed-parameter vs. fit-parameter normalizations. Horizontal lines indicate the fraction of profiles in each case where the fixed-parameter normalization underestimates relative to the fit-parameter normalization. That these are all close to 0.5 indicates that the approximations of 500 m for ℓ and 0.7 for b introduce negligible bias.

model provides a metric of goodness-of-fit. Repeating this process many times produces a distribution of RMSE for each model relative to the simulated data, i.e. $p_T(\text{RMSE})$ and $p_F(\text{RMSE})$.

The median RMSE for the true model will be ν by construction, and the false model's RMSE will on average be larger by construction, say $\nu + \varepsilon$. For any one simulated profile, however, the RMSE of the false model may be smaller, because the random noise can result in the false model having a better fit. This difference in these models' median RMSE, ε , may or may not be significantly detectable relative to the variation in RMSE for individual simulated profiles. The probability distributions $p_T(\text{RMSE})$ and $p_F(\text{RMSE})$ quantify uncertainty in the actual goodness-of-fit of f_T and f_F to the simulated data; if these distributions overlap, we cannot be 100% confident that f_T is a better fit, i.e. that $\varepsilon \geq 0$. The distributions yield a bootstrap percent confidence in the 'true' model versus the 'false' one, depending on their overlap. This process is systematically repeated with each model as 'true' and 'false' to yield a 'true'-model-versus-'false'-model table of tolerable measurement variabilities for a given confidence level (as in Figure 4-1), or confidence levels of 'false' model rejection for a given measurement variability (by systematically increasing ν until the estimated percent confidence goes below the chosen confidence level).

The routine can be described stepwise as follows:

1. Choose a set of $f(z)$ data, a bootstrap iteration number N , and a measurement variability level ν . (The $f(z)$ data can be either a single profile or a composite of profiles. For our analyses we found $N = 10^5$ to be sufficient. We chose $\nu = 0.2$ in the following analyses, corresponding to the lower end of reported measurement variability for sediment trap and ^{234}Th measurements [Buesseler and Others, 2000, Buesseler and Others, 2007, Stanely et al., 2004]; ν can also be directly estimated from composite profile data, such as that from Figure 4-1, where $\nu \approx 0.37 - 0.4$.)

2. Choose a 'true' model f_T and a 'false' model f_F . Select parameters for f_T by fitting to the $f(z)$ data, then select parameters for f_F by fitting to f_T sampled at the depths and resolution of the $f(z)$ data. (For our analyses we fit by minimizing relative error as described in Section 4.1.10 above. We therefore also simulated measurement variability multiplicatively, i.e. superimposed multiplicative Gaussian measurement error with log-standard deviation ν , and calculated RMSE in terms of relative error. We also repeated these analyses minimizing absolute error, simulating measurement variability additively, and

calculating RMSE in terms of absolute error. Doing so yielded similar results. f_F is fit to f_T rather than to $f(z)$ because the objective is to determine how closely different models can fit each other relative to measurement variability, which is included in $f(z)$.)

3. Simulate N measurement profiles sampled from f_T by sampling f_T at the depths and resolution of the $f(z)$ data N times and each time superimposing random variability with characteristic magnitude ν . Compute the RMSE for f_T and f_F for each of the N simulated profiles to generate the probability distributions $p_T(\text{RMSE})$ and $p_F(\text{RMSE})$.

4. Determine the percent confidence with which f_F can be rejected in favor of f_T from maximum percentile range at which $p_T(\text{RMSE})$ and $p_F(\text{RMSE})$ are disjoint. (For instance if the 75th percentile of $p_T(\text{RMSE})$ equals the 25th percentile of $p_F(\text{RMSE})$, f_F can be rejected with 50% confidence).

5. Repeat steps 2-4 for each valid combination of true and false models. (Not including cases where the true model is the same as or a special case of the false model, e.g. exponential vs. exponential, ballast, double, or stretched).

6. If determining tolerable variability thresholds for a chosen confidence level, repeat steps 1-5 above, systematically increasing ν until the estimated percent confidence is lower than the chosen confidence level.

Figure 4-4 illustrates the routine visually. A true model (there, a power-law, whose parameters are taken from Figure 4-1) is fit by a false model (there, a ballast); data is then generated at sample depths (there $z = [125, 250, 500, 1000, 2000, 4000]\text{m}$ and $\nu = 0.20$). RMSEs of many simulated profiles (only one of which is shown in Figure 4-1) result in probability distributions $p_T(\text{RMSE})$ and $p_F(\text{RMSE})$, and the percent confidence of rejecting the false model is determined by the overlap of these probability distributions. Each model is used as both true and false, to generate a model-model-table. For this artificial profile, an exponential can be rejected versus most other models with reasonable confidence, but other model-model comparisons are generally well below typical significance thresholds.

Table 4.2 summarizes the result of applying the statistical routine to the profile database. With an assumed measurement variability of 20% (i.e. $\nu = 0.2$) and chosen confidence level of 90%, the table shows the number of profiles out of the total of 722 for which each model could be significantly rejected relative to each other. In general, very few profiles have the necessary vertical resolution and spacing. Significant rejection is largely restricted to profiles that have both high vertical resolution ($n \geq 6$) as well as a combination of shallow,

% Confidence in rejecting false model

T vs:	E	P	B	R	D	S	I
E	.	99	.	99	.	.	53
P	90	.	39	7	2	2	57
B	96	46	.	68	.	50	82
R	88	5	35	.	1	<1	4
D	85	17	54	11	.	11	44
S	88	2	38	2	5	.	50
I	31	28	7	2	1	1	.

Figure 4-4: a) Illustration of the statistical routine. b) Percent bootstrap confidences in rejecting the ‘false’ (column) model in favor of the ‘true’ (row) one, from applying the statistical routine illustrated (a) to an artificial profile with observations at $z = [125, 250, 500, 1000, 2000, 4000]$ m and 20% measurement variability. Initials are first letter of each model, except I refers to ExpInt model.

T vs:	E	P	B	R	D	S	I
E	·	6	·	7	·	·	0
P	6	·	1	0	2	0	0
B	13	7	·	9	·	4	7
R	7	0	1	·	2	0	3
D	11	5	1	8	·	3	3
S	2	1	1	2	3	·	0
I	0	0	2	4	4	0	·

Table 4.2: Number of profiles (out of a total of 722) for which the percent bootstrap confidence in rejecting the ‘false’ (column) model in favor of the ‘true’ (row) one is greater than 90%, assuming a measurement uncertainty of 20%. Initials are first letter of each model, except *i* refers to the ExpInt model.

mesopelagic, and bathypelagic observations.

Table 4.3 illustrates that increasing vertical resolution uniformly increases the percent confidences, but only to a limited extent. When the vertical resolution in Figure S3 is increased to $n = 20$ (logarithmically spaced measurements from 100-4000m), the percent confidences all increase, but many of the models can fit each other so closely that the percent confidence of rejection is still very small in many cases. Table 4.4 illustrates the importance of vertical extent of measurement profiles for comparing curves; when the same $n = 20$ measurements are logarithmically spaced over 100-1000m instead of 100-4000m, the percent confidence of rejection of curves is drastically reduced; not even the exponential can be significantly rejected in many cases. It is also worth noting that this level of vertical resolution is rarely, if ever, achieved in real observations.

In general the application of this routine to many artificial and actual measurement profiles broadly demonstrates that these models can fit each other so closely relative to the variability in observations that the small differences in their goodness-of-fit are insufficient to significantly reject one model in favor of another. The exception to this is the exponential model for profiles extending into the bathypelagic, where the exponential is known to perform poorly. We find these results to be robust across different combinations of model parameters, different types and magnitudes of simulated variability, and different measurement resolution and spacing.

T vs:	E	P	B	R	D	S	I
E	·	99	·	99	·	·	70
P	99	·	53	11	3	10	76
B	99	57	·	77	·	61	95
R	98	7	46	·	7	1	66
D	97	18	70	12	·	11	64
S	98	3	52	3	9	·	68
I	70	98	37	99	2	2	·

Table 4.3: Same as Figure S3b except $n = 20$.

T vs:	E	P	B	R	D	S	I
E	·	20	·	4	·	·	8
P	54	·	7	2	1	1	20
B	15	27	·	14	·	5	2
R	39	2	4	·	1	1	8
D	44	7	1	3	·	4	12
S	46	<1	6	1	1	·	14
I	12	12	1	2	<1	9	·

Table 4.4: Same as Table 4.3 except $\max(z) = 1\text{km}$.

4.2 Distributional model for particle flux attenuation

4.2.1 Overview

Sinking particulate matter comprises a major branch of the ocean’s ‘Biological Carbon Pump’ [Buesseler and Boyd, 2009]. While marine particle fluxes are complex and variable, idealized models of the depth-attenuation of flux are useful not only to parameterize particle flux in Earth System Models [Gloege et al., 2016], but also to diagnose spatial and temporal patterns in flux observations [Marsay et al., 2015] and to generate testable theories [Armstrong et al., 2001]. Here we propose a mechanistic, distributional model for the depth-attenuation of marine particle flux, that takes into account measurements of both the flux profile and the settling velocity distribution (SVD) of particulate matter. SVD measurements are best captured by a reciprocal distribution. This distribution then implies the exponential integral (ExpInt) model $f(z) = f_o E_1(z/\ell_o)$ if we assume that the bulk flux profile is the superposition individual particles’ fluxes and controlled by the initial variation in settling velocity. The ExpInt model also fits measured flux profiles equivalently to or better than standard models. Though necessarily oversimplistic, it is therefore unique in capturing observations of both flux vs. depth and of the SVD, and thus constitutes a particle flux

parameterization that is more consistent with and constrained by empirical knowledge. Furthermore, the emergence of the reciprocal distribution underlying the ExpInt model can be understood as resulting from the large number of processes that produce sinking particulate material. The model also generates multiple hypotheses that are testable *in situ*.

4.2.2 Particle flux complexity

In the ocean, particle flux is nothing if not complex and heterogeneous. Total fluxes comprise myriad sinking particles which are themselves the complex result of myriad biological, chemical, and physical factors [de la Rocha and Passow, 2007]. These particles' physical properties vary widely, from their size and shape to their density and porosity [Kajihara, 1971, Alldredge and Gotschalk, 1988]. Individual particles themselves are comprised of a complex suite of organic and inorganic materials, which may represent a continuum ranging from highly labile to utterly refractory [Hedges et al., 2001, Dittmar, 2008, Jiao and Others, 2010]. Beyond this, the remineralization of particulate material is plausibly related to properties of the ambient fluid like temperature or oxygen concentration [Mooy et al., 2002], may occur via consumption by different bacteria using different metabolic/chemical pathways even within a single particle [Bianchi et al., 2018], may occur at different rates due to the physical microstructure of the particle [Rothman and Forney, 2007], and may in turn alter the shape and other physical properties of the particle [Mayor et al., 2014]. Settling velocity is also related to the density of the ambient fluid, which of course changes with depth due to stratification [MacIntyre et al., 1995]. Individual particles may also interact with each other through myriad aggregation and disaggregation events [Burd and Jackson, 2009, Jackson and Burd, 1998, Alldredge et al., 1990], may be consumed or otherwise altered by free-swimming zooplankton [Steinberg et al., 2008, Kiørboe, 2000, Steinberg, 1995], etc.

From a modeling perspective, we may broadly classify all of these processes and properties into four 'Axes of Complexity' (AoC):

AoC1. Those that determine particles' 'initial' settling velocities at a reference depth, and hence the total settling velocity distribution (SVD) of particulate matter at that depth.

AoC2. Those that determine the remineralization rates of particles' material subsets at a reference depth, and hence the total remineralization rate distribution of particulate matter at that depth.

AoC3. Those that determine how a particle's settling velocity changes as it sinks (i.e.

changes with depth/time), or how a material subset of a particle’s remineralization rate changes as it sinks.

AoC4. Those by which particles interact with each other or with other non-ambient components of the water column such as zooplankton.

The first two of these determine the initial settling-remineralization balance; the third determines the evolution of the settling-remineralization balance for individual particles; the fourth determines the extent to which the total flux can be decomposed into the microscopic fluxes of individual particles. In total, then, we may consider particle flux as comprised of diverse particles sinking at variable and changing rates, themselves comprised of diverse material being remineralized at variable and changing rates, which may interact with each other and with other components of the system.

Although an impressive body of work documents the importance of all of these processes and properties on the depth-attenuation of particle fluxes, much remains unknown or only coarsely quantified. Little is known about the quantitative effect of zooplankton and particle-particle interactions (i.e. AoC4) [Kjørboe, 2000, Burd and Jackson, 2009]. There is evidence for increases [Berelson, 2001, Villa-Alfageme and Others, 2016] in *average* settling velocity with depth, but whether these are the result of changes to or initial variation in individual particles’ settling velocities is unclear (i.e. AoC3 or AoC1). Similarly, while a decrease in bulk apparent first-order remineralization rate k [d^{-1}] is a consistent feature of marine sediments [Middelburg, 1989], it is unclear how k varies on the much faster timescales over which particles sink through the water column, whether this variation would be due to changes to or initial variation in the remineralization of material subsets of particles (i.e. AoC3 or AoC2), or even in the case of sediments how much variation in k is due to differences in lability or accessibility of organic matter [Rothman and Forney, 2007, Jiao and Others, 2010]. While initial variation in k for material subsets of particles can in principle be estimated [Boyd et al., 2015, Forney and Rothman, 2012], such measurements are challenging *in situ* and would require extended deployments; to our knowledge these have not been attempted.

There is, however, one axis of complexity for which quantitative information is available: the settling velocity distribution (SVD, with respect to mass flux) at a reference depth – AoC1. Herein we focus on particulate organic carbon (POC) flux (see Discussion), as this is the flux of most interest to marine biogeochemists. To date, particle flux models have

neglected variability in particles' settling velocities, beyond in the coarsest way possible (i.e. by resolving two pools of sinking material – see Section 4.1). If, however, data types beyond bulk flux vs. depth are required to generate multiple constraints by which to compare such models, SVD data are an ideal candidate. Here we explore the implications of available SVD data for particle flux modeling.

4.2.3 Settling velocity distribution data

Indented Rotating Sphere Carousel (IRSC) sediment traps [Figure 4-5] can collect particles based on discrete settling-velocity ranges, and can therefore provide *in situ* SVD estimates [Peterson et al., 2005, Trull et al., 2008, Ebersbach et al., 2011]. IRSC traps use a cylinder to collect particles, then isolate trapped material from the collection portion of the trap to minimize washout of solid and dissolved materials as well as contamination of trapped particulate matter by free-swimming animals. IRSC traps are traditionally used for time-series observations of particle fluxes, rotating the indented rotating sphere and sample carousel simultaneously at a fixed time interval. However, by rotating the indented rotating sphere once each day and then rotating the sample carousel through eleven sample chambers at appropriate time intervals, *in situ* separation of particles by settling velocity can be accomplished [Peterson et al., 2005]. This modification corresponds to a range in settling velocities of up to three orders of magnitude; exact values depend on the particular study, but the ranges are on the order of $\sim 1\text{m/d}$ to $\sim 1000\text{ m/d}$ [Peterson et al., 2005, Trull et al., 2008, Ebersbach et al., 2011]. We note that this procedure does involve some sample manipulation that may bias the SVD measurements, as material is first collected on and then released from the indented rotating sphere; we will assume here that such bias is negligible. Though the deployment of these traps, especially in their 'settling velocity mode,' has been somewhat limited for logistical and historical reasons, data are available from the Mediterranean (MedFlux) [Peterson et al., 2005], subtropical (ALOHA) and subpolar (K2) Pacific [Trull et al., 2008], and Southern Ocean (P2) [Ebersbach et al., 2011]. Furthermore, as we will discuss below, these data tell a consistent story.

First, it is important to note the character of the POC flux distributions from different IRSC trap deployments [Figure 4-6]. Broadly speaking, all are similar in exhibiting a very wide range, and more importantly a non-negligible contribution to total POC flux from any individual measured window of settling velocities. In other words, although their distribu-

Figure 4-5: Schematic of an IRSC sediment trap in settling velocity mode; figure adapted from [Peterson et al., 2005]. Settling particles are collected by the intended rotating sphere (1), which then rotates 180° to dump material into the skewed funnel (2), which is then separated by settling velocity via successive rotations of the sample carousel (3).

Figure 4-6: Empirical cumulative distribution functions (CDFs) for the inverse of the settling velocity (m or $1/w$) with respect to POC flux from deployments of modified indented rotating sphere carousel sediment traps at three oceanographic stations at different times and depths. Compare to Figure 4 from [Trull et al., 2008]. Note that the ALOHA CDF only has five values because collection cups had to be aggregated due to low total flux [Trull et al., 2008]. Viewed right-to-left these distributions are also complementary CDFs for w .

tions quantitatively differ, it is apparent that particles with sinking speeds across the entire measured range contribute appreciably to total flux in each case. This is wholly inconsistent with a standard assumption implicit in idealized models of particle flux attenuation: that sinking matter can be treated as homogeneous, i.e. one or two pools of material (hence one or two characteristic settling velocities for particle flux – see Section 4.1). Any model of particle flux incorporating information about variation in settling velocities must therefore be able to incorporate – indeed, must be grounded in – this substantial variability.

4.2.4 Distributional model

Here we present a means by which to incorporate such variability in particle properties into a model for particle flux. We rely on an integral transform, which supposes that the total flux f [g/m²/d] can be described as the continuous superposition of the sinking and remineralization of particles' material subsets. Consider particle flux f to comprise a diverse continuum of sinking particles. At a shallow reference depth z_o [m] (e.g. 100m), each particle contributing to the total flux $f(z_o)$ has a settling velocity w [m/d]. Each particle is also comprised of a continuum of material, and each subset of material has an apparent first-order remineralization rate k [1/d] at that depth. For each subset of material within each particle, these w and k may vary with time or depth – the former case can be mapped into changing with depth because $z = \int_0^t w(s)ds$ (where s is a dummy variable). Thus each subset of material within each particle has an attenuation function $\alpha(z|k, w)$. The continuous superposition of these attenuation functions¹², weighted by the probability $q(k, w)$ that a unit of material will initially have the remineralization rate k and settling velocity w , then determines $f(z)$ via:

$$f(z) = f(z_o) \int_0^\infty \int_0^\infty q(k, w) \alpha(z|k, w) dk dw \quad (4.29)$$

Note that by assuming we can write total particle flux as a continuous superposition of individual particles' microscopic fluxes, we have neglected AoC4 above; Eq. (4.29) essentially formalizes the statement that total flux is the net result of the individual fluxes of particles' material subsets, and therefore that the depth-attenuation of total flux is the net result of how these that material subsets' individual fluxes attenuate. While the above equation is sufficiently general to allow for the consideration of the other three axes, it is intractable as written. We then make two major simplifying assumptions. Assume for simplicity that all sinking material is remineralized at approximately the same rate κ , i.e. $q(k) \approx \delta_\kappa(k)$, and that w and k for particles' material subsets are approximately constant with depth. In other words, assume that AoC1 determines the flux profile. These assumptions are discussed below. While there may be some evidence to support these [Hedges et al., 2001, Berelson, 2001, Villa-Alfageme and Others, 2016, Schneider et al., 2003], both can be relaxed and/or

¹²As the two subsets of material with the same initial k and w will not necessarily sink and remineralize through the water column the same, $\alpha(z|k, w)$ should be considered the *average* attenuation of material with that initial k and w .

tested, especially the first and arguably more contentious assumption; importantly the final result holds even under a more plausible description of $q(k)$ (see Section 4.2.7). Under these assumptions the attenuation function then becomes an exponential and we are then left with the simpler Laplace transformation:

$$f(z) = f(z_o) \int_0^\infty q(m) e^{-\kappa z m} dm = \mathcal{L}[f(z)] \quad (4.30)$$

where for compactness we have defined $m := 1/w$ [d/m] as the inverse settling velocity and absorbed κ into the definition of $\mathcal{L}[\cdot]$. Thus the Laplace transform maps the SVD¹³ into $f(z)$ and considers particle flux to be heterogeneous and linear¹⁴ (as opposed to homogeneous and nonlinear as for other particle flux models). The Laplace transform has been used successfully to model other flows within the carbon cycle [Forney and Rothman, 2012]; it is typically described for quantities that vary with time rather than depth, but the formalism here is otherwise similar. Notably, Eq. (4.30) generalizes each of the standard particle flux parameterizations, meaning for each of these representations of $f(z)$, there is a $q(m)$, referred to as its inverse Laplace transform ($q(m) = \mathcal{L}^{-1}[f(z)]$), such that Eq. (4.30) applied to that $q(m)$ yields that $f(z)$ model [Table 4.5].

Under the assumption then that the flux-depth relationship is dominantly determined by AoC1, we may specify three criteria by which to evaluate idealized particle flux models:

- the distribution $q(m)$ captures settling velocity distribution observations.
- the associated model $f(z) = \mathcal{L}[q(m)]$ captures flux profile observations.
- the parameter values are consistent between fits to either data type.

Here we compare the fits of different probability distributions $q(m)$ to the SVD data in Figure 4-6, and the fits of their Laplace transforms $f(z)$ to a large collection of profiles (see Section 4.1.8). We evaluate the standard particle flux models (exponential, power-law, ballast, rational, double exponential) from the literature as well as one additional model – the exponential integral (ExpInt). This model is the Laplace transform of the reciprocal distribution for m :

$$q(m) = \chi/m \quad (4.31)$$

¹³Technically the inverse of the SVD.

¹⁴i.e. all decay is modeled as exponential.

Model $f(z)$	Distribution $q(m)$
Exponential: $e^{-z/\ell}$	Delta: $\delta_\mu(m)$
Power: z^{-b}	Power: m^{b-1}
Ballast: $e^{-z/\ell} + c$	2× delta: $\delta_\mu(m) + \delta_0(m)$
Rational: $1/(z + \zeta)$	Exponential: $e^{-\zeta m}$
Double: $e^{-z/\ell_1} + e^{-z/\ell_2}$	2× delta: $\delta_\mu(m) + \delta_{\mu'}(m)$
ExpInt: $E_1(z/\ell_o)$	Reciprocal: m^{-1}

Table 4.5: Models of particle flux attenuation and their inverse Laplace transform, i.e. the distribution which when substituted into Eq. (4.30) yields that model.

where m is confined to the range $(\frac{1}{\max(w)}, \frac{1}{\min(w)})$ and χ is a normalizing constant dependent on this range. The Laplace transform of this distribution¹⁵ for $q(m)$ yields the ExpInt model:

$$f(z) = f_o E_1(z/\ell_o) \quad (4.32)$$

where $\ell_o = \max(w)/\kappa$, f_o is the flux at $z \approx \frac{1}{4}\ell_{max}$, and $E_1(\cdot)$ is the exponential integral:

$$E_1(x) := \int_x^\infty s^{-1} e^{-s} ds \quad (4.33)$$

(where s is again a dummy variable). It is also worth noting that the inverse Laplace transform is the canonical example of an ‘ill-conditioned inverse problem,’ meaning that very similar $f(z)$ profiles can map into very different $q(m)$ distributions, while very different $q(m)$ distributions can map into virtually identical $f(z)$ profiles via the Laplace transform [Forney and Rothman, 2012]. This gives reason to expect much larger differences in $q(m)$ data and in models’ ability to fit those data than in $f(z)$ data. It should also be noted that the particle flux models from the literature were not derived via Eq. (4.30) as here¹⁶ (see Section 4.1); these are largely either empirical or assume a particular behavior in k and a constant w . Here instead we are concerned with the other limit case, that the flux profile is determined only from variation in w .

The data in Figure 4-6 are clearly not well-described by a delta function or the sum of two delta functions. Barring a strong relationship between k and w , then, the exponential,

¹⁵ $\min(w)$ is largely irrelevant and therefore can be neglected in this formulation because its influence on the normalizing constant χ is absorbed into the parameter f_o in Eq. (4.32) and for a fixed κ material settling at small velocities $w \rightarrow 0$ will be remineralized near z_o .

¹⁶For instance, the ballast model [Armstrong et al., 2001] certainly was not derived by assuming a fraction of particulate material sinks infinitely fast!

ballast, and double exponential models are inconsistent with these data¹⁷. We believe such a relationship to be implausible, and that the most likely case is that w and k are independent or very weakly related, because particulate material likely undergoes numerous transformations before sinking out of the upper ocean [Burd and Jackson, 2009] and marine snow aggregates are made up of variable material [Fowler and Knauer, 1986], so both fast- and slow-sinking particles are likely to include both quickly- and slowly-remineralized material. However, the distributions associated with the power-law, rational, and ExpInt models are all broad and so may describe the SVD data well. In the following section we compare the fits of the distributions associated with each of these $f(z)$ models (the power-law, exponential, and reciprocal distribution, respectively). While a large number of possible distributions could potentially be fit to these data, we restrict ourselves to distributions that either are associated to existing particle flux models or which there is a reason to expect may apply.

4.2.5 Fit to settling velocity distribution data

Cumulative Distribution Function (CDF) statistics measure deviations between a hypothesized distribution and data's empirical CDF [Flannery et al., 1992, Stephens, 1974]. These statistics can therefore be used to quantitatively compare the ability of our different distributions (power-law, exponential, and reciprocal) of interest to capture the observed $q(m)$ distributions. We use the Kolmogorov-Smirnov (d), Kuiper (v), and Anderson-Darling (a) statistics (see Section 3.1.20). These differ in the extent to which they weigh median versus tail values; v is the most balanced, while d weighs median values most heavily and a weights tail values most heavily. We fit each hypothesized distribution to each of the empirical distributions in Figure 4-6, estimating the parameters by minimizing each statistic, and then compare the best-fitting distributions. For the reciprocal distribution, we choose $\max(w)$ as the maximum measured velocity and either choose $\min(w)$ as the minimum measured velocity or allow $\min(w)$ to be a free parameter. We do not fit the delta and double-delta distribution as the inability of these distributions to capture the $q(m)$ data is visually apparent. Figure 4-7 shows an example of a fit, where the CDF statistic used is Kuiper's statistic

¹⁷Indeed, because it is the remineralization lengthscale w/k and neither w nor k alone that determines the depth-attenuation of flux, any of the models on the left-hand-side of Table 4.5 can result from any of the distributions on the right-hand-side of Table 4.5 given the appropriate relationship between k and w . As stated above, we assume here that k and w are independent – this assumption not only is most plausible in our estimation but also is justifiable by parsimony. This assumption could be tested by separating particulate material by settling velocity, e.g. by elutriation [Peterson et al., 2005], and determining whether the bulk remineralization rate varied across different settling velocity classes.

v and the data are from station P2 at 320m [Ebersbach et al., 2011].

For all three of the CDF statistics, the distribution corresponding to the ExpInt model, i.e. the reciprocal distribution, fits best for six of seven distributions despite its fit having zero degrees of freedom (the distributions corresponding to the power and rational distributions have one degree of freedom each); see Table 4.6. Furthermore, the best-fit b -values for the power law model are all in the range $b \in (0.17, 0.44)$, much smaller than typically found by fitting $f(z)$ data. As expected¹⁸, differences in goodness-of-fit metrics for $q(m)$ data are generally very large in contrast to the small differences in goodness-of-fit between the models' fits to $f(z)$ data (see Section 4.1). When the reciprocal distribution is permitted a free parameter by allowing $\min(w)$ to vary rather than remain fixed, the ExpInt model fits best for all data and all CDF statistics.

We also fit a lognormal distribution $\ln\mathcal{N}(\mu, \sigma)$ to these data and found depending on the empirical CDF and statistic either the reciprocal distribution to fit better or the difference in fit between the lognormal and reciprocal distributions to be negligible given the former's two additional free parameters. In empirical applications, the reciprocal and lognormal distributions often behave similarly when σ of the latter is large enough [Montroll and Shlesinger, 1982]. A lognormal fit to these data and to the $f(z)$ data discussed next, via $\mathcal{L}[\ln\mathcal{N}(\mu, \sigma)]$, may also be also plausible but in our estimation the additional model complexity is not justified given the resolution of available data.

4.2.6 Fit to flux-depth data

Above we see that the distribution corresponding to the ExpInt model captures SVD data markedly better than those associated with standard particle flux models. Here we show that it also captures $f(z)$ observations equivalently or better. While the difference in different particle flux models' goodness-of-fit to individual profiles is in most instances not statistically significant [Section 4.1], it is still instructive to compare models' ability to fit large collections of profiles. Figure 4-8 shows the distribution of r^2 values that result from fitting all of the two-parameter (power, rational, exponential and ExpInt) models in Table 4.5 to 722 profiles

¹⁸Unfortunately, the standard techniques for calculating a p-value from these statistics are not appropriate here, because these data are not continuous and do not have a sample size. Instead, these statistics must be used heuristically to compare goodness-of-fit for different distributions – higher values meaning worse fits. In general p-values tend to decrease exponentially for smaller CDF-statistic values [Stephens, 1974].

Model	MedF'lux '03	MedF'lux '05	K2-1	K2-2	P2 170m	P2 320m	ALOHA
Power	$d=.169$ (.23)	$d=.138$ (.25)	$d=.264$ (.22)	$d=.172$ (.29)	$d=.189$ (.26)	$d=.126$ (.27)	$d=.082$ (.25)
	$v=.329$ (.17)	$v=.266$ (.18)	$v=.529$ (.23)	$v=.339$ (.23)	$v=.351$ (.37)	$v=.251$ (.26)	$v=.161$ (.20)
	$a=.283$ (.32)	$a=.222$ (.31)	$a=.591$ (.35)	$a=.328$ (.44)	$a=.553$ (.26)	$a=.173$ (.31)	$a=.096$ (.25)
Rational	$d=.284$ (.22)	$d=.263$ (.17)	$d=.124$ (.95)	$d=.255$ (.22)	$d=.176$ (.27)	$d=.186$ (.32)	$d=.056$ (.95)
	$v=.398$ (.93)	$v=.440$ (.97)	$v=.190$ (.160)	$v=.324$ (.65)	$v=.268$ (.16)	$v=.298$ (.66)	$v=.085$ (.81)
	$a=.603$ (.20)	$a=.550$ (.16)	$a=.570$ (.57)	$a=.545$ (.20)	$a=.369$ (.26)	$a=.411$ (.29)	$a=.123$ (.97)
ExpInt	$d=.157$	$d=.087$	$d=.329$	$d=.118$	$d=.127$	$d=.047$	$d=.037$
	$v=.193$	$v=.167$	$v=.329$	$v=.221$	$v=.147$	$v=.066$	$v=.037$
	$a=.157$	$a=.116$	$a=.362$	$a=.149$	$a=.373$	$a=.047$	$a=.037$
ExpInt free min(w)	$d=.120$ (1.6)	$d=.081$ (.66)	$d=.120$ (.23)	$d=.113$ (1.6)	$d=.069$ (3.4)	$d=.038$ (2.0)	$d=.013$ (1.0)
	$v=.170$ (.53)	$v=.111$ (.32)	$v=.184$ (.39)	$v=.216$ (2.4)	$v=.129$ (4.9)	$v=.062$ (3.0)	$v=.024$ (4.8)
	$a=.137$ (1.0)	$a=.087$ (.58)	$a=.212$ (9.4)	$a=.144$ (1.0)	$a=.106$ (4.1)	$a=.047$ (1.7)	$a=.014$ (3.7)

Table 4.6: Results of fitting the distributions corresponding to the power-law, rational, and ExpInt models to each CDF from Figure 4-6, using all three CDF statistics. Values in parentheses are the parameters for the best-fit distributions: b for the power-law model, ζ [m] for the rational model, the, and min(w) [m/d] for the ExpInt (free min(w)) model.

Figure 4-7: Fits of exponential, power-law, and reciprocal distributions (corresponding to rational, power-law, and ExpInt models) to one CDF from Figure 4-6. v is Kuiper's statistic.

of POC flux from Section 4.1¹⁹ [Cael and Bisson, 2018]. As these models all have the same number of free parameters, r^2 is a fair metric for this comparison. The ExpInt model's distribution of r^2 is nearly identical to that of the power-law model, with a median of 0.87 in both cases, whereas the rational and exponential models have medians of 0.83 and 0.82 respectively. Assuming $\max(w) = 1000\text{m/d}$, the median κ found from these fits is 0.44d^{-1} ; this value is reasonable and consistent with other estimates [Iversen and Ploug, 2010, McDonnell et al., 2015], but should only be considered order-of-magnitude accurate as the uncertainty is a factor of four, as estimated from the log-standard-deviation of the distribution of κ values.

Figure 4-8 also shows the comparison of the ExpInt model to the ballast and double exponential models. However, in these cases we only fit profiles with $n \geq 4$ and $n \geq 5$ measurements, respectively, so as to leave at least one degree of freedom in the fit (see Section 4.1), and because the model fits have different degrees of freedom, a modification of the r^2 statistic is necessary. The F -statistic is commonly used in this application [Twardowski et al., 2004]:

$$F := \frac{r^2/D_f}{(1-r^2)/D_e} \quad (4.34)$$

where D_f is the degrees of freedom of the model fit and D_e is the degrees of freedom of the error, i.e. $n - D_f$. By this metric the ExpInt model is superior to both the ballast and double exponential models. The median F value for the ballast model is 2.6 (vs. 4.4 for the ExpInt model for profiles with $n \geq 4$) and 0.87 for the double exponential model (vs. 5.3 for the ExpInt model for profiles with $n \geq 5$). We also note that the power-law fits to profiles result in a much broader range of b values, ranging from ~ 0 to 2.9 with a median of 0.65 (as compared to a median of 0.26 for the fits to $q(m)$ data). When b is restricted to $b \in (0.17, 0.44)$ as for the fits to $q(m)$ above, the median r^2 decreases from 0.87 to 0.61 – far worse than the other two-parameter models.

Thus the ExpInt model not only fits $q(m)$ better than other models, but also fits $f(z)$ data as well or better than other models, while the parameters are consistent between the fits to different data types and with other estimates of the bulk first-order remineralization rate κ .

¹⁹We fit these profiles according to the same procedures as in Section 4.1. It is also important to note that zooplankton-related fluxes are included in these measurements, which we do not consider here.

Figure 4-8: a) Probability density functions (PDFs) of r^2 values for fits of the power-law, rational, exponential, and ExpInt models to the 722 profiles in the database described in Section 4.1. b,c) PDFs of F values for fits of the ballast (double) and ExpInt models to the profiles with $n \geq 4$ ($n \geq 5$) measurements – total of 277 (187) profiles. In each case, PDFs are estimated using the kernel method [Hill, 1985].

4.2.7 Discussion

What about variation in k ? Above we have assumed for simplicity that all sinking material can be modeled as remineralized at the same rate, $q(k) \approx \delta_\kappa(k)$. This is a major simplifying assumption – in reality it is much more likely that the distribution of remineralization rates is quite broad. This assumption may be relaxed: assuming still that k and w are independent, if another distribution is substituted for δ_κ we can modify Eq. (4.30) according to:

$$f(z) = f(z_o) \int_0^\infty q(m)q(k)e^{-kzm} dkdm \quad (4.35)$$

which still constitutes a mechanistic and distributional model for particle flux, though it may not have a simple closed form such as the ExpInt model. However, the most plausible distribution to assume for k itself may be a reciprocal distribution, i.e. $q(k) \sim 1/k$ [Rothman and Forney, 2007]. In this case, because the product distribution of independent reciprocal distributions is itself a reciprocal distribution, the ExpInt model still holds – though this of course changes the interpretation of the lengthscale parameter ℓ_o .

Why the reciprocal distribution for w ? The above analyses raise the question: why does $q(m)$ appear to be best described as a reciprocal distribution? A simple probabilistic argument may explain (though we note the connection is somewhat abstract and therefore tenuous). The Central Limit Theorem, which gives rise to normal and lognormal distributions, is one of the most powerful and useful concepts in probability theory [Pitman, 1999] (see Chapter 3). It has an obverse: random samples chosen from random probability distributions are collectively described by the reciprocal distribution [Friar et al., 2016, Hill, 1995, Berger and Hill, 2008, Hamming, 1970]. In other words, if random samples are taken from random probability distributions (in such a way that the overall process is scale neutral), and the results combined, then the resulting combined samples converge to the reciprocal distribution. Thus as the normal distribution is a robust result of aggregating samples, the reciprocal distribution is the robust result of aggregating distributions – the reciprocal distribution is therefore sometimes termed “the distribution of distributions” [Friar et al., 2016]. (This phenomenon is best known for giving rise to the logarithmic distribution of first digits in many types of data, known as Benford’s Law, but underlying this phenomenon is the reciprocal distribution.)

By analogy we may then argue that if total flux comprises a variety of different sorts

of particulate material generated by different sorts of processes – marine snow aggregates of various types, fecal pellets from various zooplankton species, intact organisms of various types [Fowler and Knauer, 1986] – we may expect the SVD to approximate a reciprocal distribution via conflating these different heterogeneous types of material as a single particulate class. While the velocity distributions associated with any one of these may take many different forms, if there are a large number of different types of processes generating these particles, their total distribution $q(m)$ (or equivalently $q(w)$) will converge to a reciprocal distribution. To illustrate, imagine that a large number of different processes $Q_1, Q_2, \dots, Q_n$ generate sinking particulate material. Each process will have an associated probability distribution $q_1(w), q_2(w), \dots, q_n(w)$ for the settling velocity of the material it generates. Now choose a random attogram ($=10^{-18}\text{g}$) a_i of sinking material. That a_i is generated by a random process Q_i , so its velocity is drawn from the distribution $q_i(w)$. Some other randomly chosen attogram a_j will be generated by Q_j and so its velocity will be randomly drawn from $q_j(w)$. If we take a large number of attograms $a_i, a_j, a_k, \dots$, their individual settling velocities will have to be defined by some total probability distribution $q(w)$. If the sub-distributions $q_1, q_2, \dots, q_n$ in aggregate lack a particular scale for settling velocity, this distribution will tend towards the reciprocal distribution in the limit of a large number of generating processes n – even if each sub-distribution (e.g. $q_1(w)$) is very different from a reciprocal distribution and even if these sub-distributions are not evenly sampled [Berger and Hill, 2008].

What about other complexities? The model we propose, even in its most general formulation, i.e. Eq. (4.29), is a gross oversimplification of true particle fluxes, as any idealized model necessary is. Additionally, sinking particulate material is only one means by which organic matter is transported in the ocean or by which the ocean’s vertical gradient in dissolved inorganic carbon is maintained [Volk and Hoffert, 1985]. Lateral advection [Siegel and Deuser, 1997], eddy subduction of both particulate and dissolved matter [Omand et al., 2015], zooplankton grazing and active transport [Turner, 2015], particle-particle interactions [Burd and Jackson, 2009], and numerous other processes undoubtedly play important roles. Furthermore, idealized models should only be considered applicable on large scales and cannot capture the substantial horizontal and temporal variability exhibited by particle flux measurements [Estapa et al., 2015]. Our intention here is to posit a model simple enough to implement in Earth System Models and to make straightforward predictions that are

testable *in situ*. In Earth System Models resolving complex ecosystems, rather than a simple two-parameter particle flux parameterization it could potentially be preferable to resolve a distribution $q(m)$ non-parametrically by generating a numerical sum-of-exponentials distribution to input into Eq. (4.30) whose characteristics depended on those of the material comprising the particulate pool. We also note that here we have referred very generally to particle fluxes but then have applied these models to POC flux data. The models we discuss above are equally applicable to other elements or to total particulate organic matter. Fluxes of different elements and material pools may be expected to have different κ values due to preferential remineralization [Schneider et al., 2003]. Based on the argument described next for why the reciprocal distribution is observed, the ExpInt model may not hold for fluxes of materials generated from specific processes, e.g. biogenic silica.

Testable hypotheses. One exciting aspect of the model presented here is that it is testable *in situ*; because the model is mechanistic (specifying f in terms of the processes that are thought to govern it) its validation or rejection can provide insight into the dominant processes determining how particle flux works. We may deduce the following hypotheses:

H1. The settling velocity w (and equivalently its inverse m) follows a reciprocal distribution with respect to mass, i.e. $q(m|z_o) \propto m^{-1}$

H2. The remineralization rate k varies little relative to w , i.e. $q(k) \approx \delta_k$

H3. Measurements at a reference depth – $f(z_o)$, $\kappa(z_o)$ and $q(m|z_o)$ – predict the flux attenuation beneath that depth, $f(z > z_o)$

H4. The depth-evolution of the settling velocity distribution is also predictable, i.e. $q(m|z > z_o) \propto q(m|z = z_o)e^{-\kappa(z-z_o)m}$

The first two hypotheses are just restatements of the assumptions that lead to Eq. (4.30) and the third is just a restatement of Eq. (4.30). The last hypothesis is a stricter test of the assumption that we may model particle flux via a Laplace transform, requiring the predictable depth-evolution of not only the total flux but also of the SVD. The simultaneous measurement of the relevant quantities would provide a much more detailed understanding of the flux-depth relationship and its governing processes and properties. Discrepancies between the measured and predicted $f(z)$ or $q(m|z)$ could potentially indicate that sinking speeds of individual particles do change appreciably with depth (AoC3), which could be evaluated by applying other integral transformations, i.e. using a different function than an exponential for $a(z|k, w)$ in Eq. (4.29) that captured depth-change(s) in w and/or k .

Measuring κ at additional depths would be particularly useful in refining the distributional approach taken here, as would the substantially more difficult measurements of $q(k)$ (e.g. by a very long deployment of a RESPIRE sediment trap [Boyd et al., 2015]), the relationship between w and an average k , or even the joint distribution $q(k, w)$. In the case where H1 or H2 is rejected, one can still test H3+4 either analytically by fitting other distributions to $q(m)$ and/or $q(k)$ or empirically by using data directly as an empirical distribution. Furthermore, properties of $q(m|z_o)$ may also be related to properties of the surface ecosystem producing sinking material, such as the particle size distribution, which may help account for e.g. why the anomalous K2-1 distribution in Figure 4-6 is so different from the K2-2 distribution measured only ~ 10 days later (but in end-of-bloom conditions rather than mid-bloom conditions).

Conclusion.

Here we have argued that the exponential integral model is a natural consequence of the assumptions that i) a large number of different generating processes contribute to total particulate material and ii) particle flux attenuation is controlled by the initial properties of this material. We have shown that this model is unique in capturing observations of both the settling velocity distribution and of particle flux vs. depth. We therefore propose that this new model be used to parameterize particle fluxes in Earth System Models and that it be tested in the field.

Chapter 5

Summary and future directions

This thesis has explored the relationships between the flux of marine particulate organic carbon (POC) and a few other key oceanographic variables. In particular, it has been shown that including the variability or heterogeneity of certain processes into mathematical models of these relationships via probability distributions can improve characterizations of existing observations. Treating export efficiency as a random variable mechanistically scaled by temperature allows one to extract the temperature dependence in export efficiency observations and to estimate an average metabolically-driven change in export efficiency due to a given temperature change. Collections of measurements of net primary production and POC flux can be more straightforwardly related than individual measurements thereof. Incorporating heterogeneity in the settling velocities of POC suggests a model for the depth-attenuation of POC flux that is mechanistic, testable, and better constrained. In this closing chapter several potential avenues for future research – some theoretical, some observational; some immediate, some more expansive – are outlined.

One interesting possible application of the model developed in Chapter 2 is to glacial-interglacial changes atmospheric carbon dioxide concentration and the influence of the biological carbon pump. From the box model in Section 2.2 a scaling was derived that the metabolic mechanism considered therein should result in a ~ 7 ppm K^{-1} decrease in atmospheric CO_2 . Combined with an approximate ~ 10 ppm CO_2 K^{-1} decrease due to the solubility pump, this predicts the ~ 70 ppm drawdown resulting from a ~ 4 K cooling during the Last Glacial Maximum. Incorporation of this biosphere-climate feedback into a model of glacial-interglacial CO_2 changes may improve understanding of the relationship

between these two all-important paleorecords. Several mechanisms have been suggested to relate changes in atmospheric CO₂ to global temperature in paleoclimate records; while the biological carbon pump is widely acknowledged to play an important role, the specific mechanism(s) by which it does so are unclear [Sigman and Boyle, 2000]. Because there are multiple, not necessarily mutually exclusive, explanations of this drawdown, a Bayesian approach to determine the most likely combination(s) of factors would be preferable. This would require a sufficiently simple representation of the ocean circulation and the carbon cycle to permit a large ensemble of simulations, into which various CO₂ drawdown mechanisms could be embedded via changes in the appropriate parameters (ascribed certain prior distributions). Such an exercise would result in probability distributions (really, a joint distribution) for the relative importances of different mechanisms for glacial-interglacial CO₂ drawdown, i.e. would suggest – under certain assumptions! – how important different processes were most likely to have been for atmospheric CO₂ variations in the past.

It would also be instructive to repeat the analyses in Section 2.1 with a larger and more self-consistent dataset than the compilation used therein, to include a relationship with primary production as well as temperature in such an analysis, or to utilize additional laboratory data, particularly on zooplankton growth rates’ temperature response. The intention in Section 2.1 was to keep things as simple as possible and to illustrate the idea developed therein with published datasets. The shallow POC flux dataset compiled in Chapter 3, however, combined with the method for estimating export efficiency employed by Henson et al. [Henson et al., 2011] (i.e. retrieving a depth-integrated net primary production estimate via a satellite algorithm wherever a POC flux measurement was made; sea surface temperature could also be retrieved in this way) would provide a dataset $\sim 4\times$ larger than that used in Section 2.1 and collected in a more consistent manner. It would be interesting to see whether the model proposed in Section 2.1 was consistent with such a dataset. Given sufficient co-located temperature, POC flux, and net primary production data, one could also combine the results of Section 2.1 and Chapter 3 to describe POC flux as a random variable whose probability distribution shifted both with net primary production and temperature. A promising indicator for this possibility is that temperature-rescaled POC fluxes are more tightly related to depth-integrated net primary production than unscaled POC fluxes for the in situ measurements used in Section 2.1. The issue with this sort of exercise however is of course that satellite primary productivity algorithms are models, not measurements,

and different satellite algorithms exhibit large differences on regional and seasonal scales [Bisson et al., 2018]. Furthermore, these POC flux measurements were made at different depths, which as discussed in Chapter 4 is not an easily resolvable problem. Regardless of the necessary caveats, however, this sort of analysis could be informative and synthesize some of the core results of Chapters 2 and 3. Another means by which to refine the model of Section 2.1 would be to utilize additional laboratory data to better estimate the differential temperature sensitivity parameter (for which the symbol and value $\beta = 0.047 \text{ K}^{-1}$ was used) and to examine the behavior of this temperature sensitivity at high temperatures ($\geq 25^\circ\text{C}$). The original values for the temperature sensitivities of phytoplankton and zooplankton of Eppley [Eppley, 1972] and Huntley and Lopez [Huntley and Lopez, 1992] respectively were used, but an appreciable amount of data has since been collected that could readily be used to make a more quantitative estimate for β and to assess how this differential temperature sensitivity itself changes with temperature – thus refining and extending the applicability of the model.

Several future research questions and applications are posed in Chapter 3, but certainly the most crucial of these are the related questions of subdivision and scale in the ocean. Measurements were coarsely split into very large regions for pragmatism and illustration, but how large of a spatiotemporal region can we consider a given measurement to represent, and what is the most appropriate way to subdivision the ocean into biomes? Both of these questions are long-standing and fundamental to oceanography, and their answers of course depend on the processes being studied and the scientific questions being addressed. Resolving either in the context of global POC flux out of the surface ocean requires a sense of the variation in POC flux and its covariates on a range of spatial and temporal scales. Analysis of high-resolution spatial or temporal measurements, such as those that are expected to become available via the EXPORTS program [Siegel and Others, 2016], may be quite useful in this regard. What evidence is available suggests appreciable variation in export on small (kilometer) scales [Estepa et al., 2015], evincing the need for a probabilistic perspective. Modeling studies coupled with spatial statistics techniques are likely the best avenue for determining the appropriate oceanic subdivisions, as is one of the objectives of the CBIOMES collaboration. As the ocean will continue to be perennially undersampled, numerical models can be utilized to identify how different spatiotemporal regions of the ocean cluster ecologically and biogeochemically. These regions could then be used as the

subregions defined in Chapter 3, and the statistics of their productivity and export could then be explored. Some of these regions are likely to be much more sparsely sampled than others, which could help identify where to focus future observational efforts; data from better-sampled regions would provide useful smaller-scale tests for the theory described in Chapter 3.

Another key pair of questions arising from Chapter 3 is whether the log-moments of different regions' productivity distributions and the $\sim 2/3$ scaling between the log-moments of POC flux and net primary production can be explained. Average productivity in a region might plausibly be explicable in terms of nutrient supply and light availability; this could certainly be tested using existing climatological information. The production-flux scaling could result from a diverse set of phenomena: anything from being an emergent property of the marine microbial ecosystem producing the material, to a simple consequence of the measurements having different dimensions or the flux measurement integrating over and thus damping the variation in the production measurement. Idealized theoretical and numerical models of different possible means by which to explain this scaling should be explored to assess their plausibility. As alluded to in Chapter 3, predicting these log-moments and their relationships would allow for a highly simplified yet accurate description of net primary production and POC flux globally.

The results of Chapter 4 illustrate that novel suites of observations are central to improving the understanding of the depth-attenuation of POC flux. Section 4.2 suggests a combination of measurements that would allow one to test the mechanistic model described therein. In particular, measuring in addition to POC flux i) the settling velocity distribution and ii) the bulk remineralization rate would yield a fully constrained model that would allow one to test to what extent POC flux attenuation characteristics are explicable in terms of heterogeneity in POC settling velocity. Making these measurements at multiple depths would significantly increase the power of such a test. Excitingly, this suite of observations is possible to measure; rotating indented sphere carousel traps provide an estimate of the settling velocity distribution in addition to measuring POC flux, while several techniques exist to measure remineralization rates. Ideally one could probe the distribution of remineralization rates of POC, though this would require a substantially extended deployment of the available instrumentation and may therefore be infeasible. Combining these measurements with other covariates thought to be mechanistically tied to settling velocity and/or

rem mineralization rate (e.g. the particle size distribution [Guidi et al., 2009], temperature [Cram et al., 2018]) could be fruitful, especially in linking properties of the surface ocean ecosystem to the fate of the material it produces.

In general it is clear from the analyses herein that when it comes to particle fluxes in the ocean, we are still very much data-limited. Large swaths of the ocean remain unsampled or undersampled, especially in the context of the variability and complexity of the environmental processes in which we have been interested. Where there are measurements, these are insufficient to parse between divergent mechanistic descriptions of these processes. Thinking carefully about where and when to make measurements so as to be most informative biogeographically – and what sets of simultaneous measurements to make so as to provide sufficient constraints to characterize the processes of interest – is essential. This thesis has provided some direction along these lines; perhaps developing further this question of what measurements to make when and where will be most important for improving the global characterization of ocean particle fluxes.

Bibliography

- [Adkins et al., 2002] Adkins, J. F., McIntyre, K., and Schrag, D. P. (2002). The salinity, temperature and $\delta^{18}\text{O}$ of the glacial deep ocean. *Science*, 298:1769–1773.
- [Alldredge and Gotschalk, 1988] Alldredge, A. L. and Gotschalk, C. (1988). In situ settling behavior of marine snow. *Limnology & Oceanography*, 33:339–351.
- [Alldredge et al., 1990] Alldredge, A. L., Granata, T. C., Gotschalk, C. C., and Dickey, T. D. (1990). The physical strength of marine snow and its implications for particle disaggregation in the ocean. *Limnology & Oceanography*, 35:1415–1428.
- [Alldredge and Silver, 1988] Alldredge, A. L. and Silver, M. W. (1988). Characteristics, dynamics and significance of marine snow. *Progress in Oceanography*, 20:41–82.
- [Amiel et al., 2002] Amiel, D., Cochran, J. K., and Hirschberg, D. J. (2002). $^{234}\text{Th}/^{238}\text{U}$ disequilibrium as an indicator of the seasonal export flux of particulate organic carbon in the north water. *Deep Sea Research Part II*, 49:5191–5209.
- [Antia, 2005] Antia, A. (2005). Solubilization of particles in sediment traps: revising the stoichiometry of mixed layer export. *Biogeosciences*, 2:189–204.
- [Archer et al., 2000] Archer, D. E., Eshel, G., Winguth, A., Broecker, W., Pierrehumbert, R., Tobis, M., and Jacob, R. (2000). Atmospheric $p\text{CO}_2$ sensitivity to the biological pump in the ocean. *Global Biogeochemical Cycles*, 14:1219–1230.
- [Armstrong et al., 2001] Armstrong, R. A., Lee, C., Hedges, J. I., Honjo, S., and Wakeham, S. G. (2001). A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals. *Deep Sea Research Part II*, 49:219–236.
- [Bacon et al., 1996] Bacon, M. P., Cochran, J. K., Hirschberg, D., Hammar, T. R., and Fler, A. P. (1996). Export flux of carbon at the equator during the EqPac time-series cruises estimated from ^{234}Th measurements. *Deep Sea Research Part II*, 43:1133–1153.
- [Baena et al., 2008] Baena, A. M. R., Boudjenoun, R., Fowler, S. W., Miquel, J. C., Masque, P., Sanches-Cabez, J. A., and Warnau, M. (2008). ^{234}Th -based carbon export during an ice-edge bloom: Sea-ice algae as a likely flux in data interpretation. *Earth and Planetary Science Letters*, 269:596–604.
- [Banse, 1990] Banse, K. (1990). New views on the degradation and disposition of organic particles as collected by sediment traps in the open ocean. *Deep Sea Research Part A*, 37:1177–1195.

- [Banse, 1992] Banse, K. (1992). Grazing, temporal changes of phytoplankton concentrations, and the microbial loop in the open sea. In Falkowski, P. and Woodhead, A., editors, *Primary productivity and biogeochemical cycles in the sea*, pages 409–440. New York: Plenum Press.
- [Bar-On et al., 2018] Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on earth. *Proceedings of the National Academy of Sciences*, 115:11842.
- [BATS:,] BATS:. Bermuda atlantic time series sediment trap data. available at <<http://batsftp.bios.edu/BATS/flux/>>, accessed on 20 April 2016.
- [Behrenfeld et al., 2016] Behrenfeld, M. J. et al. (2016). Revaluating ocean warming impacts on global phytoplankton. *Nature Climate Change*, 6:323–330.
- [Behrenfeld and Falkowski, 1997] Behrenfeld, M. J. and Falkowski, P. G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology & Oceanography*, 42:1–20.
- [Berelson, 2001] Berelson, W. M. (2001). Particle settling rates increase with depth in the ocean. *Deep Sea Research Part II*, 49:237–251.
- [Berger and Hill, 2008] Berger, A. and Hill, T. P. (2008). A basic theory of Benford’s law. *Probability surveys*, 8:1–126.
- [Berges, 1997] Berges, J. A. (1997). Ratios, regression statistics, and “spurious” correlations. *Limnology & Oceanography*, 42:1006–1007.
- [Bianchi et al., 2018] Bianchi, D., Weber, T. S., Kiko, R., and Deutsch, C. (2018). Global niche of marine anaerobic metabolisms expanded by particle microenvironments. *Nature Geoscience*, 11:263.
- [Bisson et al., 2018] Bisson, K., Siegel, D. A., DeVries, T., Cael, B. B., and Buesseler, K. O. (2018). How data set characteristics influence ocean carbon export models. *Global Biogeochemical Cycles*, 32:1312–1328.
- [Boyd, 2015] Boyd, P. W. (2015). Toward quantifying the response of the oceans’ biological pump to climate change. *Frontiers in Marine Science*, 2:77.
- [Boyd et al., 2015] Boyd, P. W., McDonnell, A. M. P., Valdez, J., LeFevre, D., and Gall, M. P. (2015). RESPIRE: An in situ particle interceptor to conduct particle remineralization and microbial dynamics studies in the oceans’ twilight zone. *Limnology & Oceanography: Methods*, 13:494–508.
- [Boyd and Trull, 2007] Boyd, P. W. and Trull, T. W. (2007). Understanding the export of biogenic particles in oceanic waters: Is there consensus? *Progress in Oceanography*, 72:276–312.
- [Britten et al., 2017] Britten, G. L., Wakamatsu, L., and Primeau, F. W. (2017). The temperature-ballast hypothesis explains carbon export efficiency observations in the southern ocean. *Geophysical Research Letters*, 44:1831–1838.
- [Buesseler, 1991] Buesseler, K. O. (1991). Do upper-ocean sediment traps provide an accurate record of particle flux? *Nature*, 353:420.

- [Buesseler, 1998] Buesseler, K. O. (1998). The decoupling of production and particular export in the surface ocean. *Global Biogeochemical Cycles*, 12:297–310.
- [Buesseler, 2004] Buesseler, K. O. (2004). ²³⁴Th based poc export. Biological and Chemical Oceanography Data System, available at <<http://www.bco-dmo.org/dataset/3164>>, accessed on 2 Nov 2015.
- [Buesseler et al., 1995] Buesseler, K. O., Andrew, J. A., Hartman, M. C., Belostock, R., and Chai, F. (1995). Regional estimates of the export flux of particulate organic carbon derived from thorium-234 during the JGOFS EqPac program. *Deep Sea Research Part II*, 42:777–804.
- [Buesseler et al., 1998] Buesseler, K. O., Ball, L., Andrews, J., Benitez-Nelson, C., Belostock, R., Chai, F., and Chao, Y. (1998). Upper ocean export of particulate organic carbon in the arabian sea derived from thorium-234. *Deep Sea Research Part II*, 5:2461–2487.
- [Buesseler et al., 2003] Buesseler, K. O., Barber, R. T., Dickson, M. L., Hiscock, M. R., Moore, J. K., and Sambrotto, R. (2003). The effect of marginal ice-edge dynamics on production and export in the southern ocean along 170W. *Deep Sea Research Part II*, 50:579–603.
- [Buesseler and Boyd, 2009] Buesseler, K. O. and Boyd, P. W. (2009). Shedding light on processes that control particle export and flux attenuation in the twilight zone of the open ocean. *Limnology & Oceanography*, 54:1210–1232.
- [Buesseler et al., 1994] Buesseler, K. O., Michaels, A. F., Siegel, D. A., and Knap, A. H. (1994). A three-dimensional time-dependent approach to calibrating sediment trap fluxes. *Global Biogeochemical Cycles*, 8:179–193.
- [Buesseler and Others, 2000] Buesseler, K. O. and Others (2000). A comparison of the quantity and quality of material caught in a neutrally buoyant versus surface-tethered sediment trap. *Deep Sea Research Part I*, 47:277–294.
- [Buesseler and Others, 2001] Buesseler, K. O. and Others (2001). Upper ocean export of particulate organic carbon and biogenic silica in the southern ocean along 170W. *Deep Sea Research Part II*, 48:4275–4297.
- [Buesseler and Others, 2006] Buesseler, K. O. and Others (2006). An assessment of particulate organic carbon to thorium-234 ratios in the ocean and their impact on the application of ²³⁴Th as a POC flux proxy. *Marine Chemistry*, 100:213–233.
- [Buesseler and Others, 2007] Buesseler, K. O. and Others (2007). An assessment of the use of sediment traps for estimating upper ocean particle fluxes. *Journal of Marine Research*, 65:345–416.
- [Buesseler and Others, 2008] Buesseler, K. O. and Others (2008). Vertigo (vertical transport in the global ocean): A study of particle sources and flux attenuation in the north pacific. *Deep Sea Research Part II*, 55:1522–1539.
- [Buesseler et al., 2009] Buesseler, K. O., Pike, S., Maiti, K., Lamborg, C. H., Siegel, D. A., and Trull, T. W. (2009). Thorium-234 as a tracer of spatial, temporal and vertical variability in particle flux in the north pacific. *Deep Sea Research Part I*, 56:1143–1167.

- [Buitenhuis et al., 2013] Buitenhuis, E. T., Hashioka, T., and Quéré, C. L. (2013). Combined constraints on global ocean primary production using observations and models. *Global Biogeochemical Cycles*, 27:847–858.
- [Burd and Jackson, 2009] Burd, A. B. and Jackson, G. A. (2009). Particle aggregation. *Annual Reviews of Marine Science*, 1:65–90.
- [Burd and Others, 2010] Burd, A. B. and Others (2010). Assessing the apparent imbalance between geochemical and biochemical indicators of meso- and bathypelagic biological activity: What the @\$#! is wrong with present calculations of carbon budgets? *Deep Sea Research Part II*, 57:1557–1571.
- [Cael and Bisson, 2018] Cael, B. B. and Bisson, K. (2018). Particle flux parameterizations: quantitative and mechanistic similarities and differences. *Frontiers in Marine Science*, 5:395.
- [Cael et al., 2018] Cael, B. B., Bisson, K., and Follett, C. L. (2018). Can rates of ocean primary production and biological carbon export be related through their probability distributions? *Global Biogeochemical Cycles*, 31:954–970.
- [Cael et al., 2017] Cael, B. B., Bisson, K., and Follows, M. J. (2017). How have recent temperature changes affected the efficiency of ocean biological carbon export? *Limnology & Oceanography: Letters*, 2:113–118.
- [Cael and Follows, 2016] Cael, B. B. and Follows, M. J. (2016). On the temperature dependence of oceanic export efficiency. *Geophysical Research Letters*, 43:5170–5175.
- [Cai and Others, 2008] Cai, P. and Others (2008). A high-resolution study of particle export in the southern south china sea based on ^{234}Th : ^{238}U disequilibrium. *Journal of Geophysical Research*, 113:C04019.
- [Cai et al., 2010] Cai, P., van der Loeff, R. M., Stimac, I., Nöthig, E. M., Lepore, K., and Moran, S. (2010). Low export flux of particulate organic carbon in the central arctic ocean as revealed by ^{234}Th : ^{238}U disequilibrium. *Journal of Geophysical Research*, 115:C10037.
- [Cai et al., 2001] Cai, P., Yipu, H., Min, C., Guangshan, L., and Yusheng, Q. (2001). Export of particulate organic carbon estimated from ^{234}Th - ^{238}U disequilibria and its temporal variation in the south china sea. *Chinese Science Bulletin*, 46:1722–1726.
- [Campbell, 1995] Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research*, 100:13237–13254.
- [Charette and Moran, 1999] Charette, M. A. and Moran, S. B. (1999). Rates of particle scavenging and particulate organic carbon export estimated using ^{234}Th as a tracer in the subtropical and equatorial atlantic ocean. *Deep Sea Research Part II*, 46:885–906.
- [Charette et al., 1999] Charette, M. A., Moran, S. B., and Bishop, J. K. (1999). ^{234}Th as a tracer of particulate organic carbon export in the subarctic northeast pacific ocean. *Deep Sea Research Part II*, 46:2833–2861.
- [Charette et al., 2009] Charette, M. A., Moran, S. B., Pike, S. M., and Smith, J. N. (2009). Investigating the carbon cycle in the gulf of maine using the natural tracer thorium 234. *Journal of Geophysical Research*, 106:11553–11579.

- [Chen et al., 2003] Chen, M., Huang, Y., Cai, P., and Guo, L. (2003). Particulate organic carbon export fluxes in the Canada basin and Bering sea as derived from $^{234}\text{Th}/^{238}\text{U}$ disequilibria. *Arctic*, 56:32–44.
- [Cochran and Others, 2000] Cochran, J. K. and Others (2000). Short-lived isotopes (^{234}Th , ^{228}Th) as indicators of POC export and particle cycling in the Ross sea, Southern Ocean. *Deep Sea Research Part II*, 47:3491–3520.
- [Collier et al., 2000] Collier, R., Dymond, J., Honjo, S., Manganini, S., Francois, R., and Dunbar, R. (2000). The vertical flux of biogenic and lithogenic material in the Ross sea: moored sediment trap observations 1996–1998. *Deep Sea Research Part II*, 47:3491–3520.
- [Collins et al., 2011] Collins, L. E., Berelson, W., Hammond, D. E., Knapp, A., Schwartz, R., and Capone, D. (2011). Particle fluxes in San Pedro basin, California: A four-year record of sedimentation and physical forcing. *Deep Sea Research Part I*, 58:898–914.
- [Coppola et al., 2006] Coppola, L., Roy-Barman, M., Mulso, S., Povinec, P., and Jeandel, C. (2006). Low particulate organic carbon export in the frontal zone of the Southern Ocean (Indian sector) revealed by ^{234}Th . *Deep Sea Research Part I*, 52:51–68.
- [Cox et al., 2000] Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., and Totterdell, I. J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408:184–187.
- [Cram et al., 2018] Cram, J. A., Weber, T., Leung, S. W., McDonnell, A. M. P., Liang, J. H., and Deutsch, C. (2018). The role of particle size, ballast, temperature, and oxygen in the sinking flux to the deep sea. *Global Biogeochemical Cycles*, 32:858–876.
- [Danovaro et al., 2011] Danovaro, R., Corinaldesi, C., Dell’Anno, A., Fuhrman, J. A., Middelburg, J. J., Noble, R. T., and Suttle, C. A. (2011). Marine viruses and global climate change. *FEMS Microbiology Reviews*, 35:993–1034.
- [de Boyer Montégut et al., 2004] de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, 109:C12.
- [de la Rocha and Passow, 2007] de la Rocha, C. and Passow, U. (2007). Factors influencing the sinking of POC and the efficiency of the biological carbon pump. *Deep Sea Research Part II*, 54:639–658.
- [Dee and Others, 2011] Dee, D. P. and Others (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137:553–597.
- [Deutsch et al., 2015] Deutsch, C., Ferrel, A., Seibel, B., Pörtner, H. O., and Huey, R. B. (2015). Climate change tightens a metabolic constraint on marine habitats. *Science*, 348:1132–1135.
- [Devol and Hartnett, 2001] Devol, A. H. and Hartnett, H. E. (2001). Role of the oxygen-deficient zone in transfer of organic carbon to the deep ocean. *Limnology & Oceanography*, 46:1684–1690.

- [Dittmar, 2008] Dittmar, T. (2008). The molecular level determination of black carbon in marine dissolved organic matter. *Organic Geochemistry*, 39:396–407.
- [Dugdale and Goering, 1967] Dugdale, R. C. and Goering, J. J. (1967). Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnology & Oceanography*, 12:196–206.
- [Dunne et al., 2005] Dunne, J. P., Armstrong, R. A., Gnanadesikan, A., and Sarmiento, J. L. (2005). Empirical and mechanistic models for the particle export ratio. *Global Biogeochemical Cycles*, 19:GB4026.
- [Durham et al., 2013] Durham, W. M., Climent, E., Barry, M., Lillo, F. D., Boffetta, G., Cencini, M., and Stocker, R. (2013). Turbulence drives microscale patches of motile phytoplankton. *Nature Communications*, 4:2148.
- [Durkin et al., 2015] Durkin, C. A., Estapa, M. L., and Buesseler, K. O. (2015). Observations of carbon export by small sinking particles in the upper mesopelagic. *Marine Chemistry*, 175:72–81.
- [Ebersbach et al., 2011] Ebersbach, F., Trull, T. W., Davies, D. M., and Bray, S. G. (2011). Controls on mesopelagic particle fluxes in the sub-antarctic and polar frontal zones in the southern ocean south of australia in summer—perspectives from free-drifting sediment traps. *Deep Sea Research Part II*, 58:2260–2276.
- [Efron, 1979] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7:1–26.
- [Eppley, 1972] Eppley, R. W. (1972). Temperature and phytoplankton growth in the sea. *Fisheries Bulletin*, 70:1063–1085.
- [Eppley, 1989] Eppley, R. W. (1989). New production: History, methods, problem. In Berger, W. H., Smetacek, V. S., and Wefer, G., editors, *Productivity of the Ocean*, pages 85–97. John Wiley.
- [Eppley and Peterson, 1979] Eppley, R. W. and Peterson, B. J. (1979). Particulate organic matter flux and planktonic new production in the deep ocean. *Nature*, 282:677–680.
- [Estapa et al., 2015] Estapa, M. L., Siegel, D. A., Buesseler, K. O., Stanley, R. H., Lomas, M. W., and Nelson, N. B. (2015). Decoupling of net community and export production on submesoscales in the sargasso sea. *Global Biogeochemical Cycles*, 29:1266–1282.
- [Falkowski et al., 2008] Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive earth’s biogeochemical cycles. *Science*, 320:1034–1039.
- [Falkowski and Ziemann, 1991] Falkowski, P. G. and Ziemann, D. (1991). Role of eddy pumping in enhancing primary production in the ocean. *Nature*, 352:55.
- [Field, 2009] Field, A. (2009). *Discovering statistics using SPSS Sage publications*. Thousand Oaks.
- [Flannery et al., 1992] Flannery, B. P., Press, W. H., Teukolsky, S. A., and Vetterling, W. (1992). *Numerical Recipes in C*. Press Syndicate of the University of Cambridge.

- [Forney and Rothman, 2012] Forney, D. C. and Rothman, D. H. (2012). Common structure in the heterogeneity of plant-matter decay. *Journal of the Royal Society Interface*, 9:0122.
- [Fowler and Knauer, 1986] Fowler, S. W. and Knauer, G. A. (1986). Role of large particles in the transport of elements and organic compounds through the oceanic water column. *Progress in Oceanography*, 16:147–194.
- [Friar et al., 2016] Friar, J. L., Goldman, T., and Pérez-Mercader, J. (2016). Ubiquity of benford’s law and emergence of the reciprocal distribution. *Physics Letters A*, 380:1895–1899.
- [Friedrich and van der Loeff, 2002] Friedrich, J. and van der Loeff, M. M. R. (2002). A two-tracer ($^{210}\text{Po}/^{234}\text{Th}$) approach to distinguish organic carbon and biogenic silica export flux in the antarctic circumpolar current. *Deep Sea Research Part I*, 49:101–120.
- [Gloege et al., 2016] Gloege, L., McKinley, G. A., Mouw, C. B., and Ciochetto, A. B. (2016). Global evaluation of particulate organic carbon flux parameterizations and implications for atmospheric $p\text{CO}_2$. *Global Biogeochemical Cycles*, 31:1192–1215.
- [Goodwin et al., 2008] Goodwin, P., Follows, M. J., and Williams, R. G. (2008). Analytical relationships between atmospheric carbon dioxide, carbon emissions, and ocean processes. *Global Biogeochemical Cycles*, 22:GB3030.
- [Guidi et al., 2009] Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., Picheral, M., and Gorsky, G. (2009). Effects of phytoplankton community on production, size and export of large aggregates: A world-ocean analysis. *Limnology & Oceanography*, 54:1951–1963.
- [Hamming, 1970] Hamming, R. W. (1970). On the distribuion of numbers. *Bell System Technical Journal*, 49:1609–1625.
- [Hargrave et al., 1994] Hargrave, B. T., von Bodungen, B., Stoffyn-Egli, P., and Mudie, P. J. (1994). Seasonal variability in particle sedimentation under permanent ice cover in the arctic ocean. *Continental Shelf Research*, 14:279–293.
- [Hedges et al., 2001] Hedges, J. I., Baldock, J. A., Gélinas, Y., Lee, C., Peterson, M., and Wakeham, S. G. (2001). Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature*, 409:801.
- [Henson et al., 2011] Henson, S. A., Sanders, R., Madsen, E., Morris, P. J., Moigne, F. L., and Quartly, G. D. (2011). A reduced estimate of the strength of the ocean’s biological carbon pump. *Geophysical Research Letters*, 38:L04606.
- [Henson et al., 2010] Henson, S. A., Sarmiento, J. L., Dunne, J. P., Bopp, L., Lima, I., Doney, S. C., John, J., and Beaulieu, C. (2010). Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences*, 7:621–640.
- [Hill, 1985] Hill, P. D. (1985). Kernel estimation of a distribution function. *Communications in Statistics - Theory and Methods*, 14:605–620.
- [Hill, 1995] Hill, T. P. (1995). A statistical derivation of the significant-digit law. *Statistical science*, 10:354–363.

- [Honda,] Honda, M. C. Jamstec sediment trap data at time-series stations: K2 and S1 for ocean SITES. JAMSTEC Environmental Biogeochemical Cycle Research Program, available at <http://ebcrpa.jamstec.go.jp/rigc/e/ebcrp/mbcrt/st_k2s1_oceansites/>, accessed on 10 April 2014.
- [Honda et al., 2002] Honda, M. C., Imai, K., Norjiri, Y., Hoshi, F., Sugawara, T., and Kusakabe, M. (2002). The biological pump in the northwestern north pacific based on fluxes and major components of particulate matter obtained by sediment-trap experiments (1997-2000). *Deep Sea Research Part II*, 49:5595–5625.
- [Honjo et al., 2008] Honjo, S., Manganini, S. J., Krishfield, R. A., and Francois, R. (2008). Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983. *Progress in Oceanography*, 76:217–285.
- [HOT:,] HOT:. Hawaii ocean time-series HOT DOGS database. available at <<http://hahana.soest.hawaii.edu/hot/hot-dogs/interface.html>>, accessed on 1 August 2016.
- [Houghton, 2003] Houghton, R. A. (2003). The contemporary carbon cycle. *Treatises on geochemistry*, 8:473–513.
- [Huntley and Lopez, 1992] Huntley, M. E. and Lopez, M. D. (1992). Temperature-dependent production of marine copepods: A global synthesis. *American Naturalist*, 140:201–242.
- [IPCC, 2014] IPCC (2014). Climate change 2014: Synthesis report. In Team, C. W., Pachauri, R. K., and Meyer, L. A., editors, *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.
- [Ishii et al., 2005] Ishii, M., Shouji, A., Sugimoto, S., and Matsumoto, T. (2005). Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the kobe collection. *International Journal of Climatology*, 25:865–879.
- [Ito and Follows, 2005] Ito, T. and Follows, M. J. (2005). Preformed phosphate, soft tissue pump and atmospheric CO₂. *Journal of Marine Research*, 63:813–839.
- [Iversen and Ploug, 2010] Iversen, M. and Ploug, H. (2010). Ballast minerals and the sinking carbon flux in the ocean: carbon-specific respiration rates and sinking velocity of marine snow aggregates. *Biogeosciences*, 7:2613–2624.
- [Jackson and Burd, 1998] Jackson, G. A. and Burd, A. B. (1998). Aggregation in the marine environment. *Environmental Science and Technology*, 32:2805–2814.
- [Jacquet et al., 2011] Jacquet, S. H. M., Lam, P. J., Trull, T., and Dehairs, F. (2011). Carbon export production in the subantarctic zone and polar front zone south of tasmania. *Deep Sea Research Part II*, 58:2277–2292.
- [Jiao and Others, 2010] Jiao, N. and Others (2010). Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nature Reviews Microbiology*, 8:593.

- [Kajihara, 1971] Kajihara, M. (1971). Settling velocity and porosity of large suspended particles. *Journal of the Oceanographic Society of Japan*, 27:158–162.
- [Karl and Lukas, 1996] Karl, D. M. and Lukas, R. (1996). The hawaii ocean time-series (hot) program: Background, rationale and field implementation. *Deep Sea Research Part II*, 43:129–156.
- [Kawakami and Honda, 2007] Kawakami, H. and Honda, M. C. (2007). Time-series observation of poc fluxes estimated from 234th in the northwestern north pacific. *Deep Sea Research Part I*, 54:1070–1090.
- [Kiørboe, 2000] Kiørboe, T. (2000). Colonization of marine snow aggregates by invertebrate zooplankton: abundance, scaling, and possible role. *Limnology & Oceanography*, 45:479–484.
- [Klaas and Archer, 2002] Klaas, C. and Archer, D. E. (2002). Association of sinking organic matter with various types of mineral ballast in the deep sea: Implications for the rain ratio. *Global Biogeochemical Cycles*, 16:63.
- [Koch, 2005] Koch, A. L. (2005). The logarithm in biology 1. mechanisms generating the log-normal distribution exactly. *Journal of Theoretical Biology*, 12:276–290.
- [Kwon et al., 2009] Kwon, E., Primeau, F., and Sarmiento, J. (2009). The impact of remineralization depth on the air-sea carbon balance. *Nature Geoscience*, 2:630–635.
- [Lalande et al., 2007] Lalande, C., Lepore, K., Cooper, L. W., Grebmeier, J. M., and Moran, S. B. (2007). Export fluxes of particulate organic carbon in the chukchi sea: A comparative study using 234Th/238U disequilibria and drifting sediment traps. *Marine Chemistry*, 103:185–196.
- [Lalande et al., 2008] Lalande, C., Moran, S. B., Wassman, P., Grebmeier, J. M., and Cooper, L. W. (2008). 234 Th-derived particulate organic carbon fluxes in the northern barents sea with comparison to drifting sediment trap fluxes. *Journal of Marine Systems*, 73:103–113.
- [Lamborg and Buesseler, 2010] Lamborg, C. and Buesseler, K. (2010). VERTIGO project sediment trap fluxes of mass, elements and phytoplankton pigments data from KM0414 and RR_K2 cruises. biological and chemical oceanography data management office (bcodmo) [16 january 2016].
- [Lamborg and Others, 2008] Lamborg, C. H. and Others (2008). The flux of bio- and lithogenic material associated with sinking particles in the mesopelagic twilight zone of the northwest and north central pacific ocean. *Deep Sea Research Part II*, 55:1540–1563.
- [Laws, 1975] Laws, E. (1975). The importance of respiration losses in controlling the size distribution of marine phytoplankton. *Ecology*, 56:419–426.
- [Laws et al., 2000a] Laws, E. A., D’Sa, E., and Naik, P. (2000a). Simple equations to estimate ratios of new or export production to total production from satellite-derived estimates of sea surface temperature and primary production. *Limnology & Oceanography: Methods*, 9:593–601.

- [Laws et al., 2000b] Laws, E. A., Falkowski, P. G., Smith, W. O., Ducklow, H., and McCarthy, J. J. (2000b). Temperature effects on export production in the open ocean. *Global Biogeochemical Cycles*, 14:1231–1246.
- [Lee, 2011] Lee, C. (2011). Sediment trap mass flux, MedFlux. Biological and chemical oceanography data system, available at <<http://www.bco-dmo.org/dataset/35619>>, accessed on 22 May 2013.
- [Lee and Others, 1998] Lee, C. and Others (1998). Particulate organic carbon fluxes: compilation of results from the 1995 US JGOFS arabian sea process study: By the arabian sea carbon flux group. *Deep Sea Research Part II*, 45:2489–2501.
- [Limpert et al., 2001] Limpert, E., Stahel, W. A., and Abbt, M. (2001). Log-normal distributions across the sciences. *BioScience*, 51:341–352.
- [Lindeman, 1942] Lindeman, R. L. (1942). The trophic-dynamic aspect of ecology. *Ecology*, 23:399–417.
- [Longhurst, 1998] Longhurst, A. R. (1998). *Ecological geography of the sea*. Academic Press.
- [López-Urrutia et al., 2006] López-Urrutia, A., Martin, E. S., Harris, R. P., and Irigoien, X. (2006). Scaling the metabolic balance of the oceans. *Proceedings of the National Academy of the Sciences*, 103:8739–8744.
- [Luo and Others, 2012] Luo, Y. W. and Others (2012). Database of diazotrophs in global ocean: Abundance, biomass and nitrogen fixation rates. *Earth System Science Data*, 4:47–73.
- [Lutz et al., 2007] Lutz, M., Caldeira, K., Dunbar, R., and Behrenfeld, M. J. (2007). Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. *Journal of Geophysical Research*, 112:C10011.
- [Lutz et al., 2002] Lutz, M., Dunbar, R., and Caldeira, K. (2002). Regional variability in the vertical flux of particulate organic carbon in the ocean interior. *Global Biogeochemical Cycles*, 16:GB1383.
- [MacIntyre et al., 1995] MacIntyre, S., Alldredge, A., and Gotschalk, C. (1995). Accumulation of marine snow at density discontinuities in the water column. *Limnology & Oceanography*, 40:449–468.
- [Maiti et al., 2013] Maiti, K., Charette, M. A., Buesseler, K. O., and Kahru, M. (2013). An inverse relationship between production and export efficiency in the southern ocean. *Geophysical Research Letters*, 40:1557–1561.
- [Marinov et al., 2008] Marinov, I., Follows, M. J., Gnanadesikan, A., Sarmiento, J. L., and Slater, R. D. (2008). How does ocean biology affect atmospheric CO₂? theory and models. *Journal of Geophysical Research: Oceans*, 113:C7.
- [Marra, 1978] Marra, J. (1978). Phytoplankton photosynthetic response to vertical movement in a mixed layer. *Marine Biology*, 46:203–208.

- [Marra, 2009] Marra, J. (2009). Net and gross productivity: Weighing in with ^{14}C . *Aquatic Microbial Ecology*, 56:123–131.
- [Marsay et al., 2015] Marsay, C. M., Sanders, R. J., Henson, S. A., Pabortsava, K., Achterbert, E. P., and Lampitt, R. S. (2015). Attenuation of sinking particulate organic carbon flux through the mesopelagic ocean. *Proceedings of the National Academy of Sciences*, 112:1089–1094.
- [Martin et al., 2011] Martin, A. P., Lampitt, R. S., Perry, M. J., Sanders, R., Lee, C., and D’Asaro, E. (2011). Export and mesopelagic particle flux during a north atlantic spring diatom bloom. *Deep Sea Research Part I*, 58:339–349.
- [Martin et al., 1987a] Martin, J. H., Fitzwater, S. E., Gordon, R. M., Hunter, C. N., and Tanner, S. J. (1987a). Iron, primary production and carbon-nitrogen flux studies during the JGOFS north atlantic bloom experiment. *Deep Sea Research Part II*, 40:115–134.
- [Martin et al., 1987b] Martin, J. H., Knauer, G. A., Karl, D. M., and Broenkow, W. W. (1987b). VERTEX: carbon cycling in the northeast pacific. *Deep Sea Research Part A*, 34:267–285.
- [Mayor et al., 2014] Mayor, D. J., Sanders, R., Giering, S. L., and Anderson, T. R. (2014). Microbial gardening in the ocean’s twilight zone: detritivorous metazoans benefit from fragmenting, rather than ingesting, sinking detritus: fragmentation of refractory detritus by zooplankton beneath the euphotic zone stimulates the harvestable production of labile and nutritious microbial biomass. *BioEssays*, 36:1132–1137.
- [McDonnell et al., 2015] McDonnell, A. M. P., Boyd, P. W., and Buesseler, K. O. (2015). Effects of sinking velocities and microbial respiration rates on the attenuation of particulate carbon fluxes through the mesopelagic zone. *Global Biogeochemical Cycles*, 29:175–193.
- [Michaels and Knap, 1996] Michaels, A. F. and Knap, A. H. (1996). Overview of the US JGOFS bermuda atlantic time-series study and the hydrostation S program. *Deep Sea Research Part II*, 43:157–198.
- [Michaels and Silver, 1988] Michaels, A. F. and Silver, M. W. (1988). Primary production, sinking fluxes and the microbial food web. *Deep Sea Research Part A*, 35:473–490.
- [Middelburg, 1989] Middelburg, J. J. (1989). A simple rate model for organic matter decomposition in marine sediments. *Geochimica et Cosmochimica Acta*, 53:1577–1581.
- [Moigne et al., 2013] Moigne, F. A. L., Henson, S. A., Sanders, R. J., and Madsen, E. (2013). Global database of surface ocean particulate organic carbon export fluxes diagnosed from the ^{234}Th technique. *Earth System Science Data*, 5:295–304.
- [Moigne and Others, 2012] Moigne, F. A. L. and Others (2012). On the proportion of ballast versus non-ballast associated carbon export in the surface ocean. *Geophysical Research Letters*, 39:L15610.
- [Montes and Others, 2012] Montes, E. and Others (2012). Vertical fluxes of particulate biogenic material through the euphotic and twilight zones in the cariacó basin, venezuela. *Deep Sea Research I*, 67:73–84.

- [Montroll and Shlesinger, 1982] Montroll, E. W. and Shlesinger, M. F. (1982). On $1/f$ noise and other distributions with long tails. *Proceedings of the National Academy of Sciences*, 79:3380–3383.
- [Mooy et al., 2002] Mooy, B. A. V., Keil, R. G., and Devol, A. H. (2002). Impact of suboxia on sinking particulate organic carbon: Enhanced carbon flux and preferential degradation of amino acids via denitrification. *Geochimica et Cosmochimica Acta*, 66:457–465.
- [Moran et al., 2012] Moran, S. B., Lomas, M. W., Kelly, R. P., Gradinger, R., Iken, K., and Mathis, J. T. (2012). Seasonal succession of net primary productivity, particulate organic carbon export, and autotrophic community composition in the eastern bering sea. *Deep Sea Research II*, 65:84–97.
- [Moran et al., 2002] Moran, S. B., Shen, C. C., Edmonds, H. N., Weinstein, S. E., Smith, J. N., and Edwards, R. L. (2002). Dissolved and particulate ^{231}Pa and ^{230}Th in the atlantic ocean: Constraints on intermediate/deep water age, boundary scavenging, and $^{231}\text{Pa}/^{230}\text{Th}$ fractionation. *Earth and Planetary Science Letters*, 203:999–1014.
- [Moran et al., 2003] Moran, S. B., Weinstein, S. E., Edmonds, H. N., Smith, J. N., Kelly, R. P., Pilson, M. E. Q., and Harrison, W. G. (2003). Does $^{234}\text{Th}/^{238}\text{U}$ disequilibrium provide an accurate record of the export flux of particulate organic carbon from the upper ocean? *Limnology & Oceanography*, 48:1018–1029.
- [Morel et al., 2007] Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., and Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (case 1) waters in the perspective of a multi-sensor approach. *Remote Sensing of Environment*, 111:69–88.
- [Morris et al., 2007] Morris, P. J., Sanders, R., Turnewitsch, R., and Thomalla, S. J. (2007). ^{234}Th -derived particulate organic carbon export from an island-induced phytoplankton bloom in the southern ocean. *Deep Sea Research Part II*, 54:2208–2232.
- [Mouw et al., 2016] Mouw, C. B., Barnett, A., McKinley, G. A., Gloege, L., and Pilcher, D. (2016). Global ocean particulate organic carbon flux merged with satellite parameters. *Earth System Science Data*, 8:531–541.
- [Murray et al., 1989] Murray, J. W., Downs, J. N., Strom, S., Wei, C. L., and Jannasch, H. W. (1989). Nutrient assimilation, export production and ^{234}Th scavenging in the eastern equatorial pacific. *Deep Sea Research Part A*, 36:1471–1489.
- [Murray et al., 1996] Murray, J. W., Young, J., Newton, J., Dunne, J., Chapin, T., Paul, B., and McCarthy, J. J. (1996). Export flux of particulate organic carbon from the central equatorial pacific determined using a combined drifting trap- ^{234}Th approach. *Deep Sea Research Part II*, 43:1095–1132.
- [Newton and Murray,] Newton, J. and Murray, J. W. Jamstec sediment trap data at time-series stations: K2 and s1 for ocean sites. Biological and chemical oceanography data system, available at <<http://www.bco-dmo.org/dataset-deployment/450839>>, accessed on 18 September 2013.
- [Omand et al., 2015] Omand, M. M., D’Asaro, E. A., Lee, C. M., Perry, M. J., Briggs, N., Cetinic, I., and Mahadevan, A. (2015). Eddy-driven subduction exports particulate organic carbon from the spring bloom. *Science*, 348:222–225.

- [Osmond, 1989] Osmond, C. B. (1989). Photosynthesis. In Briggs, W. R., editor, *Photosynthesis*, pages 5–17. Liss, New York.
- [Palevsky and Doney, 2018] Palevsky, H. I. and Doney, S. C. (2018). How choice of depth horizon influences the estimated spatial patterns and global magnitude of ocean carbon export flux. *Geophysical Research Letters*, 45:4171–4179.
- [Pei and Laws, 2013] Pei, S. and Laws, E. A. (2013). Does the ^{14}C method estimate net photosynthesis? implications from batch and continuous culture studies of marine phytoplankton. *Deep Sea Research Part I*, 82:1–9.
- [Peterson, 1980] Peterson, B. J. (1980). Aquatic primary productivity and the ^{14}C - CO_2 method: A history of the productivity problem. *Annual Review of Ecology and Systematics*, 11:359–385.
- [Peterson et al., 2005] Peterson, M. L., Wakeham, S., Lee, C., Askea, M. A., and Miquel, J. C. (2005). Novel techniques for collection of sinking particles in the ocean and determining their settling rates. *Limnology & Oceanography: Methods*, 3:520–532.
- [Pitman, 1999] Pitman, J. (1999). *Probability*. Springer New York.
- [Planchon et al., 2013] Planchon, F., Cavagna, A. J., Cardinal, D., André, L., and Dehairs, F. (2013). Late summer particulate organic carbon export and twilight zone remineralisation in the atlantic sector of the southern ocean. *Biogeosciences*, 10:803–820.
- [Primeau, 2006] Primeau, F. (2006). On the variability of the exponent in the power law depth dependence of poc flux estimated from sediment traps. *Deep Sea Research Part I*, 53:1335–1343.
- [Queré and Others, 2007] Queré, C. L. and Others (2007). Saturation of the southern ocean CO_2 sink due to recent climate change. *Science*, 316:1735–1738.
- [Reynolds et al., 2007] Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climatology*, 20:5473–5496.
- [Rose and Caron, 2007] Rose, J. M. and Caron, D. A. (2007). Does low temperature constrain the growth rates of heterotrophic protists? evidence and implications for algal blooms in cold waters. *Limnology & Oceanography*, 52:886–895.
- [Rothman and Forney, 2007] Rothman, D. H. and Forney, D. C. (2007). Physical model for the decay and preservation of marine organic carbon. *Science*, 316:1325–1328.
- [Sanders et al., 2010] Sanders, R., Morris, P. J., Poulton, A. J., Stinchcombe, M. C., Charalampopoulou, A., Lucas, M. I., and Thomalla, S. J. (2010). Does a ballast effect occur in the surface ocean? *Geophysical Research Letters*, 37:L08602.
- [Santschi et al., 1999] Santschi, P. H., Guo, L., Walsh, I. D., Quigley, M. S., and Baskaran, M. (1999). Boundary exchange and scavenging of radionuclides in continental margin waters of the middle atlantic bight: Implications for organic carbon fluxes. *Continental Shelf Research*, 19:609–636.

- [Sarmiento, 2013] Sarmiento, J. L. (2013). *Ocean biogeochemical dynamics*. Princeton University Press.
- [Sarmiento and Toggweiler, 1984] Sarmiento, J. L. and Toggweiler, J. R. (1984). A new model for the role of the oceans in determining atmospheric $p\text{CO}_2$. *Nature*, 308:621–624.
- [Savoye et al., 2004] Savoye, N., Buesseler, K. O., Carinal, D., and Dehairs, F. (2004). ^{234}Th deficit and excess in the southern ocean during spring 2001: Particle export and remineralization. *Geophysical Research Letters*, 31:L12301.
- [Schlesinger, 1991] Schlesinger, W. H. (1991). *Biogeochemistry: An Analysis of Global Change*. Academic Press.
- [Schneider et al., 2003] Schneider, B., Schlitzer, R., Fischer, G., and Nöthig, E. M. (2003). Depth-dependent elemental compositions of particulate organic matter (pom) in the ocean. *Global Biogeochemical Cycles*, 17:GB1871.
- [Shimmield et al., 1995] Shimmield, G. B., Ritchie, G. D., and Fileman, T. W. (1995). The impact of marginal ice zone processes on the distribution of ^{210}Pb , ^{210}Po and ^{234}Th and implications for new production in the bellingshausen sea, antarctica. *Deep Sea Research Part II*, 42:1313–1335.
- [Shockley, 1957] Shockley, W. (1957). On the statistics of individual variations of productivity in research laboratories. *Proceedings of the IRE*, 45:279–290.
- [Siegel et al., 2014] Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., and Boyd, P. W. (2014). Global assessment of ocean carbon export by combining satellite observations and food-web models. *Global Biogeochemical Cycles*, 28:181–196.
- [Siegel and Deuser, 1997] Siegel, D. A. and Deuser, W. G. (1997). Trajectories of sinking particles in the sargasso sea: Modeling of statistical funnels above deep-ocean sediment traps. *Deep Sea Research Part I*, 44:1519–1541.
- [Siegel and Others, 2013] Siegel, D. A. and Others (2013). Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sensing of Environment*, 135:77–91.
- [Siegel and Others, 2016] Siegel, D. A. and Others (2016). Prediction of the export and fate of global ocean net primary production: The exports science plan. *Frontiers in Marine Science*, 3:22.
- [Siegenthaler and Wenk, 1984] Siegenthaler, U. and Wenk, T. (1984). Rapid atmospheric CO_2 variations and ocean circulation. *Nature*, 308:624–626.
- [Sigman and Boyle, 2000] Sigman, D. M. and Boyle, E. A. (2000). Glacial/interglacial variations in atmospheric carbon dioxide. *Nature*, 407:859.
- [Silsbe et al., 2016] Silsbe, G. M., Behrenfeld, M. J., Halsey, K. H., Milligan, A. J., and Westberry, T. K. (2016). The cafe model: A net production model for global ocean phytoplankton. *Global Biogeochemical Cycles*, 30:1756–1777.
- [Smayda, 1970] Smayda, T. J. (1970). Normal and accelerated sinking of phytoplankton in the sea. *Oceanography and Marine Biology*, 8:353–414.

- [Smith and Morowitz, 2016] Smith, E. and Morowitz, H. J. (2016). *The Origin and Nature of Life on Earth: The Emergence of the Fourth Geosphere*. Cambridge University Press.
- [Stanely et al., 2004] Stanely, R. H. R., Buesseler, K. O., Manganini, S. J., Steinberg, D. K., and Valdes, J. R. (2004). A comparison of major and minor elemental fluxes collected using neutrally buoyant and surface-tethered traps. *Deep Sea Research Part I*, 51:1387–1395.
- [Steinberg, 1995] Steinberg, D. K. (1995). Diet of copepods (*scopalatum vorax*) associated with mesopelagic detritus (giant larvacean houses) in monterey bay, california. *Marine Biology*, 122:571–584.
- [Steinberg et al., 2008] Steinberg, D. K., Mooy, B. A. V., Buesseler, K. O., Boyd, P. W., Kobari, T., and Karl, D. M. (2008). Bacterial vs. zooplankton control of sinking particle flux in the ocean’s twilight zone. *Limnology & Oceanography*, 53:1327–1338.
- [Stephens, 1974] Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69:730–737.
- [Stewart and Others, 2007] Stewart, G. and Others (2007). Comparing POC export from $^{234}\text{Th}/^{238}\text{U}$ and $^{210}\text{Po}/^{210}\text{Pb}$ disequilibria with estimates from sediment traps in the northwest mediterranean. *Deep Sea Research Part I*, 54:1549–1570.
- [Suess, 1980] Suess, E. (1980). Particulate organic carbon flux in the oceans – surface productivity and oxygen utilization. *Nature*, 288:260–263.
- [Sverdrup, 1953] Sverdrup, H. (1953). On vernal blooming of phytoplankton. *Journal du Conseil International pour l’Éxploration de la Mer*, 18:287–295.
- [Thomalla et al., 2008] Thomalla, S. J., Poulton, A. J., Sanders, R., Turnewitsch, R., Holligan, P. M., and Lucas, M. I. (2008). Variable export fluxes and efficiencies for calcite, opal, and organic carbon in the atlantic ocean: A ballast effect in action? *Global Biogeochemical Cycles*, 22:GB2982.
- [Thurnell, 2016] Thurnell, R. C. (2016). Sediment trap data, CARIACO ocean time series data. available at <<http://imars.marine.usf.edu/CAR/index.html>>, accessed on 30 July 2016.
- [Timoth et al., 2013] Timoth, D. A., Wong, C. S., Barwell-Clarke, J. E., Paige, J. S., White, L. A., and Macdonald, R. W. (2013). Climatology of sediment flux and composition in the subarctic northeast pacific ocean with biogeochemical implications. *Progress in Oceanography*, 116:95–129.
- [Torres-Valdés et al., 2013] Torres-Valdés, S., Painter, S. C., Martin, A. P., Sanders, R., and Felden, J. (2013). Fluxes of sedimenting material from sediment traps in the atlantic ocean. *Earth System Science Data Discussions*, 6:123–145.
- [Torres-Valdés et al., 2014] Torres-Valdés, S., Painter, S. C., Martin, A. P., Sanders, R., and Felden, J. (2014). Data compilation of fluxes of sedimenting material from sediment traps in the atlantic ocean. *Earth System Science Data*, 6:123–145.

- [Trull et al., 2008] Trull, T. W., Bray, S. G., Buesseler, K. O., Lamborg, C. H., Manganini, S., Moy, C., and Valdes, J. (2008). In situ measurement of mesopelagic particle sinking rates and the control of carbon transfer to the ocean interior during the vertical flux in the global ocean (vertigo) voyages in the north pacific. *Deep Sea Research Part II*, 55:1684–1695.
- [Turner, 2015] Turner, J. T. (2015). Zooplankton fecal pellets, marine snow, phytodetritus and the ocean’s biological pump. *Progress in Oceanography*, 130:205–248.
- [Twardowski et al., 2004] Twardowski, M. S., Boss, E., Sullivan, J. M., and Donaghay, P. L. (2004). Modeling the spectral shape of absorption by chromophoric dissolved organic matter. *Marine Chemistry*, 89:69–88.
- [van der Loeff et al., 2011] van der Loeff, M. M. R., Cai, P. H., Stimac, I., Bracher, A., Middag, R., Klunder, M. B., and van Heuven, S. M. (2011). ²³⁴Th in surface waters: Distribution of particle export flux across the antarctic circumpolar current and in the weddell sea during the geotraces expedition zero and drake. *Deep Sea Research Part II*, 58:2749–2766.
- [van der Loeff et al., 1997] van der Loeff, M. M. R., Friedrich, J., and Bathmann, U. (1997). Carbon export during the spring bloom at the antarctic polar front, determined with the natural tracer ²³⁴Th. *Deep Sea Research Part II*, 44:457–478.
- [Villa-Alfageme and Others, 2016] Villa-Alfageme, M. and Others (2016). Geographical, seasonal, and depth variation in sinking particle speeds in the north atlantic. *Geophysical Research Letters*, 43:8609–8616.
- [Volk and Hoffert, 1985] Volk, T. and Hoffert, M. I. (1985). Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO₂ changes. In Sundquist, E. T. and Broecker, W. S., editors, *The carbon cycle and atmospheric CO₂: Natural variations archean to present*, pages 99–110. American Geophysical Union.
- [Wainer et al., 2006] Wainer, H., Gessaroli, M., and Verdi, M. (2006). Visual revelations: Finding what is not there through the unfortunate binning of results: The mendel effect. *Chance*, 19:49–52.
- [Westberry et al., 2008] Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles*, 22:GB2024.
- [Wickham, 2014] Wickham, H. (2014). Tidy data. *Journal of Statistical Software*, 59:1–23.
- [Wiener, 1949] Wiener, N. (1949). *Extrapolation, interpolation, and smoothing of stationary time series, vol. 7*. MIT Press.
- [Williams and Follows, 2011] Williams, R. G. and Follows, M. J. (2011). *Ocean dynamics and the carbon cycle*. Cambridge University Press.
- [York, 1966] York, D. (1966). Least-squares fitting of a straight line. *Canadian Journal of Physics*, 44:1079–1086.

[Zhou et al., 2012] Zhou, K., Noddler, S., Minhan, D., and Hall, J. A. (2012). Insignificant enhancement of export flux in the highly productive subtropical front, east of new zealand: A high resolution study of particle export fluxes based on ^{234}Th : ^{238}U disequilibria. *Biogeosciences*, 9:973–992.