Department of Physics

University of Strathclyde

A Mie-based flow cytometric size and real refractive index determination method for natural marine particle populations

Jacopo Agagliate

2017

A thesis submitted in accordance with the requirements for the degree

Doctor of Philosophy

Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

Acknowledgements

First of all, I wish to express my most sincere gratitude to my supervisor David McKee, for his energy and enthusiasm and for all the patience, time and dedication that he offered both to me and to my research. Little of this study could have been accomplished without his precious advice and assistance.

I then wish to thank my colleagues at the University of Strathclyde: Derek, for all the random chats we've had in the lab about work, science and life; and Ina, for her kind friendship, help and support, and for giving me a roof for a while, back when I was still looking for one. My time at the University would have been much, much more drab without you. I also wish to give my thanks to all the many other fellow scientists that both here and abroad, whether through data collection, practical help or counsel have made this doctoral work possible.

I would further like to express my gratitude to the Marine Alliance for Science and Technology for Scotland (MASTS), not only for funding my research, but also for creating a wonderful environment for young scientists to learn, discuss and share their work in.

To all my friends, near and far, present or past: we may talk with each other every day or once a month, or we may just end up never really talking again; gods know I’m awful at maintaining social ties, and I'm truly sorry for that. But whoever you are, know that I’m grateful for whatever time we've had the chance to share in this universe.

To my family at large: I feel like I take all you guys for granted most of the time, but the truth is that I’m more grateful than I could ever express and likely more than I can even realise. Thank you for being there for me, always.

And finally, to mum and dad: thank you for your unconditional love, and for always reminding me that it is still okay to dream a little. I swear I don't forget it on purpose!

Nomenclature

ADM: Anomalous diffraction method.

AOP: Apparent optical property.

ChlA: Chlorophyll-A content.

FC: Flow cytometry; or, flow cytometric.

FFPF: Fournier-Forand phase function.

FRRf: Fast repetition rate fluorometry.

FWHM: Full width at half maximum.

HPLC: High-performance liquid chromatography.

IOP: Inherent optical property.

iRI: Imaginary part of the complex refractive index.

ISM: Inorganic suspended matter.

OSM: Organic suspended matter.

PMT: Photomultiplier tube.

POC: Particulate organic carbon.

PRID: Particle (real) refractive index distribution.

PSD: Particle size distribution.

RI: Complex refractive index.

RMS%E: Root-mean-square percentage error.

RMSE: Root-mean-square error.

rRI: Real part of the complex refractive index.

SEM: Scanning electron microscope.

SPF: Scattering phase function.

TSM: Total suspended matter.

VSF: Volume scattering function.

Abstract

Following the path of combining Mie theory and flow cytometry to assign size and refractive index to suspended particles in the steps of Ackleson & Spinrad (1988) and, more recently, Green et al. (2003a, 2003b), a Mie-based flow cytometry (FC) method was developed to retrieve particle size distributions (PSDs) and real refractive index (rRI) information in natural waters. The need for a technique capable of directly assessing both size and real refractive index of the particles was first established by carrying out a sensitivity analysis of the effect a spectrally complex refractive index and log-normal variations to commonly employed PSD models have on the optical behaviour of the particle population. The Mie-based FC method proper was then developed and tested, initially against standard particles of known diameter and rRI and secondly on two datasets, one of algal culture samples (AC dataset) and one of natural seawater samples collected in UK coastal waters (UKCW dataset).

The method retrieved PSDs and real refractive index distributions (PRIDs) for both datasets. FC PSDs were validated against known algal sizes for AC samples and against independent PSDs measured via laser diffractometry for UKCW samples. PRIDs were then combined with FC PSDs and fed into Mie-based forward optical modelling to reconstruct bulk IOPs. These achieved broad agreement with independent IOP measurements, lending further support to the results of the FC method and to the employment of Mie theory within the context of optical modelling of natural particle populations. Furthermore, the unique insight offered by the FC method in terms of PSD and PRID determination allowed for the assessment of the individual contribution of particle subpopulations to the bulk IOPs, both by size (small/large particle fractions) and by particle type (inorganic/organic/fluorescent fractions). Lastly, PSDs and PRIDs were combined with literature-derived models of particle density, cell organic carbon and chlorophyll-A content, in an effort to explore the biogeochemical properties of the particle populations within the UKCW dataset. The models successfully estimated independent measurements of particulate suspended matter and (after an optimisation procedure) of organic carbon and chlorophyll-A content.

Table of Contents

1. Introduction

In this first chapter, an overall look at the background and goals of this thesis will be given, providing an introduction to the optical properties of water, to their relationship with the physical properties of suspended particle populations, and to the principal tools that will be used throughout this work to explore both: flow cytometry and Mie theory. The second half of the chapter will also provide a comprehensive list of relevant marine optics formulae, concepts and definitions, constituting a useful theoretical basis for the chapters to follow.

1.1 Overview

1.1.1 Marine optics, IOPs and flow cytometry

The optical properties of the ocean are the result of many factors. Some of them are internal to the medium, such as temperature, salinity, composition of the dissolved material and suspended particles within it. The rich diversity of this suspended and dissolved content heavily affects light behaviour. Some other factors are instead external, such as intensity, direction and polarization of the incident light field. It is in this separation between internal and external factors that the two traditional groupings of optical properties, *inherent* and *apparent optical properties*, have their origin. Inherent optical properties (*IOPs*) are those that are solely dependent on the physical characteristics of the medium. Apparent optical properties (*AOPs*) are those that also depend on the angular structure of the light field.

Of all IOPs, two are fundamental in the sense that all others can be derived from them: the *spectral absorption coefficient, a,* and the *spectral volume scattering function* (VSF), β. The spectral absorption coefficient a represents the wavelength- or λ-dependent fraction of light absorbed within a unit volume per unit of distance travelled in the medium, where λ is the wavelength. The spectral volume scattering function β represents the λ-dependent polar angular distribution of scattered intensity per unit of incident irradiance per unit volume, i.e. how much light is scattered into each angle, in bulk, by the water, with the angle ranging from zero to π radians (where zero is the direction of the incident light).

Both a and β are fundamentally dependent on the characteristics of the suspended particles and dissolved content of the water. Particles in suspension interact with the incoming light, scattering, absorbing and sometimes re-emitting it as fluorescence; dissolved substances chiefly contribute to total absorption. Together they produce a net effect which contributes to the bulk optical properties of the medium: a simple example is the difference between the brown, opaque, particle-rich waters of rivers and estuaries and the blue, relatively transparent waters of the open sea, where particle content is often scarce.

Establishing a robust relationship between suspended particle composition and bulk IOPs remains an outstanding challenge, and one that has been intrinsic to the field for more than a century. In 1908, Gustav Mie published his famous analytical solution for light scattering from uniform spherical particles, which stemmed from the study of scattering from gold colloid solutions (Mie, 1908). By the 50s and 60s studies were already investigating familiar topics, e.g. particle size distributions in polydisperse systems through light scattering measurements (Chin *et al.*, 1955), single particle scattering from latex beads suspended in water (Ashley & Cobb, 1958), relative contributions of particle scattering and molecular scattering in sea water samples (Morel, 1965, expanding on earlier Jerlov, 1953 and Jerlov, 1963), measurements at small angles of the volume scattering function in various particle suspensions (Bauer & Morel, 1967). Comprehensive textbooks about particle scattering from that era, such as van de Hulst (1957), are still relevant to this day. More recent works continue to explore the nature of the size distribution of suspended particles in the ocean (e.g. Risović, 1993; Jonasz & Fournier, 1996) and its effects on the optical properties of the water (Ulloa *et al.*, 1994; Volten *et al.*, 1998; Peng & Effler, 2007). Yet, despite the long history of the field, many questions are left ananswered and many aspects of suspended particulate physics and dynamics are poorly understood and documented. The most important physical properties of the particle population in the context of marine optics are the *size* and *complex refractive index* of the particles, which are however also difficult to accurately determine over the entire range of optically relevant sizes.

The complex refractive index (RI) is at the core of scattering and absorption. Light is scattered at the interface between mediums with differing real parts of the refractive index (rRI), a basic example being Snell's law of refraction,

$$n_1 \sin \vartheta_i = n_2 \sin \vartheta_t \qquad (1.1)$$

where n_1 and n_2 are the real refractive indices encountered by light as it travels through the interface, ϑ_i is the angle of the incident light relative to the normal of the interface and ϑ_t is the angle of the transmitted light relative to the same normal (Fig. 1.1).

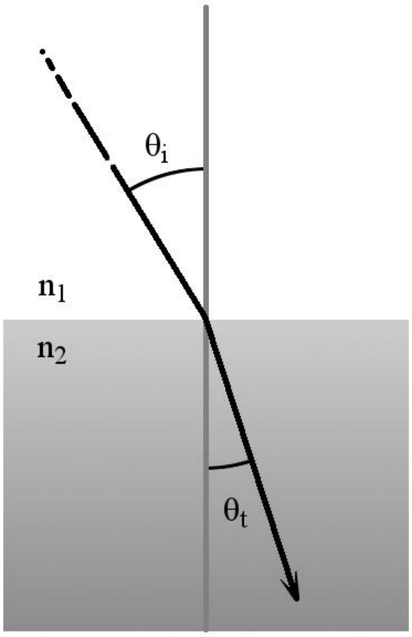


Figure 1.1 – Refraction of light passing through the interface of two materials with real refractive indices n_1 and n_2, as described by Snell's law.

Absorption of light is determined by the imaginary part of the refractive index (iRI) of the medium, a fundamental example being the relationship

$$a = \frac{4\pi n_i}{\lambda}, \qquad (1.2)$$

where n_i is the imaginary part of the refractive index. At the same time, the complex refractive index of the particles is difficult to determine in the case of mixed populations: in most cases experimental results produce bulk values for the whole assemblage of particles, and accordingly most literature values are given for suspensions of a single type of particle or provide an average for a whole population. Twardowski *et al*. (2001) offer a useful compendium of the real part of the refractive indices of typical marine particle population constituents (Tab. 1.1), expanding on Aas (1996) which provides real refractive indices for various phytoplankton types based on their mixture of basic components and adapting a list of relevant inorganic RIs from Lide (1997).

	Material or particle type	n
a	Chlorophyll-rich waters	1.04-1.05
	Sediment-rich waters	1.14-1.18
b	Feldspar (8)	1.14-1.18
	Clays (5)	1.16-1.20
	Mica (5)	1.17-1.23
	Amphibole (5)	1.23-1.25
	Quartz	1.15
	Calcite	1.24
	Amorphous silica (opal)	1.08
c	Cellulose	1.163
	Coccolithophorids	1.036-1.068
	Dinoflagellates	1.035-1.065
	Brown flagellates	1.034-1.065
	Red algae	1.034-1.064
	Blue-green algae	1.033-1.062
	Diatoms	1.033-1.062
	Green algae	1.032-1.061

Table 1.1 – Values of the real part of the complex refractive index for some of the main costituents of marine particle populations, given relative to water. Reference: a) Twardowski *et al*. (2001); b) from Lide (1997) via Twardowski *et al*. (2001). The number of mineral types included within each class is indicated within parentheses; c) from Aas (1996).

Relevant examples of the methods used in literature for real refractive index retrieval include old techniques such as immersion of particles in various media until scattering disappears (McCrone *et al.*, 1967; Bryant *et al.*, 1969); bulk derivation from volume scattering functions and size distributions (Gordon & Brown, 1972; Zaneveld *et al.*, 1974); and bulk derivation from attenuation and absorption efficiencies (Bricaud *et al.*, 1988; Stramski & Kiefer, 1990; Stramski & Morel, 1990). Further bulk methods include the algorithm presented in Twardowski *et al.* (2001), and the anomalous diffraction method (Bricaud & Morel, 1986; Stramski & Mobley, 1997) which has also the advantage of calculating a value for the imaginary part of the RI. Values for the imaginary part of the refractive index are unfortunately quite sparse in literature, e.g. Ahn *et al.* (1992), Babin *et al.* (2003).

Information about the size of the particles present in a suspension is conveyed in what is known as a *particle size distribution* (PSD). PSDs represent the concentration of particles within each size class of the particle population. These size classes are somewhat arbitrary subdivisions based on a definition of "size", which can be variably represented through volume, equivalent area or relevant length of each particle. The choice is most often guided by either the context of the study, the shape of the particles or the nature of available measurements. PSDs can be combined with RIs to feed directly into IOP calculations, using optical models to describe the interaction between light and particles: one of the most widely used optical models is Mie's solution to Maxwell's equations, known as *Mie theory* for short, an analytical solution to scattering under the approximation of spherical, homogeneous spheres. Yet, PSDs too are difficult to determine completely due to the wide range of optically relevant sizes, spanning from tens of nanometres to a few millimetres (Reynolds *et al.*, 2010; Davies *et al.*, 2014). Submicron scales are still poorly investigated, with only a few suitable techniques applicable e.g. SEM (Peng & Effler, 2007). These are however only really suitable for minerogenic particles, as biological samples require extensive preparation, and are therefore not broadly suitable for natural waters.

Most studies use some parametrization of the PSD, of which the most common form in the case of natural waters is the power law or Jungian distribution (Junge, 1963), adopted following studies which found small particles to far outnumber large particles in the sea (Bader, 1970; Sheldon *et al.*, 1972; Ulloa *et al.*, 1994). A power law distribution is an open-ended, logarithmically linear spectrum, with ill-defined average size and average concentration always close to that of the smallest size considered. Other techniques involve the reconstruction of PSDs, either via inversion of the volume scattering function, e.g. using laser diffractometers such as the LISST series of instruments (Agrawal & Pottsmith, 1994; Traykovski *et al.*, 1999; Agrawal & Pottsmith, 2000) which use Mie theory to calculate the scattering kernel relating energy received by the detectors to the number density of particles, or by iterative reconstruction of the measured VSF via modelled contribution of log-normal particle populations (Zhang et al., 2011). Overall, availability and reliability of PSD and RI determination are still largely partial to improvement, despite their fundamental influence on optical properties.

A particular technique in the context of PSD and particle RI studies is *flow cytometry.* It is one of few techniques (another important example being Coulter counters) to investigate particle properties on a particle-by-particle basis rather than in bulk. It was originally developed (and is still largely employed) for biomedical applications (Shapiro, 1988; Dubelaar & Jonker, 2000), but was soon applied to marine biology as well. It was then applied in the field of marine optics for the determination of particulate properties for the first time about three decades ago (Ackleson & Spinrad, 1988; Cunningham & Buonaccorsi, 1992). Flow cytometers analyze the particle population present in a suspension by focusing the fluid sample core within a high-speed sheath flow, pushing the particles in single file in front of a light source (most often a laser). Scattered and fluoresced light is collected over two angle ranges (one close to the direction of the laser, the other orthogonal to it), with the collected light offering information on a variety of particle characteristics such as size, major axis length, shape, fluorescence, internal structure and (indirectly) refractive index, allowing for multi-parameter analysis of the particle population. Advantages of flow cytometry include not only particle-by-particle multivariate analysis, but also high sample throughput and relatively large working ranges (~1-100 μm).

The use of flow cytometry for the determination of size and refractive index on an individual particle basis was pioneered by Ackleson & Spinrad (1988) and subsequently expanded upon (Green *et al*., 2003a; Green *et al*., 2003b). It has most recently been employed for the determination of scattering and backscattering cross-sections (Duforêt-Gaurier *et al*., 2015; Moutier *et al*., 2016).

1.1.2 Rationale and goals of the study

The rationale behind this thesis is to apply flow cytometry for the determination of PSDs and RIs of particle populations, with particular attention to developing the capability to operate on complex natural populations. Specific goals are:

- Development of a method capable of determining size and real refractive index of particles simultaneously and on a particle-by-particle basis.
- Validation of the capabilities of the method and application on a dataset made up of a variety of natural water samples, without any ad hoc adjustment of the method itself.
- Assessment of consistency with optical properties through forward modelling of IOPs via Mie theory and comparison with independently obtained scattering, backscattering and absorption coefficients.
- Assessment of consistency with associated biogeochemical data from water sample analysis.

Flow cytometry is well suited for the task: the technique by its very nature analyses particles individually and can retrieve information on the composition as well as the size of the particulate. Furthermore, Davies *et al*. (2014) show that a large part of the contribution to total scattering and backscattering comes from particles in a size range comprised between 0.1 and 10 µm (Fig. 1.2), and this size range is well covered by flow cytometers.

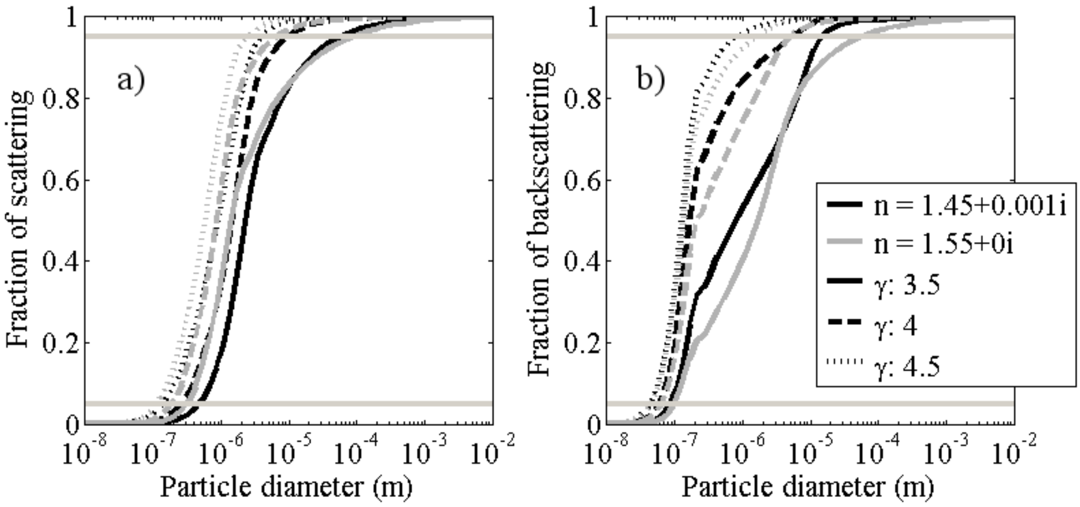


Figure 1.2 – Cumulative (a) scattering and (b) backscattering for theoretical organic (n = 1.45+0.001i) and inorganic (n = 1.55+0i) Junge distributions of particles, where γ is the slope. Adapted from Davies *et al*. (2014). The horizontal lines mark 5% and 95% of contribution to the total scattering.

This is not the entire optically relevant size range, which Davies *et al*. put between 0.05 and 2000 μm. Another goal is then to see if the use of Junge-type extensions in the fashion of Green *et al*. (2003b) is an effective tool in the bid for optical closure, i.e. the successful agreement between optical properties as determined by independent measurement techniques (in this case, IOPs as reconstructed from FC-derived particle properties and as measured *in situ*).

Finally, Mie theory is the optical model of choice in this study. Mie theory is analytical: given a set of input parameters, the output of the model is exact for particles of any size and any refractive index. At the same time, the hypotheses for which Mie theory holds are strict and in most cases represent an approximation of reality. The particles are required to be perfectly spherical and homogeneous; additionally, the incident light field is required to be an infinite plane wave, i.e. the particles must be completely bathed in the incident radiation field whatever their size.

Faceted mineral particles, non-spherical objects like flocs and aggregates, many types of diatoms are all examples of typical components of marine particle populations which violate Mie theory assumptions. A large number of more complex and accurate optical models have been developed:

- *Concentric layered spheres*, an extension of Mie scattering to the case of multiple, concentric spheres (Aden & Kerker, 1951; Fenn & Oser, 1965; Kitchen & Zaneveld, 1992; Fuller, 1993b).
- *Non-concentric layered spheres*, a generalization of the layered spheres model to include non-concentric spheres (Fikioris & Unuzoglu, 1979; Borghese *et al.*, 1992; Fuller, 1993a; Ngo, Christesen & Videen, 1996).
- *Spheroids*, both homogeneous and layered (Asano & Yamamoto, 1975; Barton, 2001; Xu *et al.*, 2007).
- *Infinitely long cylinders* (Sharma *et al.*, 1997), useful to approximate particles with extreme aspect ratios such as long chain-forming diatoms and rod-shaped phytoplankton.
- *Discrete dipole approximation (DDA)*, in which the scatterer (of arbitrary shape) is substituted by a set of point dipoles. The dipoles interact with each other and the incident light field, and a system of linear equations is generated to describe these interactions. Solving the system produces the dipole polarizations, from which scattering is derived (Purcell & Pennypacker, 1973; Draine & Flatau, 1994; Yurkin & Hoekstra, 2007).
- *T-matrix*, a method which revolves around the computation of the transfer matrix relating incident field coefficients and scattered field coefficients in non-spherical particles (Waterman, 1965; Waterman, 1971; Mishchenko *et al.*, 1996).
- *Packed spheres* (*generalized Mie Theory)*, extending Mie theory scattering calculations from the single sphere case to that of multiple packed spheres (Gouesbet & Grehan, 1999; Quinten, 2011).
- *Finite-difference time-domain method (FDTD)*, based on a time-domain solution of the Maxwell equations (Yang & Liou, 1996a).

- *Improved geometric optics method (IGOM)*, based on the implementation of ray tracing (Yang & Liou, 1996b).

Yet, despite the abundance of available models and the strict assumptions of Mie scattering, the fundamental advantage of Mie theory remains its role as a good first approximation in particle populations for which no previous structural or shape information is available. Furthermore, suspended particle populations are most often randomly oriented. In very large numbers (like those present at small scales in natural particle populations) particles for which Mie theory cannot individually apply tend to collectively behave like an ensemble of equivalent spheres (Morel, 1991). Finally, given the magnitude of potential errors in key IOPs previously used to assess the performance of Mie theory and recent advances in IOP corrections (McKee *et al.*, 2013), there is an opportunity to re-evaluate the utility of Mie theory for simulating bulk IOPs for natural particle populations.

The succesful development of a robust method for the determination of particle size and real refractive indices in natural waters would further the case of flow cytometry as an important complement to classic methods like Coulter counting, and, increasingly, laser diffractometry. In particular the establishment of viable means of retrieval of RIs on a particle-by-particle basis would also offer a wealth of new information on the biogeochemical complexity of mixed particulates (normally tackled in bulk) and be a step forward towards bridging the divide between particle optics and particle dynamics.

1.2 Marine optics fundamentals

In this section a compendium of useful definitions and formulae will be provided as reference to the reader.

The formulae in this section have been mostly adapted from Mobley's "Light and Water" (Mobley, 1994), specifically Chapters 1 and 3. The subsection about cross-sections and efficiencies has also been informed by Chapter 2 of van de Hulst's "Light Scattering by Small Particles" (van de Hulst, 1957).

1.2.1 Geometrical radiometry

Before introducing the IOPs proper, it is useful to consider the definitions for a few radiometric quantities: radiance, irradiance and intensity of the incident light.

1.2.1.1 Radiance

The (unpolarized) *spectral radiance* of light is defined as

$$L(\vec{x}, t, \hat{\xi}, \lambda) \equiv \frac{\Delta Q}{\Delta t \Delta A \Delta\Omega \Delta\lambda} \qquad (\mathrm{J\ s^{-1}\ m^{-2}\ sr^{-1}\ nm^{-1}}) \qquad (1.3)$$

where ΔQ is the amount of radiant energy and Δt, ΔA, ΔΩ and Δλ are time, area, solid angle and wavelength intervals respectively. A *spectral radiant power* or *radiant flux* is then also defined as

$$\Phi \equiv \frac{\Delta Q}{\Delta t \Delta\lambda}. \qquad (\mathrm{J\ s^{-1}\ nm^{-1}}) \qquad (1.4)$$

By combining eqs. (1.3) and (1.4) in the limit of infinitesimal intervals of the parameters considered and given the equivalence J s^{-1} = W, the definition of spectral radiance can be then rewritten as

$$L(\vec{x}, t, \hat{\xi}, \lambda) \equiv \frac{\partial^2\Phi}{\partial A\, \partial\Omega}\ . \qquad (\mathrm{W\ m^{-2}\ sr^{-1}\ nm^{-1}}) \qquad (1.5)$$

Spectral radiance ties together the spatial, temporal, angular and wavelength information of light, and is the fundamental radiometric quantity in the sense that the other quantities can be derived by it.

1.2.1.2 Irradiance

The area ΔA is more suitably described as the area projected by the light beam as it crosses the surface, $\Delta S \cos\theta$, where ΔS is the surface crossed by the light beam and θ the (polar) angle of incidence. This is relevant when considering the *spectral irradiance*,

$$E(\vec{x},t,\lambda) \equiv \frac{\partial \Phi}{\partial A}, \qquad (\mathrm{W\ m^{-2}\ nm^{-1}}) \qquad (1.6)$$

which represents radiant flux per unit surface and is obtained as the result of integrating the radiance over a certain solid angle range:

$$E(\vec{x},t,\lambda) = \int L(\vec{x},t,\hat{\xi},\lambda)\cos\theta\, d\Omega$$

$$= \int_{\varphi=0}^{2\pi}\int_{\theta_1}^{\theta_2} L\left(\vec{x},t,\hat{\xi},\lambda\right)\cos\theta\sin\theta\, d\theta d\varphi \qquad (\mathrm{W\ m^{-2}\ nm^{-1}}) \qquad (1.7)$$

where φ is the azimuthal angle. The specific spectral irradiance considered is dependent on the angle integration range. Depending on the choice of θ_1 and θ_2, the *downward spectral planar irradiance*

$$E_d(\vec{x},t,\lambda) = \int_{\varphi=0}^{2\pi}\int_{\theta=0}^{\pi/2} L\left(\vec{x},t,\hat{\xi},\lambda\right)\cos\theta\sin\theta\, d\theta d\varphi \qquad (1.8)$$

and the *upward spectral planar irradiance*

$$E_u(\vec{x},t,\lambda) = \int_{\varphi=0}^{2\pi}\int_{\theta=\pi/2}^{\pi} L\left(\vec{x},t,\hat{\xi},\lambda\right)\cos\theta\sin\theta\, d\theta d\varphi \qquad (1.9)$$

can be variously defined. Finally, in the limit of the area crossed by the light being an infinitesimal point the "effective" area seen by the light is the same for all directions, and ΔA = ΔS, defining corresponding versions of what is known as the *spectral scalar irradiance*

$$E_0(\vec{x},t,\lambda) = \int L(\vec{x},t,\hat{\xi},\lambda)\, d\Omega$$

$$= \int_{\varphi=0}^{2\pi}\int_{\theta 1}^{\theta 2} L\left(\vec{x},t,\hat{\xi},\lambda\right)\sin\theta\, d\theta d\varphi\,. \qquad (\mathrm{W\ m^{-2}\ nm^{-1}}) \qquad (1.10)$$

1.2.1.3 Intensity

Similarly to spectral irradiance $E(\vec{x},t,\lambda)$, the *spectral intensity* $I\left(\vec{x},t,\hat{\xi},\lambda\right)$ is defined as:

$$I(\vec{x},t,\hat{\xi},\lambda) \equiv \frac{\partial \Phi}{\partial \Omega}. \qquad (\mathrm{W\ sr^{-1}\ nm^{-1}}) \qquad (1.11)$$

It is obtained as the result of the integration of the spectral radiance over the area ΔA which "sees" the solid angle ΔΩ, i.e.

$$I(\vec{x},t,\hat{\xi},\lambda) = \int L(\vec{x},t,\hat{\xi},\lambda)\, d\mathrm{A}. \qquad (\mathrm{W\ sr^{-1}\ nm^{-1}}) \qquad (1.12)$$

1.2.2 Inherent optical properties

As described in the first section of this first chapter, the inherent optical properties or IOPs of a water volume are those properties which are only dependent on the inherent physical characteristics of the medium, independently from the characteristics of the light field which will interact with it.

1.2.2.1 Absorptance and scatterance

Consider $\Phi_i(\lambda)$ as the spectral radiant power *incident* on a small volume ΔV of water of thickness Δr and area ΔA. $\Phi_a(\lambda)$, $\Phi_s(\lambda)$ and $\Phi_t(\lambda)$ are respectively the fraction of power *absorbed*, *scattered* and *transmitted* by that volume. Assuming that no inelastic scattering or wavelength dispersion occurs, by conservation of energy

$$\Phi_i(\lambda) \equiv \Phi_a(\lambda) + \Phi_s(\lambda) + \Phi_t(\lambda). \qquad (1.13)$$

The fraction of incident radiant power that is absorbed within the volume is known as *spectral absorptance*

$$A(\lambda) \equiv \frac{\Phi_a(\lambda)}{\Phi_i(\lambda)} \qquad (1.14)$$

and the fraction of incident radiant power that is scattered as it travels through the volume is analogously defined as *spectral scatterance*

$$B(\lambda) \equiv \frac{\Phi_s(\lambda)}{\Phi_i(\lambda)}. \qquad (1.15)$$

Spectral transmittance is similarly defined but will not be used in the following.

1.2.2.2 Attenuation, absorption, scattering coefficients

Marine optics most often deals with absorption and scattering coefficients, which represent absorption and scatterance per unit distance within the medium.

The *spectral absorption coefficient* is defined as

$$a(\lambda) = \lim_{\Delta r \to 0} \frac{A(\lambda)}{\Delta r} \qquad (\mathrm{m}^{-1}) \qquad (1.16)$$

while the *spectral scattering coefficient* is defined as

$$b(\lambda) = \lim_{\Delta r \to 0} \frac{B(\lambda)}{\Delta r}. \qquad (\mathrm{m}^{-1}) \qquad (1.17)$$

Their sum is known as the *spectral attenuation coefficient*

$$c(\lambda) = a(\lambda) + b(\lambda). \qquad (\mathrm{m}^{-1}) \qquad (1.18)$$

1.2.2.3 Volume scattering function and phase function

Consider $B(\psi,\lambda)$ as the angular distribution of the spectral scatterance of (1.15), i.e. the fraction of the incident radiant power scattered through an angle ψ into a solid angle $\Delta\Omega$ centred on ψ. The angle ψ is the scattering angle, with values comprised between 0 and π. Consider then the expression

$$\begin{aligned}\beta(\psi,\lambda) &= \lim_{\Delta r \to 0} \lim_{\Delta\Omega \to 0} \frac{\mathrm{B}(\psi,\lambda)}{\Delta r \Delta\Omega} \\ &= \lim_{\Delta r \to 0} \lim_{\Delta\Omega \to 0} \frac{\Phi_s(\psi,\lambda)}{\Phi_i(\lambda) \Delta r \Delta\Omega}.\end{aligned} \qquad (\mathrm{m}^{-1}\ \mathrm{sr}^{-1}) \qquad (1.19)$$

Since the spectral power scattered into $\Delta\Omega$ can be interpreted as the spectral radiant intensity scattered into direction ψ multiplied by the solid angle, $\Phi_s(\psi,\lambda) =$

$\mathrm{I}_s(\psi,\lambda)\Delta\Omega$; the corresponding incident irradiance is $E_i(\lambda) = \Phi_s(\psi,\lambda)\Delta\mathrm{A}$. Eq. (1.19) can therefore be rewritten as

$$\beta(\psi,\lambda) = \lim_{\Delta\mathrm{V}\to 0} \frac{\mathrm{I}_s(\psi,\lambda)}{E_i(\lambda)\Delta\mathrm{V}}. \qquad (\mathrm{m}^{-1}\ \mathrm{sr}^{-1}) \qquad (1.20)$$

This is known as the *spectral volume scattering function*, or VSF for short, and it represents the angular distribution of the spectral scattered intensity per unit incident irradiance per unit volume. Together with the absorption coefficient *a(λ)* the volume scattering function is all that is required to inform the radiative transfer equations.

Integration the VSF over all directions produces the total spectral scattered intensity per unit incident irradiance per unit volume, i.e the spectral scattering coefficient *b(λ)*:

$$b(\lambda) = 2\pi \int_0^{\pi} \beta(\theta,\lambda) \sin\theta \, d\theta \, . \qquad (\mathrm{m}^{-1}) \qquad (1.21)$$

The form of this last equation is due to the fact that scattering in natural waters is azimuthally symmetric about the incident direction, at least when light is unpolarized and the scatterers are randomly oriented.

Different integration ranges for ψ produce different scattering coefficients, the most relevant of which is the *spectral backscattering coefficient*

$$b_b(\lambda) = 2\pi \int_{\pi/2}^{\pi} \beta(\theta,\lambda) \sin\theta \, d\theta \, . \qquad (\mathrm{m}^{-1}) \qquad (1.22)$$

Scattering and backscattering coefficients can also be combined to produce a *spectral backscattering ratio*, which is the relative magnitude of scattering in the backwards direction and total scattering:

$$\tilde{b}_b(\lambda) = \frac{b_b(\lambda)}{b(\lambda)}. \qquad (1.23)$$

The relationship between VSF and scattering coefficient can also be used to define the *spectral volume scattering phase function* (SPF)

$$\tilde{\beta}(\theta,\lambda) = \frac{\beta(\theta,\lambda)}{b(\lambda)} \, . \qquad (\mathrm{sr}^{-1}) \qquad (1.24)$$

This separates the VSF into two components, one containing the strength of the scattering, $b(\lambda)$, and one containing information on the angular distribution of the scattering, $\tilde{\beta}(\theta,\lambda)$. Finally, the combination of (1.21) and (1.24) also provides the normalization condition for the scattering phase function :

$$2\pi \int_0^{\pi} \tilde{\beta}(\theta,\lambda) \sin\theta \, d\theta = 1 \, . \qquad (1.25)$$

It is also to be noted that IOPs are intrinsically additive, i.e. their values are the result of the addition of the single contributions from each of the constituents in the medium. The absorption coefficient of a volume of water can be written as

$$a(\lambda) = a_w(\lambda) + \sum_{i=1}^{n} a_i(\lambda) \, , \qquad (1.26)$$

where $a_w(\lambda)$ is the spectral absorption coefficient of pure water and n the number of constituents. Similarly, the expression for the VSF can be written as

$$\beta(\theta,\lambda) = \beta_w(\theta,\lambda) + \sum_{i=1}^{n} \beta_i(\theta,\lambda) \, . \qquad (1.27)$$

This constitutes the basis for the idea of a single-particle VSF, a concept which is useful in the context of forward modelling of IOPs from particulate size and RI data.

1.2.3 Efficiency factors and optical cross sections

Consider σ_b as the cross-sectional area required to remove from the incident light beam as much light as the total scattering from a certain particle. σ_b is then known as the *scattering cross-section* of that particle. Analogously, σ_a can be defined as the *absorption cross-section* of that particle and, in conformity with eq. (16), σ_c as the *attenuation cross-section* of the particle, i.e.

$$\sigma_c = \sigma_a + \sigma_b \,. \qquad (1.28)$$

By the same definition, spectral coefficients are the result of the total contribution of all cross-sections of the particulate. For example, in the case of absorption

$$a(\lambda) = N_C \sigma_a \qquad (1.29)$$

where N_C is the number concentration of particles, in this simple example all considered identical.

Optical cross-sections are generally functions of the orientation of the particle and of the polarization of the incident light: consequently, if G is the *geometrical cross-section* of the particle, dimensionless *efficiency factors* Q_a, Q_b and Q_c can be defined for absorption, scattering and attenuation respectively, where

$$Q_a \equiv \frac{\sigma_a}{G} \,, \qquad (1.30)$$

$$Q_b \equiv \frac{\sigma_b}{G} \,, \qquad (1.31)$$

$$Q_c \equiv \frac{\sigma_c}{G} \,. \qquad (1.32)$$

Once again

$$Q_c = Q_a + Q_b \,. \qquad (1.33)$$

1.2.4 Particle size distributions

Particle size distributions (PSDs) have been traditionally approximated with a number of different mathematical descriptions, which can be divided into continuously increasing and peaking distributions. The most common example of the former is the power law distribution, while the most common example of the latter is the log-normal distribution. The two are not employed indifferently: power laws are often used to model natural seawater particle populations, while log-normals are better suited to model cultures or subpopulations. However, much as a population is the sum of its subpopulations so a power law distribution can be envisioned as the sum of a number of log-normal distributions (e.g. Jonasz & Fournier, 1996). In practical terms, a PSD is defined as

$$N(D) = N'(D)dD, \qquad (1.34)$$

where *N'(D)* is the underlying, "true" distribution of particle sizes (i.e. the *density function* of the PSD) and *dD* is the width of the size class bin (Jonasz & Fournier, 2007).

1.2.4.1 Power law distributions

The traditional mathematical description for PSDs of natural seawater particle populations is the *power law*, or *Junge distribution* (Stramski & Kiefer, 1991; Ulloa *et al*., 1994) (Fig. 1.3). In this study the form

$$N'(D) = kD^{-\gamma} \qquad (1.35)$$

is used, where k is the scaling factor, γ is the slope and D is the diameter of the particle. The distribution is not just typical of natural particulate mixtures, but also of certain particulate aerosols, and it is in fact from there that it was originally adopted (Junge, 1963). Sheldon *et al*. (1972) found that in marine samples the concentration of particulate material was roughly the same across logarithmically equal particulate size intervals, so that it followed that small particles are far more numerous than large ones. The resulting distribution, the power law, is linear when viewed in a log-log scale. Junge-type size distributions are scale invariant: they lack a well-defined average size and their average concentration overwhelmingly depends on the concentration of the lowest size considered. No matter the scale at which they are observed, they are always self similar, so that if in a certain volume a certain type of particle is twice as frequent as a second type which is twice as large, a third type that is half the size of the first will be twice as frequent as the first and four times as frequent as the second, and so on and so forth.

1.2.4.2 Log-normal distributions

A second important way of approximating particle size distributions is through *log-normal distributions* or sums of log-normal distributions (Jonasz, 1983; Jonasz & Fournier, 1996) (Fig. 1.3). In this study, the form

$$N'(D) = \frac{e^{-\frac{(\ln D-\mu)^2}{2\sigma^2}}}{\sigma D\sqrt{2\pi}} \tag{1.36}$$

was used. The log-normal distribution is the distribution of a variable the logarithm of which is normally distributed, and μ and σ are respectively the mean and standard deviation of that normal distribution. In the study of particles and sediments, log-normals have been used to represent peaked distributions of monotypical populations, and their sum to represent mixed populations made of discrete groups, e.g. Jonasz & Fournier (1996) use an algorithm to reconstruct experimental PSDs into sums of log-normals, while Zhang *et al.* (2011) present a procedure to deconvolute VSF measurements into the contribution of various, log-normally distributed particle groups. Worthy of mention is a similar approach by Risović (1993) which substitutes gamma distributions to log-normals.

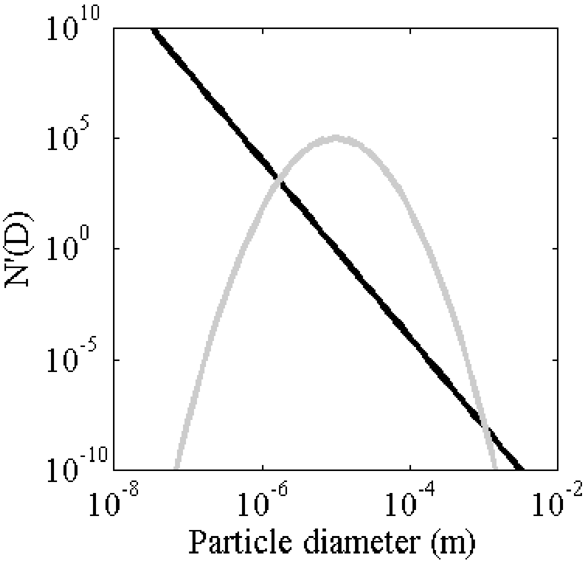


Figure 1.3 – Examples of power law (black solid line) and log-normal (grey solid line) distributions as they appear in log-log scale.

1.3 Summary

The major goals of marine optics are the precise determination of IOPs and AOPs of seawater and understanding how these can be exploited to assess and monitor the composition and concentration of the particulate and dissolved materials present within it. These two goals are fundamentally tied, as optical properties and IOPs in particular are strongly dependent on the physical properties of these materials.

The most important properties of the particle population in the context of marine optics are the size and complex refractive index of the particles, which are however also difficult to assess over the entire range of optically relevant sizes. Of particular interest in the context of PSD and particle RI studies is then flow cytometry, a technique which has the potential to retrieve both size and real refractive indices simultaneously on a particle-by-particle basis. The rationale behind this thesis work is to apply flow cytometry to the determination of PSDs and rRIs of particle populations, with particular attention to developing the capability to operate on complex natural populations: by modelling light scattering within the instrument, signal response in the sensors can be tied to size and rRI of known standards and then extended to determine size and rRI of particles in general.

To do so, an optical model of scattering is needed. Of the many available, Mie theory is the most frequently used optical model in marine optics. The hypotheses for which Mie theory holds are strict, requiring both homogeneity and sphericity of the particles: faceted mineral particles, complex organic structures, non-spherical objects like flocs and aggregates, many types of diatoms are all examples of typical components of marine particle populations which violate Mie theory assumptions. However, Mie theory is advantageous because it is analytical: given a set of input parameters, the output of the model is exact for particles of any size and any refractive index. Furthermore, Mie theory is a good first approximation for particle populations for which no previous information is available. Finally, natural marine particles are randomly orientated, and in very large numbers tend to behave like an ensemble of equivalent spheres (Morel, 1991). While this result is not strictly relevant to the reconstruction of the scattering properties of single particles within a flow cytometer (part of the interest was in fact to assess how far Mie theory could be taken in terms of size and RI determination in increasingly non-spherical particles, e.g. elongated phytoplankton cells and chains), it is one that validates the use of Mie theory for the reconstruction of total particulate IOPs from the physical properties of the particle population. For all these reasons, Mie theory is the optical model of choice in this study.

In the next chapter, a detailed look at the derivation and formulation of Mie theory will be given, together with a description of how it can be applied in forward modelling fashion to calculate IOPs from the size and refractive index distributions of a particle population.

2. Mie theory

This chapter aims to provide a short, yet comprehensive derivation of Mie theory from Maxwell's equations of electromagnetism. Mie theory is at the core of this thesis work, and forms the basis of the flow cytometric method of size and refractive index determination which will be described in Chapter 6. The second half of this chapter will focus on Mie forward modelling, and on how it can be used to calculate IOPs using the size and real refractive index properties of a suspended particle population as inputs. A thorough description of Mie theory and its derivation is available in Chapter 9 of van de Hulst (1957). An equally valid and possibly even more detailed description and derivation can also be found in Chapter 4 of Bohren & Huffman (1998). Much of the following is informed by these works.

2.1 Derivation of Mie theory

The goal of Mie theory is to provide an analytical solution to the problem of scattering by homogeneous spheres of arbitrary size. As any optical scattering problem is at its core an electromagnetic one which involves the determination of the incident, internal and scattered light fields, a general, exact solution necessarily involves Maxwell's equations of electromagnetism. Coincidentally, this justifies the proper definition of Mie theory as Mie's solution to Maxwell's equations. Note that the derivation presented in the following is given for scatterers suspended in vacuum; the results however have general validity once wavelength of the incident light and absolute RI of the scatterer are scaled relative to the absolute RI of the surrounding medium.

2.1.1 Maxwell's equations

Let

$$\nabla \times \vec{H} = ikm^2\vec{E} \tag{2.1}$$

$$\nabla \times \vec{E} = -ik\vec{H} \tag{2.2}$$

be Maxwell's equations in a periodic field with circular frequency ω, where

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda}$$

is the *propagation constant* in vacuum with c the speed of light and λ the wavelength and

$$m^2 = \varepsilon - \frac{4\pi i\sigma}{\omega}$$

is the (squared) complex refractive index of the spherical scatterer with ε the dielectric constant and σ the conductivity. The notation m (instead of n) for the complex refractive index is used here to avoid confusion with further indices. $\vec{E}$ and $\vec{H}$ represent the electric field strength and the magnetic field strength respectively. It should be noted that the following derivation is carried out using Gaussian units, and that the (relative) magnetic permeability μ was set at 1; this limits the derivation to the special case of non-magnetic particles, which holds true for all particles considered in this study. Under this latter condition the magnetic field strength $\vec{H}$ is also exactly equal to magnetic flux density $\vec{B}$.

Let then

$$\Delta\psi + k^2m^2\psi = 0 \qquad (2.3)$$

$$\Delta\vec{A} + k^2m^2\vec{A} = 0 \qquad (2.4)$$

respectively be the scalar and vector wave equations. In spherical coordinates (r, θ, φ) eq. (2.3) is separable and has solutions of the form

$$\psi_{ln} = \begin{Bmatrix} \cos l\varphi \\ \sin l\varphi \end{Bmatrix} P_n^l(\cos\theta) z_n(mkr)\,. \qquad (2.5)$$

The first factor may be either a sine or a cosine, P_n^l are Legendre polynomials and z_n is any spherical Bessel function. Indices n and l are integers and $n \geq l \geq 0$.

At the same time, given eqs. (2.1-2.2), $\vec{E}$ and $\vec{H}$ satisfy eq. (2.4) and further solutions are given by

$$\vec{M}_\psi = \nabla \times (\vec{r}\psi) \tag{2.6}$$

$$mk\vec{N}_\psi = \nabla \times \vec{M}_\psi \, , \tag{2.7}$$

where ψ is the generic scalar wave equation solution described by eq. (2.5). By then taking u and v as two of these solutions and $\vec{M}_u$, $\vec{N}_u$, $\vec{M}_v$ and $\vec{N}_v$ as their derived vector fields (via eqs. 2.6-2.7), by substitution it can be finally found that

$$\vec{E} = \vec{M}_v + i\vec{N}_u \tag{2.8}$$

$$\vec{H} = m(-\vec{M}_u + i\vec{N}_v) \, , \tag{2.9}$$

and that these satisfy eqs. (2.1-2.2). The extended forms of $\vec{M}_\psi$ and $\vec{N}_\psi$ are irrelevant to the problem at hand, and it is sufficient to know that they are functions of their respective scalar wave equation solutions.

2.1.2 *Incident, scattered and internal field solutions*

The next step involves Mie theory proper, i.e. the scattering of a plane wave by a homogeneous sphere. Assume that the external medium is vacuum and that the incident radiation is linearly polarized. The origin is fixed at the centre of the sphere with the positive z-axis along the direction of propagation of the incident wave. The incident wave is therefore described by

$$\vec{E} = \vec{u}_x e^{-ikz+i\omega t} \tag{2.10}$$

$$\vec{H} = \vec{u}_y e^{-ikz+i\omega t} \, , \tag{2.11}$$

where $\vec{u}_x$ and $\vec{u}_y$ are the unitary vectors along the x-axis and y-axis respectively. The scalar wave equation solutions which provide eqs. (2.10-2.11) via (2.8-2.9) are found to be

$$u = e^{i\omega t} \cos\varphi \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} P_n^1(\cos\theta) j_n(kr) \tag{2.12}$$

$$v = e^{i\omega t} \sin\varphi \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} P_n^1(\cos\theta) j_n(kr)\,, \tag{2.13}$$

with j_n the spherical Bessel function derived from the Bessel function of the first kind, $J_{n+1/2}$.

The field outside the sphere is given by eqs. (2.12-2.13) plus the scattered wave. The latter has solutions

$$u = e^{i\omega t} \cos\varphi \sum_{n=1}^{\infty} -a_n (-i)^n \frac{2n+1}{n(n+1)} P_n^1(\cos\theta) h_n^{(2)}(kr) \tag{2.14}$$

$$v = e^{i\omega t} \sin\varphi \sum_{n=1}^{\infty} -b_n (-i)^n \frac{2n+1}{n(n+1)} P_n^1(\cos\theta) h_n^{(2)}(kr)\,, \tag{2.15}$$

with $h_n^{(2)}$ the spherical Hankel function derived from the Hankel function of the second kind, $H_{n+1/2}^{(2)}$. Coefficients a_n and b_n are to be determined.

Finally for the wave within the sphere

$$u = e^{i\omega t} \cos\varphi \sum_{n=1}^{\infty} mc_n (-i)^n \frac{2n+1}{n(n+1)} P_n^1(\cos\theta) j_n(mkr) \tag{2.16}$$

$$v = e^{i\omega t} \sin\varphi \sum_{n=1}^{\infty} md_n (-i)^n \frac{2n+1}{n(n+1)} P_n^1(\cos\theta) j_n(mkr)\,. \tag{2.17}$$

Coefficients c_n and d_n are also to be determined.

2.1.3 Mie coefficients

Under the assumption of sharp boundary conditions between homogeneous media,

$$\vec{n} \times (\vec{H}_2 - \vec{H}_1) = 0 \tag{2.18}$$

$$\vec{n} \times (\vec{E}_2 - \vec{E}_1) = 0 \tag{2.19}$$

where $\vec{n}$ is the normal to the surface of the sphere. Accordingly, the field components E_θ, E_φ, H_θ and H_φ have to have equal values on either side of the surface of the sphere. Components E_θ and $E_\varphi$both contain the terms v and $\frac{1}{m}\frac{\partial(ru)}{\partial r}$ while H_θ and H_φ both contain mu and $\frac{\partial(rv)}{\partial r}$, and the boundary conditions (2.18-2.19) are ensured by the continuity of these four functions.

Modified Bessel functions called Riccati-Bessel functions are now introduced to help simplify subsequent calculations:

$$\psi_n(z) = z j_n(z) = \sqrt{\frac{\pi z}{2}} J_{n+\frac{1}{2}}(z) \tag{2.20}$$

$$\chi_n(z) = -z n_n(z) = -\sqrt{\frac{\pi z}{2}} N_{n+\frac{1}{2}}(z)\,, \tag{2.21}$$

where $N_{n+1/2}$ is the Bessel function of the second kind.

The Hankel function of the second kind is also tied to the two Bessel functions by

$$H_n^{(2)}(z) = J_n(z) - iN_n(z)\,, \tag{2.22}$$

providing a third Riccati-Bessel function:

$$\zeta_n(z) = \psi_n(z) + i\chi_n(z)\,. \tag{2.23}$$

Eqs. (2.18-2.23) can then be put together with the expressions for the key terms within field components E_θ, E_φ, H_θ and H_φ to rewrite the boundary conditions in terms of ψ_n, χ_n and ζ_n and finally obtain

[mu]: $\psi_n(x) - a_n\zeta_n(x) = mc_n\psi_n(y)$

$$\left[\frac{1}{m}\frac{\partial(ru)}{\partial r}\right]: \quad \psi'_n(x) - a_n\zeta'_n(x) = c_n\psi'_n(y)$$

$$[v]: \qquad \psi_n(x) - b_n\zeta_n(x) = d_n\psi_n(y)$$

$$\left[\frac{\partial(rv)}{\partial r}\right]: \quad \psi'_n(x) - b_n\zeta'_n(x) = md_n\psi'_n(y)$$

$$\Rightarrow \quad a_n = \frac{\psi'_n(y)\psi_n(x) - m\psi_n(y)\psi'_n(x)}{\psi'_n(y)\zeta_n(x) - m\psi_n(y)\zeta'_n(x)} \tag{2.24}$$

$$\Rightarrow \quad b_n = \frac{m\psi'_n(y)\psi_n(x) - \psi_n(y)\psi'_n(x)}{m\psi'_n(y)\zeta_n(x) - \psi_n(y)\zeta'_n(x)}, \tag{2.25}$$

where $x = ka = \frac{2\pi a}{\lambda}$, $y = mka$ and a is the radius of the spherical scatterer. Equipped with these definitions and the knowledge that the spherical Hankel function $h_n^{(2)}$ has asymptotic behaviour

$$\lim_{kr\to\infty} h_n^{(2)}(kr) = \frac{i^{n+1}}{kr}e^{-ikr}$$

the expression for the scattered wave is ultimately obtained as

$$u = -\frac{i}{kr}e^{-ikr+i\omega t}\cos\varphi \sum_{n=1}^{\infty} a_n \frac{2n+1}{n(n+1)}P_n^1(\cos\theta) \tag{2.26}$$

$$v = -\frac{i}{kr}e^{-ikr+i\omega t}\sin\varphi \sum_{n=1}^{\infty} b_n \frac{2n+1}{n(n+1)}P_n^1(\cos\theta)\,. \tag{2.27}$$

Lastly, by defining

$$\pi_n(\cos\theta) = \frac{P_n^1(\cos\theta)}{\sin\theta} \tag{2.28}$$

$$\tau_n(\cos\theta) = \frac{d}{d\theta}P_n^1(\cos\theta) \tag{2.29}$$

the resulting scattered field components are

$$E_\theta = H_\varphi = -\frac{i}{kr} e^{-ikr+i\omega t} \cos\varphi\, S_2(\theta) \tag{2.30}$$

$$E_\theta = H_\varphi = -\frac{i}{kr} e^{-ikr+i\omega t} \sin\varphi\, S_1(\theta)\,. \tag{2.31}$$

The functions defined as

$$S_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} [a_n \pi_n(\cos\theta) + b_n \tau_n(\cos\theta)] \tag{2.32}$$

$$S_2(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} [b_n \pi_n(\cos\theta) + a_n \tau_n(\cos\theta)] \tag{2.33}$$

are the *scattering amplitude functions* for polarization parallel and perpendicular to the plane of scattering. These are the core quantities to be computed in any Mie scattering problem, and can be further manipulated to give

$$i_1 = |S_1(\theta)|^2 \tag{2.34}$$

$$i_2 = |S_2(\theta)|^2. \tag{2.35}$$

These two terms provide the value of the scattered intensity, as will be shown in the next paragraph.

2.1.4 Efficiency factors and scattered intensity

Once a_n, b_n, $S_1(\theta)$, $S_2(\theta)$, i_1 and i_2 are known, scattered intensity and all three efficiency factors for attenuation, scattering and absorption (1.30-1.32) of a homogeneous sphere of known radius and refractive index can be calculated exactly. These are

$$Q_c = \frac{2}{x^2}\sum_{n=1}^{\infty}(2n+1)Re(a_n + b_n)\,, \tag{2.36}$$

$$Q_b = \frac{2}{x^2}\sum_{n=1}^{\infty}(2n+1)[|a_n|^2 + |b_n|^2] \tag{2.37}$$

and for absorption from (1.33) simply

$$Q_a = Q_c - Q_b\,. \tag{2.38}$$

Parameter x is the same $x = ka = \frac{2\pi a}{\lambda}$ size parameter that appears in eqs. (2.20-2.23). Note that the summations of eqs. (2.36-2.37), just like those of eqs. (2.12-2.17), (2.26-2.27) and (2.32-2.33), are defined for an infinite number of terms; these are required to provide the exact Mie solution. In practice, the series are truncated at maximum value n_{max}, which in this study was defined as $n_{max} = 2 + x + \sqrt[3]{x}$ (rounded up to the nearest integer).

The calculation of the scattered intensity requires some further considerations. As will be described in detail in Chapter 3, the laser beam within the flow cytometer used in this study is split into two parallel beams with overlapping distributions, with linear polarizations set perpendicular to each other at +45° and -45°. In the ideal case all particles processed by the flow cytometer will flow through the centre of the combined distribution, intercepting equal amounts of light from both polarizations. For polarization perpendicular to the plane of scattering, scattered intensity is given by

$$I_{\perp} = \frac{i_1}{k^2 r^2} I_0 \tag{2.39}$$

where I_0 is the incident intensity. Analogously, for polarization parallel to the plane of scattering, scattered intensity is given by

$$I_{\parallel} = \frac{i_2}{k^2 r^2} I_0\,. \tag{2.40}$$

The light scattered in any one direction from the particles in the flow cytometer will contain varying fractions of intensity comprised between the extremes of (2.39-2.40), and due to the relative geometry of the +45°/-45° polarization such fractions will compensate each other, so that the total scattered intensity will be $\propto (i_1 + i_2)$. This is true for both side scattering, which doesn't select for polarization, and forward scattering, for which the twin photodiodes filter +45°/-45° polarizations to give information on laser beam alignment. Accordingly, total scattered intensity is calculated using the formula for unpolarized light

$$I = \frac{i_1 + i_2}{2k^2r^2} I_0 \,. \qquad (2.41)$$

Correct proportionality between this modelled intensity and measured intensity is then ensured by scaling factors, as will be described in Chapter 6.

2.2 Mie forward modelling

Just as optical models can be used to derive the physical properties of the particle population from its scattering behaviour, the physical properties of the particles (once known) can be used to shed light on their bulk scattering behaviour. As will be shown in Chapter 6, optical modelling of the scattering within the flow cytometer provides the means to determine the PSDs and particle refractive index distributions (PRIDs) within a sample; and as will be shown in Chapter 7, forward optical modelling using the PSDs and PRIDs so determined provides the means to estimate the IOPs of the sample. This also offers a potential way to test the reliability of flow cytometry as a technique for determining diameters and rRIs of natural mixed particle populations, as successful comparison with IOPs obtained using other techniques establishes *optical closure* between different instruments: if all measurements are found to agree with each other, the various techniques can be confidently used to inform and support one another. In doing so there is also an implicit evaluation of the extent to which Mie theory, with all its assumptions, is sufficient to support this kind of study.

2.2.1 Theory

The foundation of Mie forward modelling rests on the additive nature of the inherent optical properties. All IOPs are the result of the sum of the single contributions from each of the constituents of a water volume, pure water itself included.

Recalling eq. (1.27) the expression of an axially symmetrical total VSF is given as

$$\beta(\theta,\lambda) = \beta_w(\theta,\lambda) + \sum_{i=1}^{n} \beta_i(\theta,\lambda)\,, \tag{1.27}$$

where $\beta_w(\theta,\lambda)$ is the VSF contribution of pure water and $\beta_i(\theta,\lambda)$ the VSF contribution of the *i*-th component in the water volume. The concept extends to integrating contributions from single particles which collectively make up the suspended particle population. From the definition of PSD given in Chapter 1, it has been shown that

$$N(D) = N'(D)dD\,, \tag{1.34}$$

where $N'(D)$ is the underlying density function of the PSD, i.e.

$$N'(D) = \frac{dN}{dD}\,. \tag{2.42}$$

It follows that eq. (1.27) is rewritten as

$$\beta(\theta,\lambda) = \beta_w(\theta,\lambda) + \int \beta_D(\theta,\lambda)\, N'(D)dD\,, \tag{2.43}$$

where $\beta_w(\theta,\lambda)$ is the VSF contribution of pure water and $\beta_D(\theta,\lambda)$ the VSF contribution of a single particle of diameter *D*. Since in turn $\beta_D(\theta,\lambda)$ is the result of the contribution of particles with different refractive indices *n*, eq. (2.43) is more rigorously written as

$$\beta(\theta,\lambda) = \beta_w(\theta,\lambda) + \iint \beta_{D,n}(\theta,\lambda) N'(D,n)\, dD dn\,, \qquad (2.44)$$

where $\beta_{D,n}(\theta,\lambda)$ is the VSF contribution of a single particle of diameter D and refractive index n and $N'(D,n)$ the corresponding value of the density function. Scattering and backscattering for the water volume are then simply calculated by integrating the total $\beta(\theta,\lambda)$ over the appropriate angle ranges, following eqs. (1.21-1.22).

Analogously, for absorption

$$a(\lambda) = a_w(\lambda) + \sum_{i=1}^{n} a_i(\lambda) \qquad (1.26)$$

and

$$a(\lambda) = a_w(\lambda) + \iint a_{D,n}(\lambda) N'(D,n)\, dD dn\,, \qquad (2.45)$$

where $a_w(\lambda)$ is the absorption contribution of pure water, $a_i(\lambda)$ the absorption contribution of the *i*-th component in the water volume and $a_{D,n}(\lambda)$ the absorption contribution of a single particle of diameter *D* and refractive index *n*.

It follows from eqs. (1.29-1.30) that for spherical particles

$$a(\lambda) = a_w(\lambda) + \iint \frac{\pi D^2}{4} Q_a^{D,n}(\lambda) N'(D,n)\, dD dn\,, \qquad (2.46)$$

2.2.1.1 Single particle VSF

Before beginning the derivation of the single particle VSF, the definitions for spectral irradiance and spectral intensity are recalled here from Chapter 1:

$$E(\vec{x},t,\lambda) \equiv \frac{\partial \Phi}{\partial A} \qquad (1.6)$$

$$I(\vec{x},t,\hat{\xi},\lambda) \equiv \frac{\partial\Phi}{\partial\Omega}. \tag{1.11}$$

By integrating over the whole radiant sphere and whole solid angle respectively and equating the results, the relationship

$$4\pi r^2 E = 4\pi I \Rightarrow E = \frac{I}{r^2} \tag{2.47}$$

is obtained, where *r* is the distance from the scattering centre. The definition for the spectral volume scattering function is also recalled here as

$$\beta(\psi,\lambda) = \lim_{\Delta V \to 0} \frac{I_s(\psi,\lambda)}{E_i(\lambda)\Delta V} \tag{1.20}$$

Given eq. (2.41), a simple substitution within eq. (1.20) in the limit of a point scatterer (which holds true when $r \gg a$ where a is the radius of the spherical particle) and considering the axial symmetry of the scattering finally produces

$$\beta(\theta,\lambda) = \frac{1}{2}\frac{1}{k^2 r^2}\frac{(i_1+i_2)I_i(\lambda)}{E_i(\lambda)} = \frac{1}{2}\frac{(i_1+i_2)}{k^2} \tag{2.48}$$

which is the *single particle spectral volume scattering function* for a homogeneous sphere under unpolarized incident light. Given eqs. (2.44), (2.46) and (2.48), the Mie forward modelling procedure therefore requires the determination of i_1, i_2 and Q_a as defined by eqs. (2.34-2.35) and (2.38) for each *n* and *D* pair. It is important to remember that *n* represents a complex refractive index, comprising both a real component rRI and an imaginary component iRI. While the former is half of the focus of the flow cytometric method described in this work (together with particle diameter), the latter cannot be determined by it; literature values will be adapted for iRI instead, the specifics of which will be discussed in Chapters 4 & 7, where applications of Mie forward modelling are presented.

2.2.2 *PSD extrapolations*

It follows from the previous section that any kind of forward modelling will necessarily require the entire optically relevant particulate to be included as the input, or the output will not be comparable with any independently measured IOP. As noted in Chapter 1, flow cytometers can cover particle diameter ranges which contribute a large fraction of the total scattering and backscattering: nevertheless, the entire optically relevant range spans from tens of nanometres to a few millimetres. Therefore, the undetectable fraction of the particle population has to be accounted for using an approximation of the PSD to extend the range of the distribution over the whole relevant range. Any of the commonly used approximations for the PSD may be used towards this end: Green *et al*. (2003a) used a best fit of the measured PSDs made using a Junge distribution model. This same approach is used in this study, and its results will be discussed in Chapter 7.

Values for the real refractive index in the Junge extensions have to be approximated as well, and various methods may be used. The most conservative approach is to assign a fixed value to the rRI, either the same on both ends of the extension (i.e. for both particles smaller and larger than the detectable diameter range) or a separate one for each. The specific rRI values may be assigned in a variety of ways, e.g. by using the average rRI value of the measured fraction of PSD on both ends of the extension; by using the average rRI values at the extremes of the measured fraction of PSD on the respective ends of the extension; or by using literature-derived rRIs values based on an hypothesis of the probable particle population composition for the size ranges covered by the extension (cf. Tab. 1.1). Details on the rRI approximations used in this study will also be given when presenting the application of Mie forward modelling to natural marine particle populations in Chapter 7. Values for the imaginary refractive index iRI, which as will be shown are assigned according to the value of the real component of the RI (i.e. according to the type of material specific to each particle), will be dependent on the rRI approximation used.

2.3 Summary

As anticipated in Chapter 1, when Mie modelling is informed by the proper characterisation of the optical setup and performance of a flow cytometer, a comparison between flow cytometric measurements of standard particles with known diameter and RI and corresponding modelled results becomes possible: the flow cytometric method for size and RI determination rests on these very premises. The first half of this theoretical framework was established by exploring in detail how Mie's solution to Maxwell's equations, generally known as Mie theory, provides the means to analytically calculate the angular structure of scattered intensity and the efficiency factors for attenuation, absorption and scattering for a spherical scatterer of any diameter and RI. It was then further described how Mie theory allows for the calculation of bulk IOPs when its results are integrated over a whole particle assemblage for which PSD and PRID are known.

Flow cytometry, the eponymous second half of the theoretical framework behind the flow cytometric method, will be the subject of the next chapter. In it, a general overview of flow cytometry both as a technique and a technology will be given, followed by specific details about the flow cytometer used throughout this study and its operation. Finally, a characterisation of the flow cytometer in its performance will be carried out, detailing correction parameters designed to ensure precise particle concentration measurements.

3. Flow Cytometry

The following chapter will focus both on flow cytometry as a measurement technique and on flow cytometers as practical instruments. The first half of the chapter will offer an overview of the methodology, and include details on the systems and functions common to all flow cytometers. It will then give specific details on the particular instrument used in this study, the CytoSense (CytoBuoy b.v., Netherlands), including procedures for its correct operation and maintenance. The second half of the chapter will then be dedicated to the characterization procedures which were used to assess the precision of the particle concentration measurement capabilities of the CytoSense.

3.1 Overview

Strictly (and etymologically) speaking, flow cytometry is the process by which a fluid stream of cells is focused so that they pass in single file through a measurement apparatus where a number of measurements of their physical and chemical characteristics is carried out. By extension, the process is applicable to biological particles in general and, particularly in the context of this study, to inorganic suspended particles as well.

The nature of the measurements carried out on the particles can be quite diverse, as are the types of sensors that can be employed. Historically, sensors used in flow cytometry have included acoustic, electronic, optical and radiation sensors (Shapiro, 1988). Optical sensors have been the most widely used and are still common; advancements in the quality and availability of CCD and CMOS imaging sensors and digital computing power have also resulted in the rise of imaging flow cytometry, based on direct image analysis of the particles (Basiji *et al.*, 2007). A number of flow cytometry systems now focus exclusively on imaging-based approaches (particularly for biology applications), using different techniques to tackle the superior resolution requirements inherent to imaging. Examples include: camera-based systems using time delay integration of multiple images (Basiji *et al.*, 2007; Elliott, 2009); timed pulse excitation of fluorescence and synchronised imaging to eliminate motion blur

(Gorthi *et al.*, 2013); and PMT-based systems for ultrafast optical imaging using a broadband femtosecond laser (Goda *et al.*, 2012). Worthy of mention is also the Imaging FlowCytobot (Olson & Sosik, 2007), a system geared towards simple brightfield imaging but designed to be deployed underwater *in situ* for extended periods of time, conducting automated monitoring of algae populations through a supervised learning algorithm (Sosik & Olson, 2007). In all cases, many kinds of measurements can be carried out on the analysed particles, characterising intrinsic, structural parameters such as size, shape, structure and concentration of the particles or determining DNA content, composition and physiology of cells and fluorescent lifetime of phytoplankton in association with fluorescent dyes and DNA stains (Shapiro, 1988). Overall, the key application of flow cytometry is the characterization of subpopulations of particles within mixed populations. This can be ultimately achieved either through identification of discrete groupings by isolating subsets within the multivariate datasets produced by the instrument (*gated analysis*) or by physically isolating and separating a specific subpopulation of interest from the particle population (*particle sorting*), although the latter technique is usually informed by the former.

Flow cytometry is akin to microscopy, as both techniques employ a light source to illuminate a specimen and collect the light transmitted, scattered or re-emitted by it. In microscopes the light forms a real image in the objective, which is then observed directly by the operator or visualized on a screen; in flow cytometers the light is collected by photodetectors, and converted into electrical signals or images to be subsequently processed. However, flow cytometers also differ from microscopes for a few fundamental reasons. Microscope operators actively navigate the sample by moving it in, out and around the field of view, limiting the processing rate of the sample to low levels. Flow cytometers instead process individual particles for a few microseconds each, providing much higher sampling rates than those available to microscopy. Accordingly, illumination in a flow cytometer must be many times more powerful than in a microscope, as flow cytometer sensors must collect in a few microseconds the light that the eye of the microscope operator can collect in a few milliseconds (Shapiro, 1988). Optical flow cytometry commonly employs lasers as light sources (Shapiro & Telford, 2009), being both powerful and versatile as monocromatic lasers can target individual fluorescent dyes and chlorophyll groups. It also generally resorts to darkfield techniques, based on collection of light scattered and fluoresced at large angles. This is again in contrast with microscopes, which commonly operate using brightfield (transmitted light) techniques. Imaging flow cytometers can straddle the divide between the two, offering either darkfield images of fluorescent emission and side scattering (from laser emission), brightfield images of transmitted light (e.g. from LED sources) or a combination of the two (Basiji *et al.*, 2007).

In general, flow cytometry has a few key advantages when compared with other techniques (Dubelaar & Jonker, 2000): high sample throughput, permitting statistically significant results within shorter sampling time than other techniques (e.g. microscopy); a single-particle approach to particle analysis, compared to the bulk approach of other methods of particle analysis (e.g. laser diffractometry); and the large variety of parameters available for the study and characterization of the particles. Flow cytometry has drawbacks as well (Dubelaar & Jonker, 2000). It is a mechanically intensive process, and the strain put on the particles as they are pumped and focused to a single file can break up chains and aggregates. Furthermore, although flow cytometers can sample thousands of particles per second, the processed sample volumes are quite small (~1-10 μl/s), so that measurements are time consuming when trying to achive statistical significance for low concentrations of particles (e.g. large particles in natural mixed populations). Finally, flow cytometers are very expensive pieces of equipment and require continuous care to be kept in working order.

3.2 Flow cytometry fundamentals

3.2.1 Flow system

In a typical optical flow cytometer the single file flow of particles is generated by coaxially injecting the sample or core flow into a flowing jacket of particle-free fluid known as sheath fluid (Fig. 3.1). To avoid turbulence at the interface between the two fluids the speed of the flow is adjusted to make the flow laminar. The transparent chamber (*flow cell*) into which sheath and core are injected tapers gradually, preventing disturbances while accelerating the flow to the required single file regime. Most often the flow is vertical and directed either naturally downwards or upwards to facilitate the removal of air bubbles. The sheath fluid must also be as clean as possible as well as close as possible to the refractive index of the medium within which the particulate is suspended. This eliminates interference with the measurements in the form of scattering from particles outside the core flow and/or refraction from the interface between core and sheath.

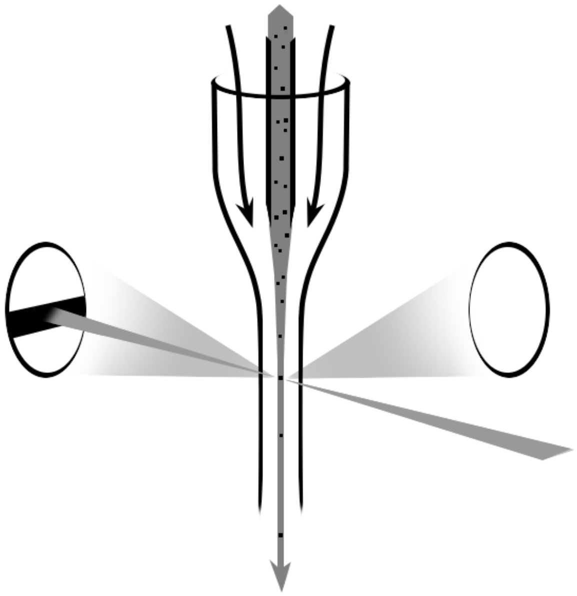


Figure 3.1 – Typical setup of a classic optical flow cytometer. The vertical sample flow is hydrodynamically focused by the sheath fluid so that particles intercept the incident laser beam one by one. Scattered light is collected at near forward angles and orthogonally to the beam, while the laser beam proper is intercepted by a beam stop before reaching the sensors.

3.2.2 Optics system

Incident laser beam and optical sensors are arranged around the flow cell in an orthogonal fashion. The light beam is focused using a cylindrical lens to be at its narrowest where it intercepts the focused core flow. It is important to note that laser beams most commonly operate to produce a Gaussian-shaped energy distribution in the beam. Width in a such a beam is defined as the distance between the points at which the intensity falls to e^{-2} times the peak, and the core flow must be confined well within this width to avoid the scattering from identical particles to be overly dependent on their position within the flow (Fig. 3.2).

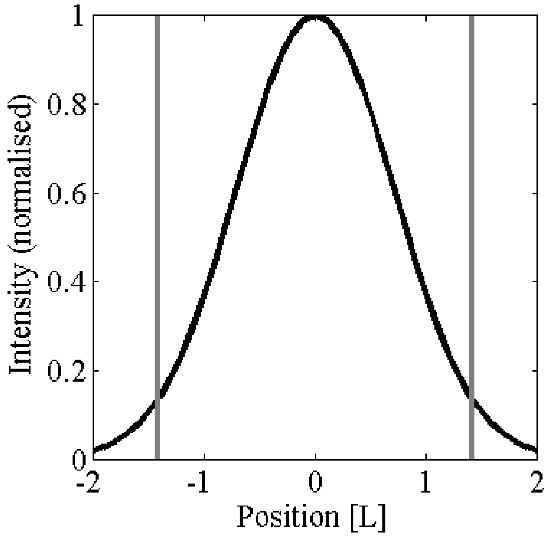


Figure 3.2 – Power distribution of the laser beam in a typical flow cytometer. The vertical lines mark the beam width, defined as the space between the points where the intensity of the beam decreases to e^{-2} times the value of its peak. The core flow must be centred and confined within a column much thinner than the beam width to avoid particle scattering being dependent on the position of the particle within the flow. Note that the x-axis values illustrate a theoretical position relative to the center of the distribution; the physical dimension is length, but the specific unit (and values) will depend on the particular instrument used.

The forward scattering detector (most often a photodiode) collects light scattered next-to-axially with respect to the incident laser; here some form of beam stop blocks the bulk of the transmitted beam while permitting light scattered at small angles to reach the detector proper (Fig. 3.1). The side scattering detectors (most often photomultiplier tubes, PMTs for short) collect light scattered and fluoresced orthogonally with respect to the laser beam, with a large optical aperture to collect as much light as possible. Scattered and fluoresced light are separated after collection: fluorescence measurements use bandpass optical filters to select a specific desired wavelength, such as chlorophyll emission wavelengths or wavelengths specific to particular dyes.

3.2.3 Signal Processing

Acquisition of particle data in a flow cytometer starts when a particle reaches the observation point, hitting the laser beam and scattering its light. The scattered (or fluoresced) light enters the detectors and produces signals which are amplified into voltage pulses. One or more of these channels are used as triggers; if the signal pulse reaches above a threshold level, the acquisition and analysis of the signal can begin.

In older optical flow cytometer models, signal analysis is analog (Shapiro, 1988) (Fig. 3.3). A comparator circuit using both signal pulse and threshold voltage as inputs discriminates signals reaching above the threshold, usually set by the operator from a computer interface. When the signal reaches above the threshold voltage the comparator outputs a logical one, which is a set voltage dependent on the electronics used; otherwise, the comparator output is logical zero, or ground. When the comparator outputs one, the electronics "know" a particle is present. Analog analysis of the signal is carried out using peak detectors and integrators: both circuits store their values in capacitors, which can be small since within the short times involved in single particle analysis the voltage stored in the capacitors drops negligibly.

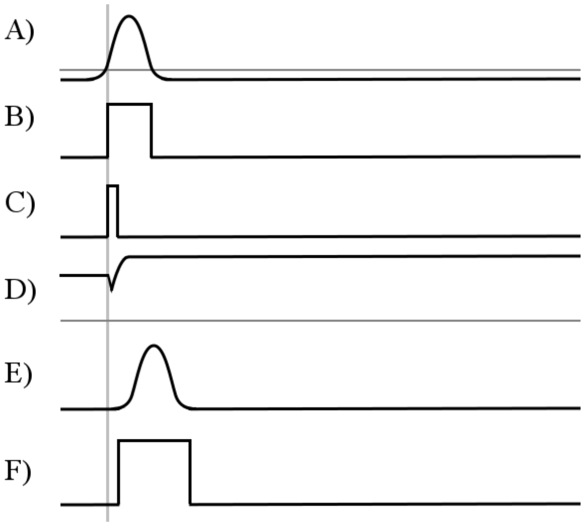


Figure 3.3 – Basic signal pulse analysis in older flow cytometers. A) Particle signal pulse and threshold level (grey line). B) Comparator output. C) Trigger pulse. D) Peak detector. E) Delayed particle signal pulse. F) Integrator gating.

The output state change in the comparator triggers a brief reset pulse which is sent to the capacitors, resetting the previously stored voltage and readying them for the new pulse. The change of state in the comparator and the subsequent reset signal are not instantaneous; taken together, their duration amounts to a few hundred nanoseconds. While this is not a problem for peak detectors, since particles take a few microseconds to transit and the peak detector capacitor is reset and ready to accept a new value well before the signal peak, integrators would miss out a chunk of the signal pulse. Because of this the original signal pulse is not fed directly into the integrator, but goes through a delay line first. Timing pulses derived from the comparator output are used to precisely gate the integrator, as precise timing is essential for the integrator as starting or ending integration at the wrong time may add spurious noise to the total value. At this point the voltage values in the peak detectors and integrators can be converted to digital and sent to a computer for further analysis.

In modern optical flow cytometer models, signal analysis is digital. A comparator circuit is used to discriminate when a signal reaches above the user-operated threshold, but no peak detectors or integrators are present. The comparator triggers the analog-to-digital conversion of the raw analog voltage signals, which are then sent to a computer for analysis. The digital pulse shape can be analysed to provide basic parameters such as baseline width (time-of-flight), full-width at half-maximum, peak height and total area, and algorithms can be developed to estimate values for a number of other parameters such as skewness and number of peaks of the curve. Imaging flow cytometers function much in the same way, with the obvious difference that particle information comes from image analysis of the snapshots triggered by the particle events.

In all cases, particle parameters are ultimately typically visualised in a characteristic dot plot, or *cytogram*; two among the many available variables are chosen as the *x* and *y* inputs and each parameter pair (corresponding to a single particle) is displayed as a dot on the graph. The cytograms are then visually inspected to identify subgroups within the particle population which can be ultimately isolated using gated selections (Adan *et al.*, 2016).

3.3 CytoBuoy CytoSense

3.3.1 Overview

The CytoBuoy CytoSense is an optical flow cytometer with brightfield imaging capabilities. It employs a square section cuvette (inner width: 1 mm) to hold the flow of core and sheath fluid to minimize edge effects and refraction from the flow cell walls, and features a two stage sheath injection system (Fig. 3.4). Fluids within the CytoSense flow cell flow upwards to facilitate air removal: bubbles are trapped and removed from the upper part of the sheath fluid injectors. In the first stage of the injector a stream of sheath fluid is injected coaxially with the core flow needle in classic flow cytometric fashion. The combined flow is then focused further by a triple injection of sheath fluid from three needles set at a 120° angle from each other.

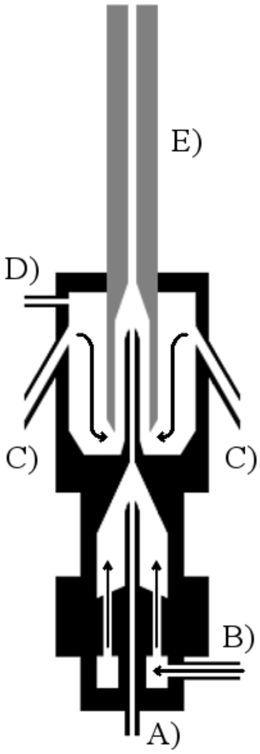


Figure 3.4 - Diagram of the flow cell and injector assembly in the CytoSense. A) Sample injector. B) Sheath fluid 1st stage injector. C) Sheath fluid 2nd stage triple injector (two shown). D) Air bypass. E) Flow cell.

This provides some degree of control on the stability of the core flow by moving the core flow relative to the three injection needles. The final speed of the focused combined flow is ~2 m/s, resulting in a sample core about 60 µm wide (the final width of the core varies depending on the flow rate of the sample pump). The CytoSense used in this study uses a single 488 nm semiconductor laser (OBIS 488LS, Coherent Inc., CA). It deals with the issue of presenting a Gaussian power distribution to the particles by splitting the laser beam into two parallel beams with overlapping distributions, producing a flat-topped total distribution of radiant power (Fig. 3.5).

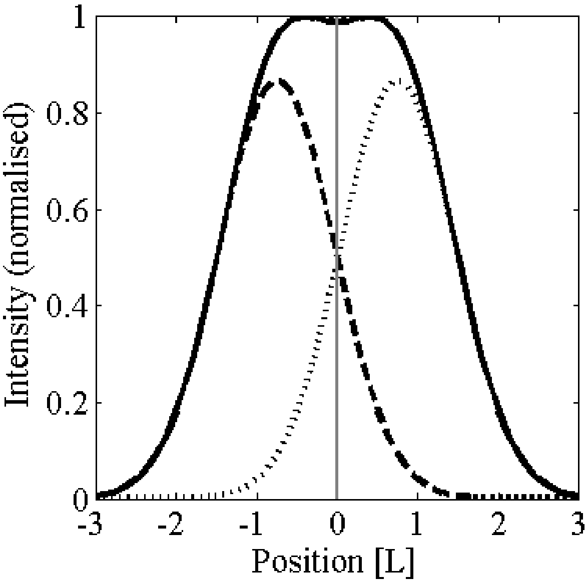


Figure 3.5 – Power distribution of the incident laser beam in the CytoSense. The laser beam is polarized to produce two gaussian distributions, one with +45° polarization (dashed line) and one with -45° polarization (dotted line). Due to their polarization, no interference takes place between the two distributions and their sum (solid line) presents a flatter profile to the incoming particles. Note that the diagram is qualitative and doesn't reflect the exact power distribution within the instrument.

To avoid interference between the incident beams their polarization is set perpendicular to each other at 45° and -45° degrees; as anticipated in Chapter 2 the combination of orthogonally polarised input beams produces scattering signals which are essentially equivalent to those that would be produced by an unpolarised input beam of the same intensity. A further advantage of the flat-topped distribution is that it provides a way of determining the position of the particle relative to the center of the beam and the means to adjust it accordingly: the forward scattering sensor is made up of a pair of twin photodiodes, each collecting light scattered to the left or right of a vertical beam stop (in the CytoSense a gap between prisms used to redirect forward-scattered light towards the detectors); particles going through the centre of the beam will scatter roughly even amounts of light into each photodiode, so that the balance between left and right forward scattering signal may be used as an indicator of beam alignment. The composite laser beam is further shaped into a flat blade of light before passing through the cuvette, with a 5 μm thickness and a ~300 μm width. This is achieved using a cylindrical lens (Fig 3.6)

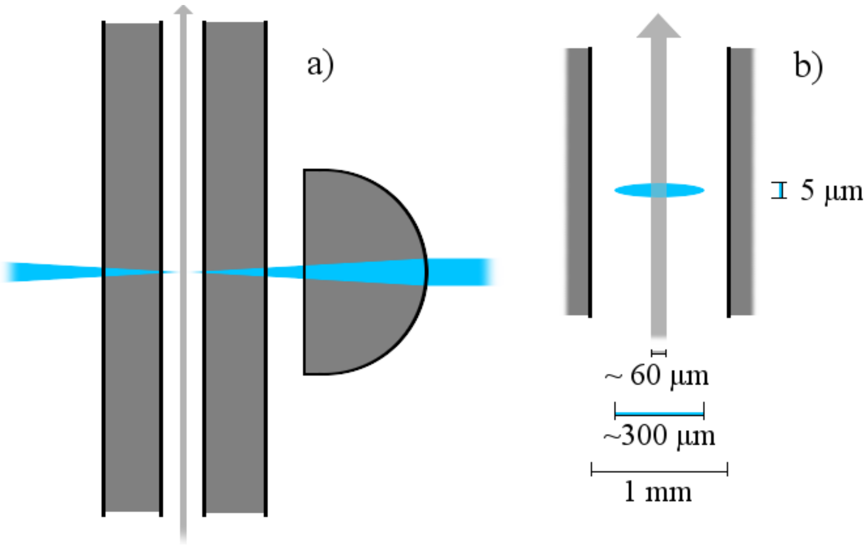


Figure 3.6 – (a) Beam shaping lens configuration within the CytoSense and (b) detail of the relative positions of sample core flow (light grey) and shaped laser beam (blue). Note that both diagrams are not to scale.

The sensor apertures within the instrument have different shapes for forward scattering and side scattering/fluorescence respectively. As mentioned, the forward scattering aperture intercepts two vertical strips of light on both sides of the laser beam using twin prisms. The light is then re-directed vertically at 90° through a circular hole before reaching the detectors. The final effective shape is that of two vertical segments of a circle collecting between 2° and 9.7° of polar angle on either side of the direction of the laser beam. The side scattering aperture is large and circular instead, designed to collect as much light as possible. It collects between 45° and 135° of polar angle. The geometry of the sensor apertures will be discussed in further detail in Chapter 6.

The fluidics system within the CytoSense is divided into 3 main lines (Fig. 3.7). The *sample line* is the shortest and is used to take the sample from its container to the flow cell. It is driven by a peristaltic pump the flow rate of which can be set by the operator. Adjusting the flow rate of the pump influences the width of the core flow. The *sheath line* is divided in two by the particulate filters. It is driven by a vacuum pump, the flow rate of which can also be set by the operator. Adjusting the flow rate of the pump influences the speed of the core flow. The first section of the sheath line takes the mixed sample waste from the flow cell towards the particle filters. Filtering happens in two stages: a first 0.45 μm main catch filter effects the bulk of the filtering, while a second 0.1 μm cleanup filter removes remaining particles, including ones produced by the wear in the sheath vacuum pump, which is located between the two filters. An overpressure outlet connects here – fluid exiting this outlet is a visual cue of clogging in the system. The second section of the sheath line re-circulates the filtered sheath fluid, feeding it back into the two-stage injector and closing the circle. Finally, the *waste line* takes excess fluid and any excess air from the main catch filter to a waste container. The waste line also includes a small 5 μm hydrophobic filter.

Unlike older flow cytometer models, the CytoSense does not use peak detectors or integrator circuits to quantify the signals produced in the sensors after triggering, but rather digitizes the signal pulses at 4 MHz and sends the resulting profiles to the computer for further analysis. The CytoSense employed in this study has three signal channels available: forward scattering, side scattering and red fluorescence (685 nm), the latter being selected using a 675 ± 15 nm bandpass filter. Particle detection can be triggered on any of the three and the threshold level can be set independently on all of them by the operator. The gain setting of the side scattering and red fluorescence photomultipliers is variable and can also be set independently on each of the two by the operator.

A description of the operation and maintenance procedures of the CytoSense flow cytometer is given in Appendix A as reference.

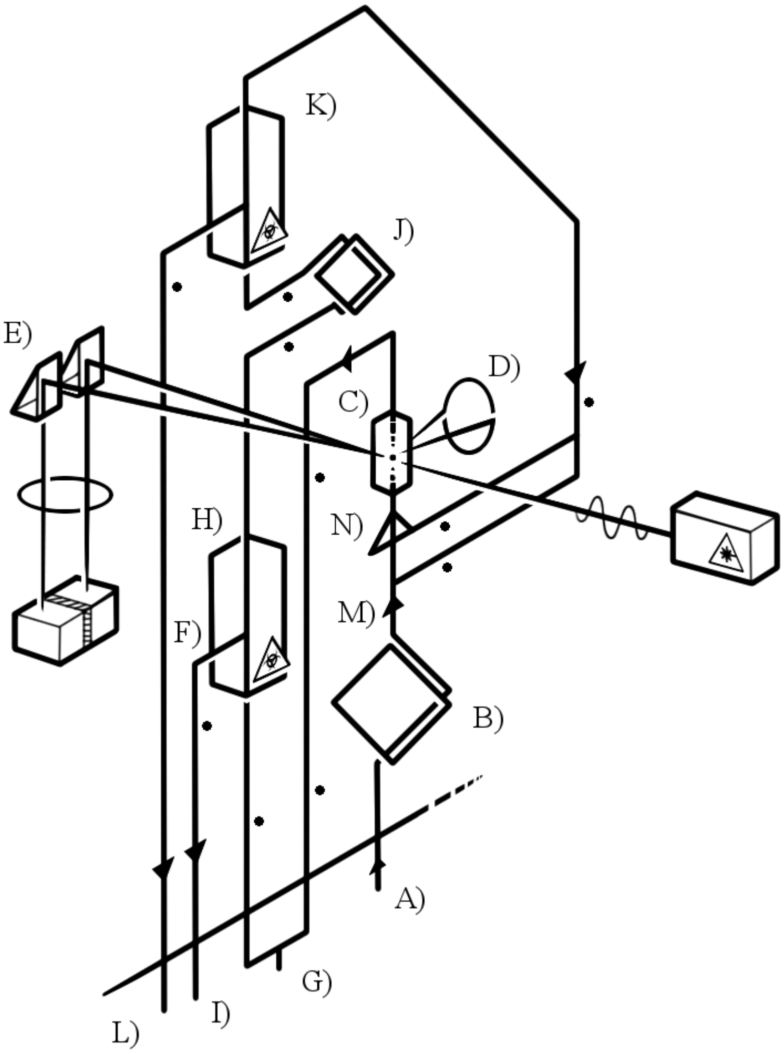


Figure 3.7 – General diagram of the fluidics system within the CytoSense and simple overlay of the sensor setup. The black dots indicate the points where the tubing sections interconnect through Luer locks. A) Sample inlet. B) Sample peristaltic pump. C) Flow cell. D) Side scattering and fluorescense sensor aperture. E) Forward scattering mirrors (note that the main laser beam is left uncollected rather than being actively removed using a beam stop. A beam stop is indeed present but located on the optical block casing wall). F) Forward scattering left/right polarization photodiodes. G) Mixed fluid external bypass (with valve for external access). H) Primary 0.45 µm glass fibre main catch filter. I) Waste outlet. J) Sheath vacuum pump. K) Secondary 0.1 µm glass fibre cleanup filter. L) Overpressure outlet. M) Sheath fluid 1st stage injector. N) Sheath fluid 2nd stage triple injector.

3.4 CytoSense characterization

This section describes the characterization procedures that were used to validate the correct functioning of the CytoSense and its capabilities for precision concentration measurements. Empirical correction factors designed to ensure the latter are also described in the following.

3.4.1 Pump flow rate calibration

3.4.1.1 Semi-automated calibration procedure

As shown in the previous sections, the CytoSense uses a peristaltic pump to carry the sample towards the flow cell assembly. The sample is injected into the sheath fluid stream, so that the speed of the sample flow within the flow cell is driven by the sheath vacuum pump. Given a fixed sample flow speed, the revolution rate of the peristaltic pump controls the amount of sample intake and the thickness of the sample core instead. Since the volume estimations provided by the instrument through its software will be the basis for all concentration measurements, a precise calibration of the correspondence between pump flow rate (given as volume flow, ml/s) and actual total sample volume intake is required, so that the instrument can give the best possible estimate of the analysed volume.

A software-driven estimation of the analysed volume is particularly necessary because *total processed volume* (sample volume used up during a measurement) and *total analysed volume* (sample volume actually analysed by the instrument) are not coincident. The discrepancy is due to data transfer overhead – the time during which particle data is acquired and sent to the computer and the instrument is not analysing (but still consuming) the sample (CytoBuoy, private communication). The discrepancy is not constant, but increases as the number of detection events increases, e.g. in high particle concentration samples or with high sensitivity settings in the PMTs.

To calibrate the peristaltic pump flow rate the CytoSense uses a semi-automated calibration procedure which can be initiated from the CytoUSB interface. The procedure should be repeated periodically, and the instrument software prompts the user to do so every few months. The procedure is based on two steps: a first automated cycle of sample intake at a high pump flow rate which takes a few minutes to complete and a second phase of precision weighing of the total sample uptake. This is done by the operator by measuring the mass of the sample before and after the calibration cycle. After the automated cycle the software prompts the operator to input the mass of the sample before and after the cycle, along with the density of the calibration liquid used. The obvious (but not obligatory) choice is to use water as the calibration sample, and two pre-sets are available for it within the software (0.998 g/cm^3 for pure water and 1.024 g/cm^3 for salt water). The density value can otherwise be set by the user to suit any other calibration liquid.

3.4.1.2 Manual validation procedure

After successfully completing the calibration procedure, it is desirable to manually verify that the total processed volume correctly reflects sample intake at a set flow rate and for a set amount of time. To do so a simple verification protocol is set up. Using the same principle used during the calibration itself, precision weighings of a water sample were made, for a few different flow rates and multiple times each. Sample runs are executed at 0.13, 0.26, 0.52, 1.04, 2.53 and 5 µl/s, with the sample mass being recorded before starting each run and then three more times at 2, 4 and 6 minutes into the run. Total processed mass values are then simply compared with expected values produced by combining nominal processed mass, pump flow rate and density, calculated as $M_{ref} = flow\ rate \cdot density \cdot 0.12$, where 0.12 is the nominal volume processed in 120 s at 1 µl/s in cm^3 (Tab. 3.1). The displayed pump flow rate was found to slightly underestimate the actual flow rate, with the effect becoming increasingly visible at high pump flow rates (best fit line: $y = 1.058x - 0.003$, adj. R-squared: 0.999, RMSE: 0.004), although the high linearity of the relationship demonstrates that correct proportionality is preserved between different flow rate settings (Fig. 3.8). Forcing the intercept of the fit line to pass through zero produces marginally worse results ($y = 1.051x$, adj. R-squared: 0.999, RMSE: 0.004), but provides a correction factor (5.1%) for the actual volume intake at all pump flow rates.

Flow rate (μl/s)		0.13	0.26	0.52	1.04	2.53	5
Mass (g):	Start	50.515	50.465	50.367	50.158	49.718	48.690
	Split 1	50.500	50.434	50.306	50.032	49.408	48.068
	Split 2	50.486	50.405	50.241	49.900	49.087	47.433
	End	50.472	50.374	50.178	49.770	48.768	46.796
	ΔM1	0.015	0.031	0.061	0.126	0.310	0.622
	ΔM2	0.014	0.029	0.065	0.132	0.321	0.635
	ΔM3	0.014	0.031	0.063	0.130	0.319	0.637
	ΔMref	0.016	0.031	0.062	0.125	0.303	0.599

Table 3.1 – Values for the mass of Milli-Q water left in the sample container during the pump flow rate calibration manual validation procedure. All measurements were taken using a precision scale at the start of each flow rate run and subsequently 2, 4 and 6 minutes into the run. At higher flow rates the mass intake projected using the calibrated values underestimates the actual intake values (ΔMref) slightly.

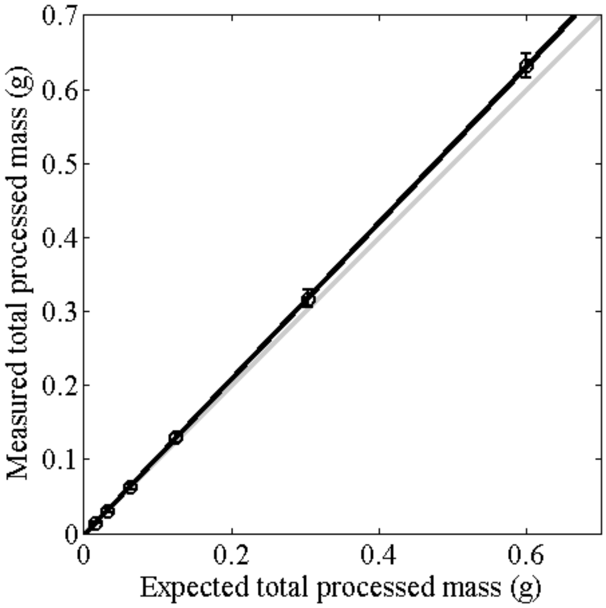


Figure 3.8 – Linear fit of the mass measurements for the pump flow rate calibration manual validation procedure as presented in Table 3.1, shown in comparison with the 1:1 line (grey line). The expected values given by the calibration underestimate the actual measured intake, but the high linearity of the relationship (best fit line: y = 1.058x - 0.003, adj. R-squared: 0.999, RMSE: 0.004) suggests that the correct proportionality is maintained between different flow rates.

3.4.2 Assessment of precision concentration measurement capabilities

The validation of the pump flow rate calibration procedure and the capability of the CytoUSB software to correctly estimate the sample volume intake enables further validation of the concentration measurement capabilities of the CytoSense instrument. This further validation can be carried out in two ways: either by using a single sample containing a known amount of particles or by using a dilution series made from a sample containing an unknown amount of particles. In the first case the validation is successful if the instrument is capable of measuring the expected amount of particles. In the second case the validation is successful if the relative concentrations measured for the series are consistent with the relative dilution.

3.4.2.1 2 µm precise counting beads

For the first case a ready-made sample of 2 µm polymer beads (EZY-CAL Microsphere Size Standards, Thermo Fisher Scientific) was used. The sample is supplied pre-counted by the manufacturer with a nominal concentration of 2000 ± 10% particles per ml, enclosed in its own bottle and complete with a magnetic stirrer. Exploratory measurements were made at 0.50, 0.99 and 2.98 µl/s. Results at low flow rates provided concentrations which were markedly lower than the nominal one, while results at higher flow rates produced concentrations closer to provided parameters (Tab. 3.2). This hints to a systematic underestimation of particle concentrations at low pump flow rates.

	Numerical concentration (ml^{-1})	
Flow rate (µl/s)	Average	St. Dev.
0.5	1127.705	78.061
0.99	1412.897	70.849
2.98	1601.83	34.351

Table 3.2 – Numerical concentration values for the EZY-CAL 2 µm polymer beads sample test run. All three values underestimate the expected value of ~2000 part./ml, with larger discrepancy at lower flow rates.

3.4.2.2 1 μm beads dilution series

For the second case a sample of 1 μm latex beads (Duke Particle Counter Size Standards, Thermo Fisher Scientific) was used, in a proportion of 1 droplet (ca. 0.05 ml) of bead suspension per 10 ml of Milli-Q water. Concentration within the bead suspension bottle is given as approximately $1x10^9$ part./ml, so that no precise quantity for the final concentration within the sample can be given. However, the nature of a dilution series is so that the operator doesn't need knowledge of the original concentration, as all subsequent concentrations in the dilutions can be normalized by the original value. The dilution series was made up of the original sample and of three dilutions with concentrations 12.5%, 25% and 50% of the original. All four samples were measured three times each with each run lasting 120 s (Tab. 3.3). Resulting concentrations were normalized by the average concentration of the three 100% concentration runs, and found to fit the 1:1 line with an adjusted R-squared value of 0.996 and an RMSE of 0.023 (Fig. 3.9).

	Num. concentration (ml^{-1})		Norm. num. concentration (ml^{-1})		
Dilution	Average	St. Dev.	Average	St. Dev.	Expected
12.5%	230.852	10.940	0.128	0.006	0.125
25%	420.904	27.666	0.233	0.015	0.250
50%	857.264	42.758	0.474	0.024	0.500
100%	1807.482	53.925	1.000	0.030	1.000

Table 3.3 – Numerical concentration values for the 1 μm polymer beads dilution series. The average values were normalised using the average concentration of the undiluted sample.

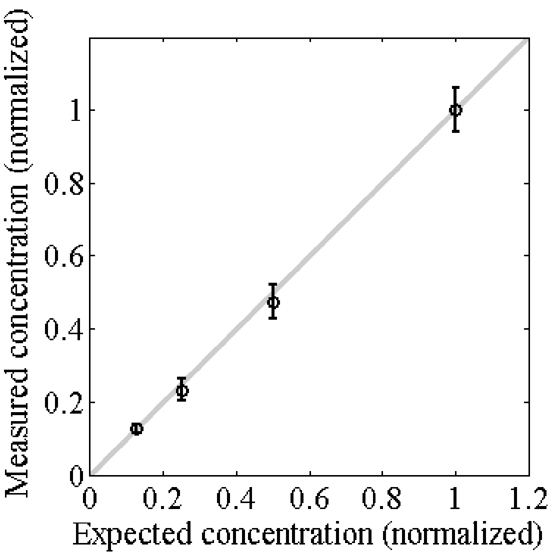


Figure 3.9 – Normalised numerical concentration values for the 1 μm polymer beads dilution series. The measured concentrations maintain the correct proportionality between dilutions (goodness of fit for the 1:1 line: adj. R-squared: 0.996, RMSE: 0.023).

3.4.3 Concentration underestimation correction

The results of the concentration measurements validation hinted at the CytoSense underestimating the actual particle concentrations when running at low pump flow rate settings. A further series of measurements was therefore carried out to characterize the underestimation and develop a correction for it, both on the EZY-CAL 2 μm polymer beads sample and on a second sample of 10 μm polymer beads (Duke Particle Counter Size Standards, Thermo Fisher Scientific).

3.4.3.1 2 μm beads measurement series

The first measurement series was carried out on the EZY-CAL 2 μm polymer beads sample for pump flow rates set at 0.99, 1.99, 2.98, 3.97, 4.97, 5.96 and 9 μl/s, with ten replicates for all settings but the last. Only seven replicates were made for 9 μl/s due to the large consumption of sample at this flow rate. The results were integrated with those presented in paragraph 3.4.2.1, and found to plateau at ~3 μl/s without ever reaching the nominal concentration expected for the sample. The expected value of ~2000 part./ml was found to be 1.088 times the plateau value obtained as the average concentration of all replicates at 4.97, 5.96 and 9 μl/s. Finally, all concentrations were normalized using the plateau value (Tab. 3.4).

	Num. concentration (ml^{-1})		Norm. num. conc. (ml^{-1})	
Flow rate (μl/s)	Average	St. Dev.	Average	St. Dev.
0.50	1127.705	78.061	0.613	0.042
0.99	1430.154	128.994	0.778	0.070
1.99	1730.309	97.764	0.941	0.053
2.98	1726.148	175.767	0.939	0.096
3.97	1766.527	113.629	0.961	0.062
4.97	1830.820	103.580	0.996	0.056
5.96	1856.900	115.952	1.010	0.063
9.00	1819.291	80.744	0.989	0.044

Table 3.4 – Numerical concentration values for the EZY-CAL 2 μm polymer beads sample at various flow rates. The average values were normalised using the average concentration of all replicates at 4.97, 5.96 and 9 μl/s. The concentration value for the plateau never reaches the expected concentration of ~2000 part./ml.

The results indicate a clear relationship between pump flow rate settings and measured concentration in the sample. Such relationship was found to be well fitted by an exponential curve of the type

$$y = ae^{-bx} + c \, , \qquad (3.1)$$

suggesting the possibility of developing a correction to compensate for the underestimation (Fig. 3.10). The correction curve is given via a least-squares fit of pump flow rate vs. normalized concentration of the reference beads, providing the means of calculating a concentration correction factor for each particular pump flow rate i.e. given a set sample pump flow rate, concentrations are corrected by dividing by the corresponding resulting factor. Curve parameters for the 2 μm polymer beads correction curve: a = -0.653, b = 1.148, c = 0.988.

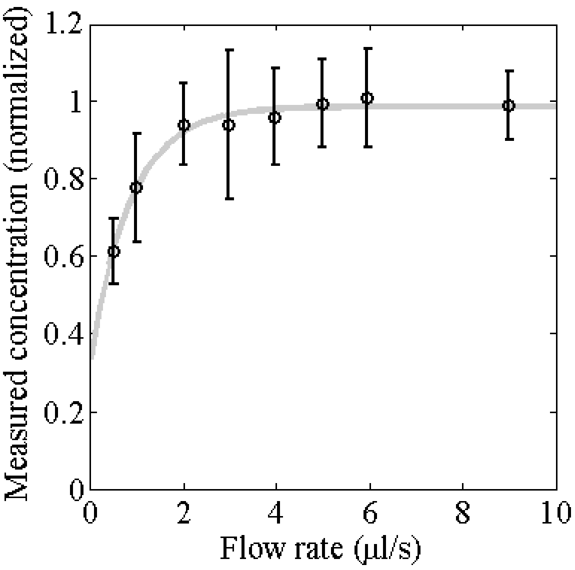


Figure 3.10 – Pump flow rate concentration correction curve ($y = ae^{-bx} + c$) for the EZY-CAL 2 µm polymer beads normalised concentration values presented in Table 3.4. Curve parameters: a = -0.653, b = 1.148, c = 0.988.

However, such relationship can be broadly applied to a mixed sample only if no large selection bias between different particle classes is present. One likely cause for underestimation may be settling/inertia of large particles, so that size variations may be subject to selection bias in concentration measurements.

3.4.3.2 10 µm beads measurement series

The same measurement protocol applied to the EZY-CAL 2 µm polymer beads sample was applied to a sample of 10 µm polymer beads (Duke Particle Counter Size Standards, Thermo Fisher Scientific) to investigate the influence of size selection bias in the underestimation of particle concentrations at low pump flow rates. The measurement series used the same pump flow rate settings as the 2 µm beads series, starting from 0.99 µl/s, and a smaller number of replicates (6 for all flow rates but 9 µl/s, for which 4 replicates were made). The actual concentration within the sample is unknown a priori, but not needed as all data are normalized to the value of the plateau (Tab. 3.5).

	Num. concentration (ml^{-1})		Norm. num. conc. (ml^{-1})	
Flow rate (µl/s)	Average	St. Dev.	Average	St. Dev.
0.99	2338.222	393.411	0.641	0.108
1.99	2879.071	388.492	0.789	0.106
2.98	3247.641	93.151	0.890	0.026
3.97	3431.940	103.743	0.941	0.028
4.97	3550.008	131.770	0.973	0.036
5.96	3677.710	44.057	1.008	0.012
9.00	3717.843	49.533	1.019	0.014

Table 3.5 – Numerical concentration values for the 10 µm polymer beads sample at various flow rates. The average values were normalised using the average concentration of all replicates at 5.96 and 9 µl/s.

When the exponential curve model is fitted to the 10 µm data, differences are visible in comparison with the 2 µm curve. The curve rises more slowly towards the plateau, indicating that larger particles require faster pump flow rates (Fig. 3.11).

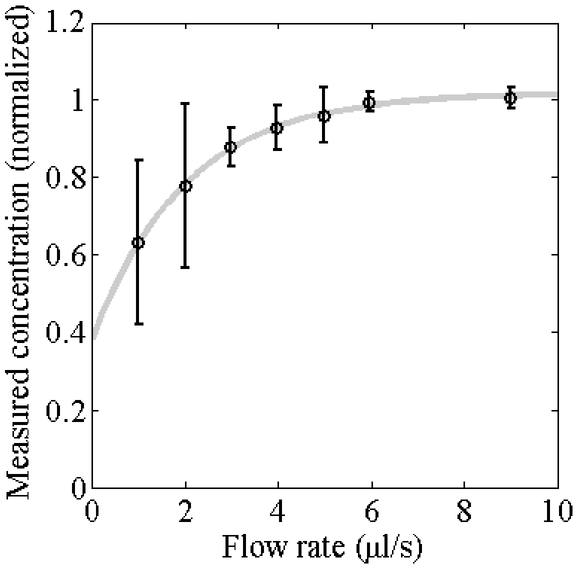


Figure 3.11 – Pump flow rate concentration correction curve ($y = ae^{-bx} + c$) for the 10 µm polymer beads normalised concentration values presented in Table 3.5. Curve parameters: a = -0.634, b = 0.496, c = 1.020.

As a result, the plateau value was obtained as the average concentration of all replicates at 5.96 and 9 µl/s pump flow rates only. Curve parameters: a = -0.634, b = 0.496, c = 1.020.

3.4.3.3 General correction curve

In a compromise between exactness and wider applicability of the method, the differences between the two exponential curve models were considered small enough for the two datasets to be merged and a single correction model to be consequently produced (Fig. 3.12), with curve parameters: a = -0.544, b = 0.623, c = 1.003.

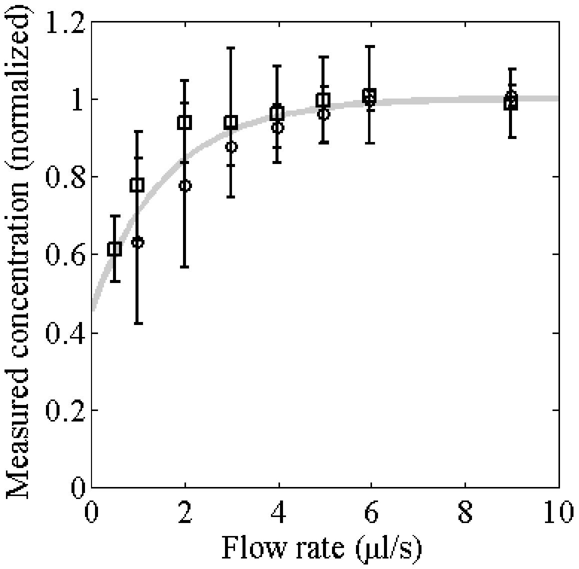


Figure 3.12 – Pump flow rate concentration correction curve ($y = ae^{-bx} + c$) for the combined 2 µm (squares) and 10 µm (circles) polymer beads normalised concentration values. Curve parameters: a = -0.544, b = 0.623, c = 1.003.

3.5 Conclusions

The CytoSense flow cytometer is found to slightly underestimate the flow rate of the sample pump, which in turn leads to an underestimation of the actual processed sample mass and, consequently, of the actual processed sample volume. It should be noted that the degree of underestimation (~5% for all flow rates) is negligible at low pump flow rates, and that the strongly linear relationship between actual and expected processed sample mass at various pump flow rates (adj. R-squared: 0.999, RMSE: 0.004) demonstrates that correct proportionality is preserved between different flow rate settings. This consistency is further demonstrated by the analysis of normalized particle concentrations in a dilution series of a 1 µm polymer beads suspension, which also produces highly linear results when compared with expected normalized concentrations (adj. R-squared: 0.996, RMSE: 0.023 for the 1:1 line). Particle concentration analysis in a 2 µm polymer beads suspension for which the particle concentration is known shows that the CytoSense consistently underestimates particle concentrations at all flow rates, and increasingly so at low pump flow rates. Particle concentrations measured by the instrument are found to plateau at concentration values on average 1.088 times lower than expected; this is interpreted as a systematic discrepancy, and the value kept as a correction factor to be applied at all flow rates in the remainder of this work. Coincidentally, this correction also accounts for the processed/analysed sample volume discrepancies as far as particle concentrations are concerned. Furthermore, the concentration underestimation vs. pump flow rate relationship is found to be well represented by an inverse exponential model ($y = ae^{-bx} + c$), which is used to develop a concentration correction curve. When correction curves are produced for both 2 µm and 10 µm polymer beads suspension, differences can be seen in the curve parameters which suggest some degree of size selection bias is taking place, likely due to the higher sinking speed of larger spheres; however, the differences are assumed small enough for a general correction curve to be calculated from a merger of the 2 µm and 10 µm datasets, with curve parameters: $a = -0.544$, $b = 0.623$, $c = 1.003$. This is of course a compromise, and in fact density as well as shape of the particles can also be expected to produce a bias in the measured concentrations. Nevertheless, the effect of particle size is the only one that could be assessed reliably, and the loss of exactness caused by the adoption of a generalised model is deemed an acceptable trade-off in exchange for a wider applicability of the method. As with the 1.088 plateau concentration ratio, the general correction curve is also used to correct flow-rate-induced concentration underestimations in the remainder of this work. In practice, the final expression for the particle concentration correction factor takes the form

$$x_{PUMP}(Q) = \frac{1.088}{y(Q)}, \tag{3.2}$$

where Q is the flow rate and y the value of the general correction curve at that flow rate.

4. Impact of particle size and RI distributions on the scattering properties of particle populations

In this chapter the Mie-based IOP forward scattering model presented in Chapter 2 is employed to explore the effects that spectral dependence of the complex refractive index and variations in the PSD have on the scattering behaviour of a particle population (within the frame of Mie theory) and on the λ-dependence of VSF and backscattering ratio $\tilde{b}_b$. The same features are also used to investigate the resilience of two frequently-used empirical approaches to the description of VSF and backscattering: the Fournier-Forand phase function approximation and the χ factor method for particle backscattering derivation.

4.1 Introduction

In Chapter 1 the VSF was shown to be one of the fundamental IOPs, as it describes the angular structure of the scattering, i.e. it provides information on which fractions of the incident light are scattered into each angle $\theta \in [0,\pi]$ (the scattering being assumed to be axially symmetric). Integration of the VSF over the appropriate angle ranges produces the scattering coefficient b and the backscattering coefficient $b_{b.}$ Measuring the VSF over the complete angular range is, however, technically challenging: this is due to the complexity of the instrument design required to make large numbers of discrete measurements over the broad angle range and the large dynamic range of the scattering signals (Sullivan & Twardowski, 2009).

A number of works over the years have faced this challenge using custom-made instruments (Jerlov, 1961; Kullenberg, 1968; Petzold, 1972; Lee & Lewis, 2003; Sullivan & Twardowski, 2009; Tan *et al.*, 2013) and while some VSF meters are now commercially available (e.g LISST-VSF, Sequoia Scientific, Inc., WA) it is still common practice to employ mathematical functions to analytically approximate the shape of the VSF, the most prominent of which is the Fournier-Forand phase function (Fournier & Forand, 1994; Mobley *et al.*, 2002). Other approaches sidestep the necessity of measuring the entirety of the VSF, for example by deriving the backscattering coefficient from a single, broad angle measurement using a coefficient known as the χ factor (Maffione & Dana, 1997; Boss & Pegau, 2001).

PSDs are especially tied to the VSF in the sense that efforts can be made to derive one from the other and vice versa. On one hand, recent works have described ways to invert volume scattering functions to reconstruct particle size distributions (Zhang *et al*., 2011; Zhang *et al*., 2014); on the other hand, all approaches to approximate the VSF have to include some definition of the properties of the PSD in their procedure. In Chapter 1 the power law or Junge distribution was introduced as the most common approximation for the PSD of mixed marine particle populations. Ulloa *et al*. (1994) showed that this kind of distribution, when matched with spectrally independent RIs, has the effect of making the backscattering ratio λ-independent, which in terms of VSF means that the angular structure of the scattering, represented by the SPF of eq. (1.24), remains unchanged across visible wavelengths.

However, this raises the question of just how sensitive this potentially helpful outcome is to deviations from the underlying assumptions of a Junge particle size distribution and a spectrally flat complex refractive index. The Mie-based IOP forward scattering model presented in Chapter 2 was therefore employed to assess the effect that a series of perturbations to these standard assumptions has on the modelled IOPs. This was done firstly by introducing spectral dependence of the complex refractive index and secondly by adding a series of log-normal curves to the power law distribution. These log-normal perturbations are intended to reproduce sub-populations with varying average size (by controlling the mode of the distribution), number of individuals (by changing the peak height of the distribution) and size variability (by controlling the scale width parameter of the distribution), broadly representative of events such as phytoplankton blooms and plumes of sediment. The effectiveness of the Fournier-Forand phase function in reproducing the shape of the scattering phase function (SPF) and the resilience of the χ factor were also both tested using the same set of log-normal variations.

Note: the backscattering ratio is defined as a percentage of the scattering coefficient; unless explicitly stated, a reported increase or decrease in its value is meant to be interpreted as such, and *not* compounded as a percentage increase or decrease of the value itself (e.g. a 2% increase of $\tilde{b}_b = 1\%$ is to be understood as $\tilde{b}_b = 3\%$, not as $\tilde{b}_b = 1.02\%$).

4.2 Theory

4.2.1 *Fournier-Forand phase function*

The Fournier-Forand phase function (FFPF) is an analytic function made to approximate the SPF of a power law size distribution of suspended particulate, specifically conceived to replicate the SPF given in the landmark Petzold (1972) paper. It is useful in situations where technical limitations impede the direct determination of the SPF over the entire angular range. This study adapts the equations and methodology given by Mobley *et al.* (2002), which built upon the original work (Fournier & Forand, 1994) by parametrizing it on backscattering ratio $\tilde{b}_b$ alone. In this formulation, the FFPF is given by

$$\begin{aligned}\tilde{\beta}_{FF}(\theta) = \frac{1}{4\pi(1-\delta)^2\delta^\nu}\Big\{&\nu(1-\delta) - (1-\delta^\nu) \\ &+ [\delta(1-\delta^\nu) - \nu(1-\delta)]\sin^{-2}\left(\frac{\theta}{2}\right)\Big\} \\ &+ \frac{1-\delta_{180}^\nu}{16\pi(\delta_{180}-1)\delta_{180}^\nu}(3\cos^2\theta - 1)\,,\end{aligned} \tag{4.1}$$

where

$$\nu = \frac{3-\mu}{2} \tag{4.2}$$

and

$$\delta(\theta) = \frac{4}{3(n-1)^2}\sin^2\left(\frac{\theta}{2}\right). \tag{4.3}$$

Here n is the real part of the refractive index and μ the slope of the power law distribution.

Equation (4.1) can be integrated to get the backscattering ratio

$$\tilde{b}_b = \left(-\frac{1}{2}\right)\frac{1-\delta_{90}^{\nu}}{(1-\delta_{90})\delta_{90}^{\nu}}, \qquad (4.4)$$

but no unique pair of n and μ exists for a given $\tilde{b}_b$ value: to proceed forward the procedure would need ancillary measurements which are most often problematic. To circumvent this, a linear relationship is postulated between the two variables (Mobley *et al*., 2002),

$$n = 1.01 + 0.1542(\mu - 3) \qquad (4.5)$$

Equations (4.4-4.5) can then be used to produce a lookup table of backscattering ratio values. The table is used to find the closest match with the backscattering ratio modelled from Mie theory, and the corresponding n and μ values are used to produce $\tilde{\beta}_{FF}(\theta)$.

4.2.2 χ factor

The technical challenges posed by VSF measurements have the effect of making backscattering coefficient measurements equally difficult. An alternative approach to the problem is to estimate the backscattering coefficient by measuring scattering $\beta(\theta)$ at a single angle or small angular range and multiplying the value by a constant factor χ. This method revolves around the idea of a special angle for which scattering $\beta(\theta)$ has a linear relationship with b_b irrespective of the properties of the particle population (Oishi, 1990; Maffione & Dana, 1997; Boss & Pegau, 2001).

Using the χ factor, b_b is obtained through

$$b_b = 2\pi\chi(\theta)\beta(\theta)\,. \qquad (4.6)$$

Since the scattering is partly due to particles and partly due to water, the χ factor is similarly divided to give

$$\chi(\theta) = \chi_w(\theta)y + \chi_p(\theta)(1-y), \tag{4.7}$$

where $y \in [0,1]$ represents the fraction of scattering due to water. It follows that the optimal angle for backscattering estimation is the angle where $\chi(\theta) = \chi_w(\theta) = \chi_p(\theta)$, because at this angle there is no need to know the relative contributions of water and particles to the overall scattering. The approach was found by Boss & Pegau (2001) to work best for $\theta = 117° \pm 3°$, with a reference χ factor value given as $\chi_p(117°) = 1.1 \pm 4\%$. A more recent work by Sullivan & Twardowski (2009) further reports the χ factor approach to work best between 110° and 120°, with values respectively $\chi_p(110°) = 1.000 \pm 0.026$ and $\chi_p(120°) = 1.097 \pm 0.032$ compatible with previous sources.

4.3 Methods

4.3.1 Spectral dependence of the complex relative refractive index

A first set of perturbations to the spectral independence of the backscattering ratio was introduced in the form of λ-dependence in the complex refractive index, given relative to water ($n_r = 1.333$). The λ-dependent RIs were applied to a single power law distribution with slope value $\gamma = 4$. Typical spectral values for the real component of the refractive index of organic matter were estimated from Fig. 2 (lower panel) of Stramski *et al.* (1988). No useful literature including spectral rRI information for mineral particles was found, and the inorganic real refractive index was set at a fixed value of 1.15 (cf. Tab 1.1). Using the same procedure, typical values for the imaginary component of the relative refractive index were estimated from Fig. 8 of Babin *et al.* (2003), for both organic and inorganic particles. Variability for both the real and imaginary components of the refractive index was considered simultaneously, so that for this sensitivity test a combined, λ-dependent complex relative refractive index was fed into the model (Tab. 4.1, Fig. 4.1).

λ (nm)	400	420	440	460	480	500	520	540
n_r (org.)	1.047	1.047	1.050	1.052	1.052	1.053	1.053	1.051
n_r (inorg.)	1.150	1.150	1.150	1.150	1.150	1.150	1.150	1.150
n_i (org.)	0.0022	0.002	0.0026	0.0023	0.0019	0.0014	0.001	0.0008
n_i (inorg.)	0.006	0.0042	0.003	0.0024	0.0019	0.0015	0.0012	0.001

λ (nm)	560	580	600	620	640	660	680	700
n_r (org.)	1.050	1.050	1.050	1.050	1.053	1.053	1.053	1.055
n_r (inorg.)	1.150	1.150	1.150	1.150	1.150	1.150	1.150	1.150
n_i (org.)	0.0007	0.0006	0.0005	0.0004	0.0005	0.0009	0.0018	0.0009
n_i (inorg.)	0.0008	0.0006	0.0005	0.0003	0.0003	0.0002	0.0002	0.0001

Table 4.1 – Spectral values for the real and imaginary parts of the refractive index as adapted from literature – Stramski *et al*. (1988), Fig. 2, lower panel, and Babin *et al*. (2003), Fig. 8. The inorganic real refractive index was set at a fixed value of 1.15.

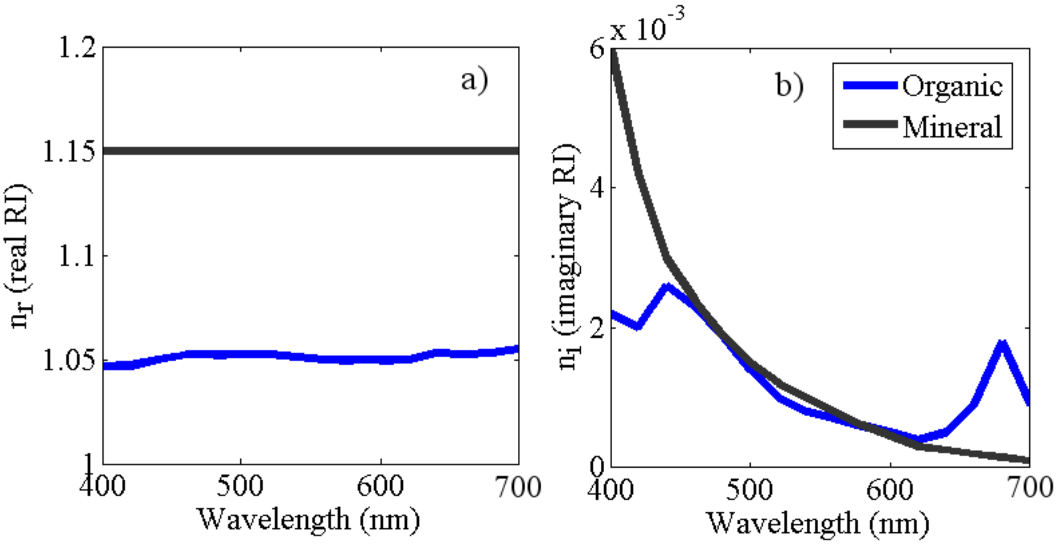


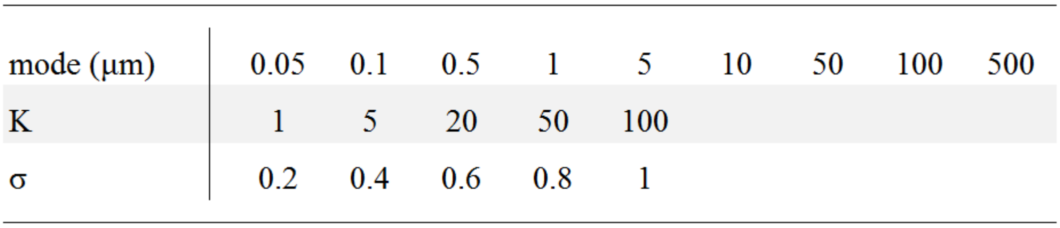
Figure 4.1 – The profiles for the organic and inorganic spectral values of (a) the real part of the refractive index and (b) the imaginary part of the refractive index presented in Tab. 4.1.

4.3.2 *Log-normal variations in the PSD*

The second set of perturbations to the spectral independence of the backscattering ratio was introduced with the addition of log-normally distributed populations to the classic power law PSD, as defined in eq. (1.36). A baseline population was represented with a fixed power law distribution with slope $\gamma = 4$ and particle diameters ranging between 10^{-8} and 10^{-4} m. The log-normal deviations were made to vary along three parameters: the mode, the scale width parameter σ and a scaling population multiplier K (Tab 4.2), defined so that the resulting log-normal distribution peaks at K times the corresponding N(mode) value in the baseline power law distribution. The mode of a log-normal is tied to the other parameters of the distribution by

$$mode \ = e^{(\mu-\sigma^2)} \,. \qquad (4.8)$$

The effects of increasing values for K (with fixed σ = 0.1) and of increasing scale width σ (with fixed K = 1) at various points along the power law distribution were considered separately (Fig. 4.2).



mode (µm)	0.05	0.1	0.5	1	5	10	50	100	500
K	1	5	20	50	100				
σ	0.2	0.4	0.6	0.8	1				

Table 4.2 – The values of the mode, population multiplier K and scale width parameter σ used in the test of sensitivity to log-normal perturbations to the power law PSD.

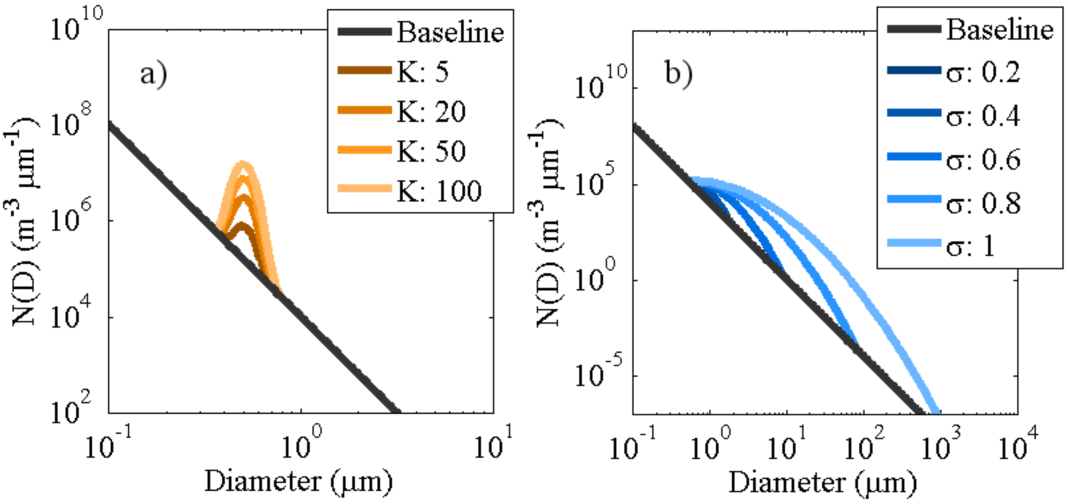


Figure 4.2 – Log-normal variations corresponding to the values of (a) population multiplier *K* and (b) width parameter *σ* presented in Tab. 4.2 (for clarity, curves are displayed for *mode* = 0.5 µm only).

4.3.3 *Quantification of λ-dependence*

The spectral backscattering ratio resulting from both types of perturbations is the outcome of the effects induced by the deviations on the angular structure of the VSF. While a purely power-law-like distribution with spectrally flat RIs has the same angular structure at all wavelengths, it is generally not so for the modified PSDs (Fig. 4.3) and/or when the complex RI is λ-dependent.

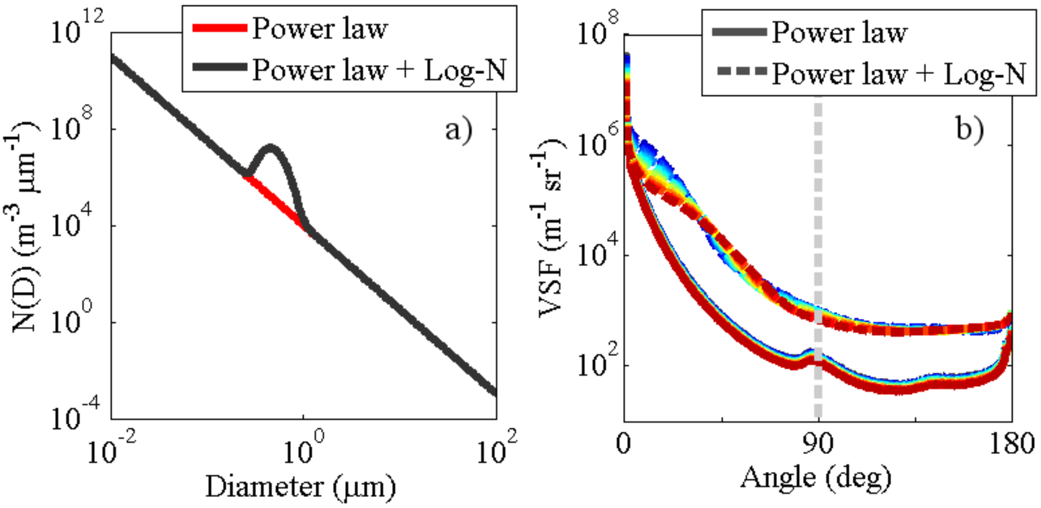


Figure 4.3 – (a) An example of log-normal distribution added to a power law size distribution and (b) the effect it has on VSF at various wavelengths (blue to dark red, going from 400 to 700 nm).

All resulting spectral backscattering ratios were fitted linearly against λ, using the slope of the fit line as a measure of λ-dependence (Fig. 4.4). For the sake of simplicity, all slopes are given using microns as the unit measure of wavelength: under this convention a slope value m = 1 corresponds to a rate of change of 1% $\tilde{b}_b$ per μm, which translates to 0.3% over the 400-700 nm wavelength range here considered. Note that although all observed spectral dependencies are mostly linear with λ, the degree to which they agree with this simple model varies; the linear best fit model was chosen as a slightly crude but rather convenient way to characterise the broader aspect of the λ-dependence.

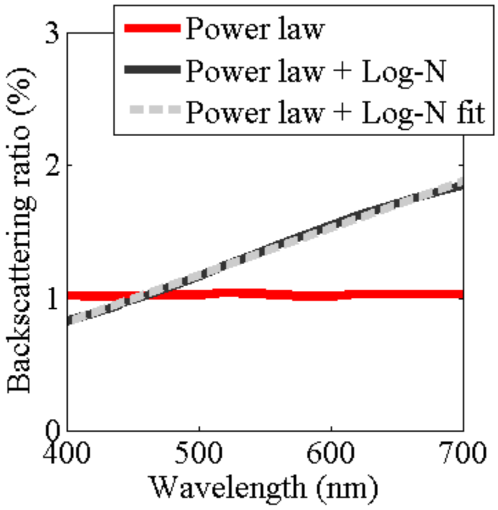


Figure 4.4 – The effect on the λ-dependence of the backscattering ratio for the same distributions presented in Fig. 4.3. The slope of the linear fit is used as a measure of the λ-dependence in terms of rate of change of the backscattering value over the 400-700 nm wavelength range considered.

4.3.4 FFPF calculations

Following eqs. (4.4-4.5) a lookup table of backscattering ratio values was generated starting from an array of 1000 μ values, linearly spaced between $\mu = 3$ and $\mu = 5$. The $\tilde{b}_b$ values provided by the forward scattering model were then compared with the values within the lookup table to find the closest match and produce the corresponding $\tilde{\beta}_{FF}(\theta)$. The FFPF approximation was tested by applying the same PSD deviations described in Paragraph 4.3.2. For each of the perturbations described and each of the 16 linearly spaced wavelengths between 400 and 700 nm a FFPF and a SPF were produced, for a total of 2460 significant FFPF-SPF pairs. The quality of the FFPF approximation was assessed by calculating a $\tilde{\beta}_{FF}(\theta)/\tilde{\beta}(\theta)$ ratio for each of these pairs over the whole angular range, providing an indication of the agreement between the two for each individual angle.

4.3.5 χ factor calculations

$\chi_p(\theta)$ was calculated using eq. (4.7) over the entire angular range for each of the PSD deviations considered in paragraph 4.3.2. Average $\chi_p(\theta)$ and its variance were determined for each individual angle across all cases to locate the angular neighbourhoods of major stability, with particular attention to the $\sim 117°$ area for which no information on the relative contribution of water to the scattering is needed (Boss & Pegau, 2001).

4.4 Results

4.4.1 IOP forward modelling validation

The IOP forward modelling used in this case-study follows the lines described in Section 2.2 of Chapter 2 and builds on the Mie theory foundations provided in Chapter 2 in general. To validate its correct functioning the model was employed to reproduce the results given by Ulloa *et al.* (1994). Three Junge distributions used in their work were replicated, with slope value $\gamma = 4$ and the real part of the complex refractive index set to $n_r = 1.04$, $n_r = 1.05$ and $n_r = 1.06$ respectively (all real refractive indices values in this chapter are given relative to water). The imaginary part was set to $n_i = 0.001$. The particle distributions thus produced were fed into the model and the results compared with those presented by Ulloa *et al.*; all three backscattering ratios were found to be λ-independent and to follow closely the results of the paper (Fig. 4.5a). It is interesting to note that although the angular distribution of the scattering (the SPF) is λ-independent under these conditions, the magnitude of the scattering (and therefore the VSF) is not (Fig. 4.5b).

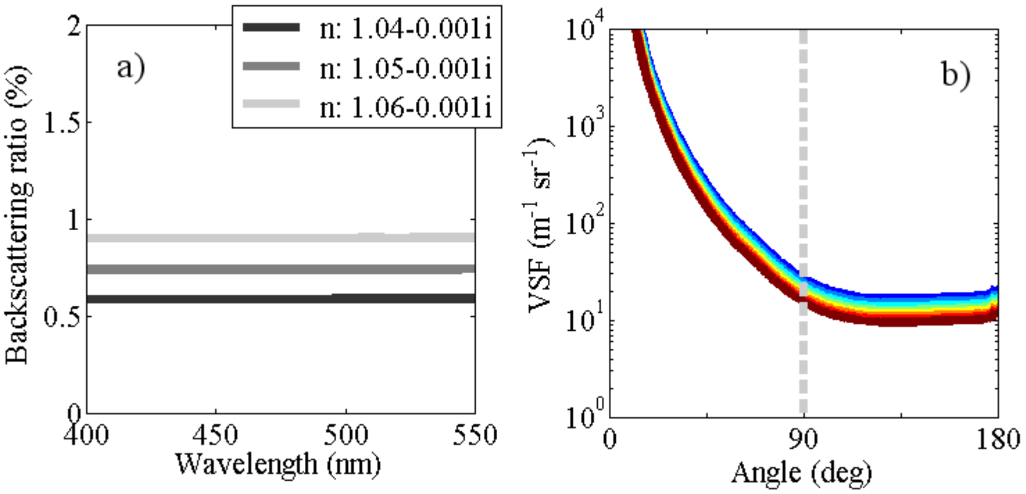


Figure 4.5 – (a) Wavelength-independence in the backscattering ratio for a power law PSD and fixed refractive indices. The scattering model successfully reproduces the results presented in Fig. 4 of Ulloa *et al*. (1994) when the same parameters are used. (b) Detail of the VSF for *n* = 1.05-0.001i. The backscattering ratio is spectrally flat because the angular distribution of the scattering is the same at all wavelengths (blue to dark red, going from 400 to 700 nm); the magnitude of the scattering, however, is *not* constant between wavelengths.

4.4.2 Spectral dependence of VSF and backscattering ratio

4.4.2.1 Effects of complex refractive index λ-dependence

Spectral dependence in the complex relative refractive index induced a range of variability of 0.145% (2.702-2.847%) over the 400-700 nm wavelength range in the backscattering ratio of inorganic-like particles, as opposed to a range of variability of 0.111% (0.709-0.820%) for organic-like particles (Fig. 4.6). The linear best-fit of the backscattering ratio produced slope values $\gamma = 0.141$ and $\gamma = 0.475$ for organic- and inorganic-like particles respectively, with corresponding R-squared values 0.147 and 0.932. This corresponds to an average rate of change of 0.042% and 0.280% respectively over the 400-700 nm range. The small R-squared value obtained for organic-like particles indicates that variability is dominated by oscillations around the mean. Overall it can be concluded that λ-dependent complex relative refractive indices did cause some variability in the backscattering ratio, but of small magnitude.

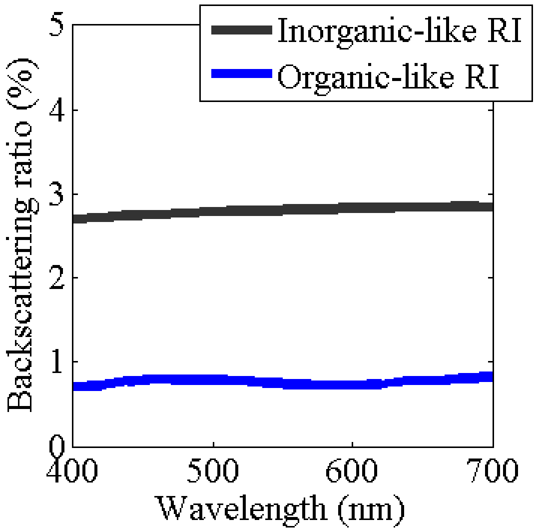


Figure 4.6 - Effect on the λ-dependence of the backscattering ratio of the introduction of λ-dependent complex refractive indices in a power law PSD (slope value γ = 4), for both organic-like and inorganic-like refractive indices.

4.4.2.2 Effects of log-normal variations in the PSD

When the height of the added log-normal populations (the population multiplier K) was made to vary, substantial variability was seen for distribution modes ranging from 0.05 to 1 μm for organic-like particles, and up to 10 μm for inorganic-like ones (Fig. 4.7a-b). Using λ = 550 nm as sample wavelength, the backscattering ratio was generally seen to increase with increasing log-normal height for particles smaller than 1 μm, and to decrease (or remain largely constant) for larger particles. In the most extreme case observed (mode = 0.1 μm, K = 100) the linear fit across all wavelengths provided slope values $m = -18.660$ for organic-like particles and $m = -29.720$ for inorganic-like ones, corresponding to a rate of change of -5.598% and -8.916% respectively over the 400-700 nm wavelength range (Fig. 4.7c-d). The sign of the slope was also found to switch between 0.1 and 0.5 μm, changing from a decreasing trend to an increasing one from blue to red wavelengths. In general, organic-like particles mirrored the behaviour of inorganic-like ones, but with smaller absolute values for both $\tilde{b}_b$ and $\tilde{b}_b$ slope. Overall, in both cases the largest effects were found for the smallest particles, and as a general rule the magnitude of the spectral slope increased as the height of the distribution increased, with slope direction varying in a progressive manner with mode value.

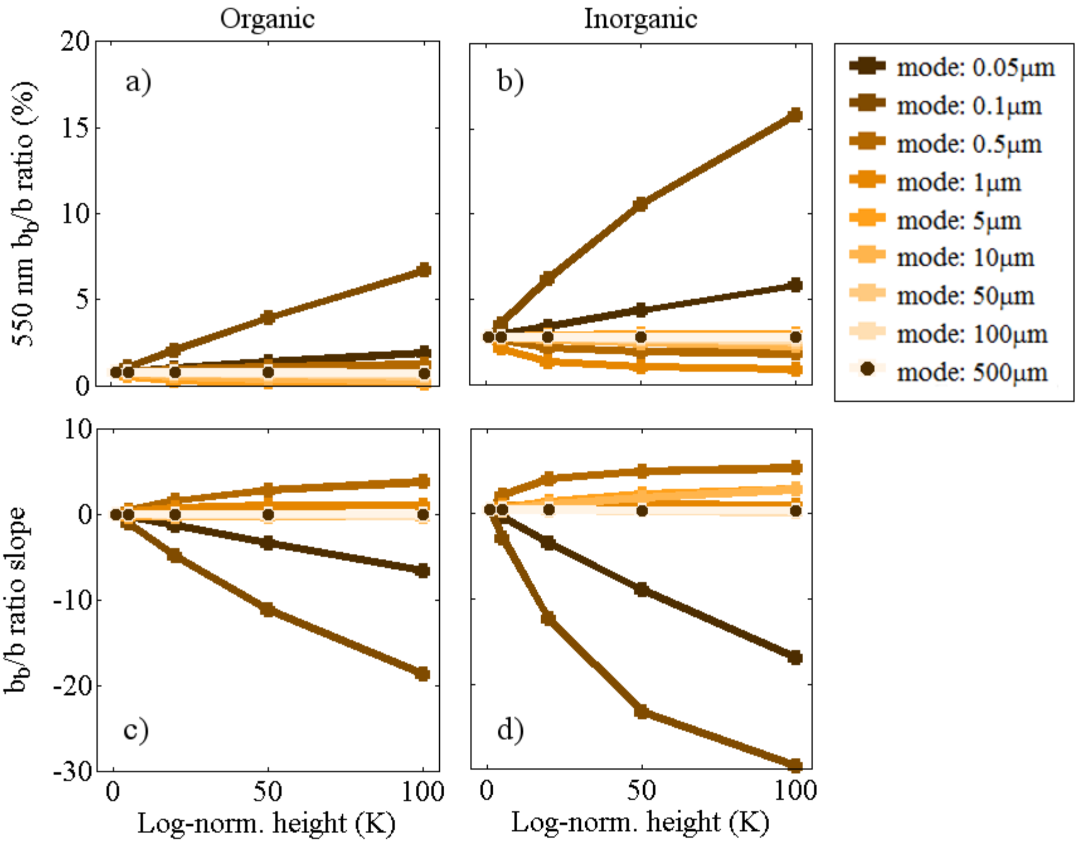


Figure 4.7 – Resulting $\tilde{b}_b$ values at the $\lambda = 550$nm sample wavelength due to log-normal deviations made to vary along the height (K), for (a) organic- and (b) inorganic-like particles. In the bottom row, resulting slopes of the linear fit of the spectral $\tilde{b}_b$ values due to the same log-normal deviations, for (c) organic- and (d) inorganic-like particles.

When the width of the log-normal (the scale width parameter σ) was made to vary, large deviations were seen to correspond to distribution modes ranging from 0.05 to 0.1 μm for organic- and up to 5 μm for inorganic-like particles. Using λ = 550 nm as sample wavelength, $\tilde{b}_b$ was generally seen to irregularly decrease across all modes, with only 0.05 and 0.1 μm producing an initial increase before settling to lower $\tilde{b}_b$ values as well. As it was the case for log-normal height variations, organic-like particles mirrored the behaviour of inorganic-like ones, but with smaller absolute values (Fig. 4.8a-b). The behaviour of the $\tilde{b}_b$ slope was found to be non-uniform; the slope for mode 0.05 μm becomes negative at first, before quickly growing highly positive and finally settling close to zero. The slope for mode 0.1 µm steadily grows at first, but eventually settles towards zero as well (Fig. 4.8c-d). For organic-like particles, all other cases remain largely constant and close to zero (Fig. 4.8c); inorganic-like particles continue to show a variety of behaviours instead, with slopes irregularly increasing to values above zero for particles between modes 0.5 and 10 µm, slightly decreasing for modes 50 and 100 µm, and constant at zero for 500 µm only (Fig. 4.8d).

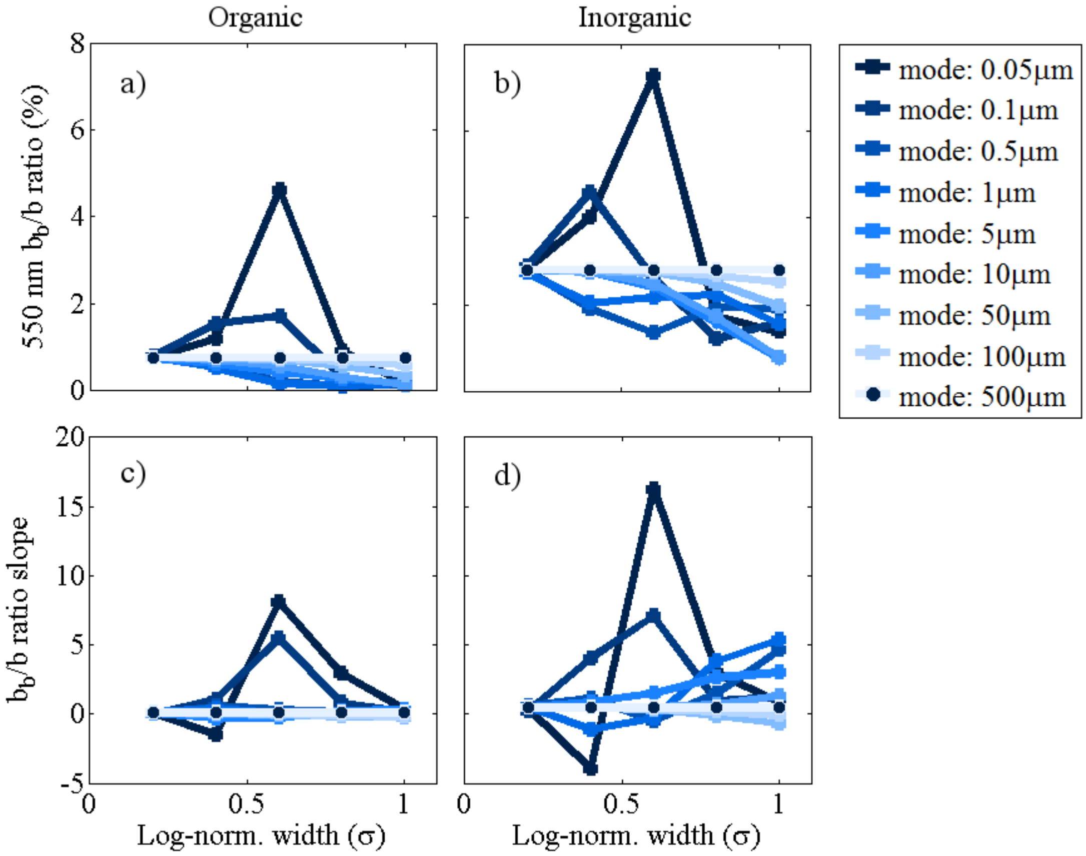


Figure 4.8 – Resulting $\tilde{b}_b$ values at the $\lambda = 550$nm sample wavelength due to log-normal deviations made to vary along the width (σ), for (a) organic- and (b) inorganic-like particles. In the bottom row, resulting slopes of the linear fit of the spectral $\tilde{b}_b$ values due to the same log-normal deviations, for (c) organic- and (d) inorganic-like particles.

The most extreme case corresponded to mode 0.05 µm, with $\tilde{b}_b$ slope ranging from a minimum value $m = -1.501$ $(m = -4.003)$ to a maximum value $m = 8.090$ $(m = 16.230)$ for organic-like (inorganic-like) particles, corresponding to $\sigma = 0.4$ and $\sigma = 0.6$ ($\sigma = 0.4$ and $\sigma = 0.6$ as well). These largest deviations (*mode* = 0.05

μm, $\sigma = 0.6$) respectively result in a rate of change of 2.427% and 4.869% in $\tilde{b}_b$ between 400 and 700 nm.

4.4.3 Resilience of the FFPF

The FFPF script was first of all tested in its capability to emulate the results presented in Mobley *et al.* (2002). The backscattering ratio value given for Petzold's data ($\tilde{b}_b = 0.0183$) provided values for slope and refractive index $\mu = 3.5826$ and $n = 1.0998$, respectively differing by 0.018% and 0.025% from the $\mu = 3.5835$ and $n = 1.10$ values given by Mobley *et al.* (2002). The values thus obtained were used to produce a corresponding $\tilde{\beta}_{FF}(\theta)$, which was compared with Petzold's SPF: the FFPF produced by the script successfully matched Petzold's data with an average FFPF/SPF ratio of 0.99 ± 0.10 across the angular range (Fig. 4.9).

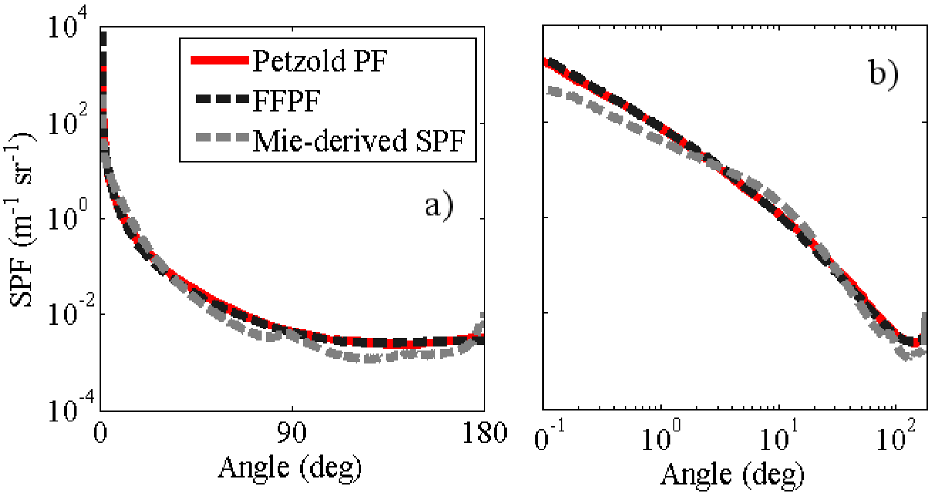


Figure 4.9 – Validation of the FFPF script and comparison with the SPF produced for the same parameters by Mie-based forward modelling in (a) log-linear and (b) log-log scales.

After validating the FFPF script, the resilience of the FFPF approximation was tested against the set of log-normal variations in the PSD, producing a variety of results (Fig. 4.10). Overall, as indicated by each $\tilde{\beta}_{FF}(\theta)/\tilde{\beta}(\theta)$ ratio, the FFPF approximated the SPF to within ±10% of its value ($0.9 < \tilde{\beta}_{FF}(\theta)/\tilde{\beta}(\theta) < 1.1$) in less than half of the 2460 significant cases for the angle range $0° - 70°$, and in less than a fifth of them for the angle range $0° - 60°$ (except for two local maxima around 1.7° and 5.9°

for organic-like and inorganic-like particles respectively). At the opposite end of the angle range, the FFPF approximated the SPF within 10% of its value in less than a third of the significant cases for the angle range $175° - 180°$. However the FFPF approximation fared considerably better when mid-backwards angles were considered, roughly in the range $90° - 165°$. Three maxima were found at $99°$, $120°$ and $162°$. At the overall maximum, located at $120°$, the FFPF approximated the SPF to within $\pm10\%$ of its value in 92.561% of the 2460 significant cases, conforming to the idea that this angular neighbourhood is one of particular stability in terms of angular structure of the VSF, even in the case of major deviations from a power law distribution (Fig. 4.10).

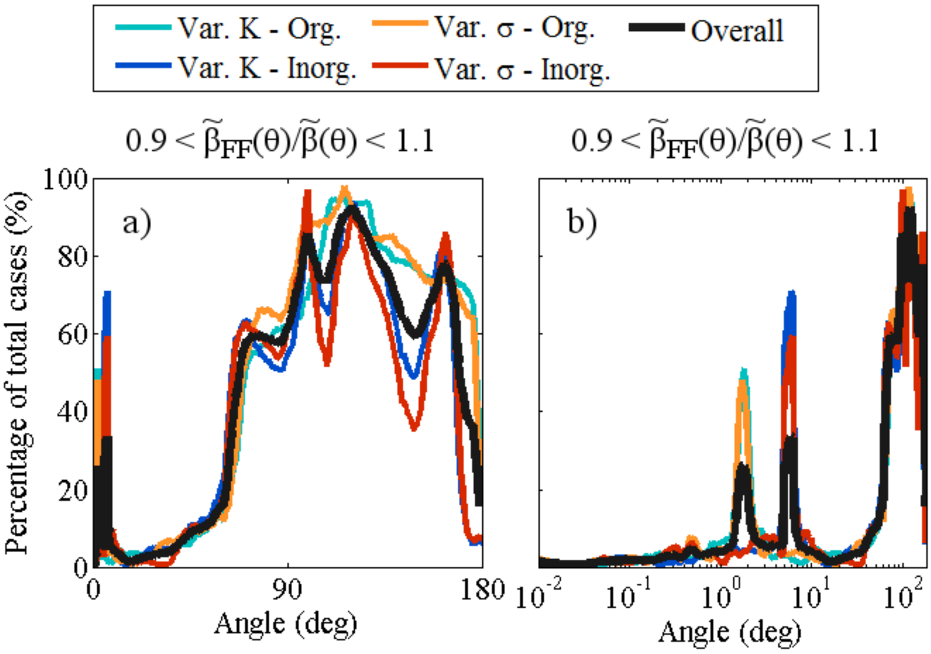


Figure 4.10 – Angular percentage of all significant cases considered for which the FFPF approximates the corresponding Mie-based SPF to within $\pm10\%$ of its value ($0.9 < \tilde{\beta}_{FF}(\theta)/\tilde{\beta}(\theta) < 1.1$), both in (a) linear and (b) logarithmic angular scale. The FFPF fares poorly at forward and narrow-backward angles, but presents three compatibility maxima at $99°$, $120°$ and $162°$. At the overall maximum ($120°$) the FFPF approximated the SPF to within $\pm10\%$ of its value in more than 90% of the cases.

When considering organic- and inorganic-like particles separately, the former were found to be better approximated by the FFPF as reflected by a general broadening of the maxima and a resulting higher rate of approximation within 10% across the

$90° - 165°$ range. When considering K and σ perturbations separately, the former were found to be better approximated by the FFPF across most of the entire angle range for both organic- and inorganic-like particles. Singularly, the σ perturbations in the inorganic case produced their global compatibility maximum at $99°$ instead of $120°$.

4.4.4 *Resilience of the χ factor*

The reference value for the χ factor is given as $\chi_{BP}(117°) = 1.100 \pm 0.044$ by Boss & Pegau (2001). A more conservative value by Oishi (1990) (when converted into an χ factor as per the definition given by Boss & Pegau) is given as $\chi_{Oishi}(120°) = 1.140 \pm 0.114$, with a corresponding $\chi_{BP}(120°) = 1.120 \pm 0.047$ and $\chi_{ST}(120°) = 1.097 \pm 0.032$ (Sullivan & Twardowski, 2009). In our results (Tab. 4.3), varying the height of the log-normal distributions produced $\chi_p(117°) = 1.052 \pm 0.056$ and $\chi_p(117°) = 1.090 \pm 0.046$ for inorganic- and organic-like particles respectively.

	Organic				Inorganic				Total	
	Height (K)		Width (σ)		Height (K)		Width (σ)			
	Mean	St. D.	Mean	St. D.	Mean	St. D.	Mean	St. D.	Mean	St. D.
$\chi_p(117°)$	1.090	0.046	1.106	0.053	1.052	0.056	1.046	0.069	1.074	0.062
$\chi_p(120°)$	1.105	0.055	1.143	0.085	1.102	0.058	1.117	0.086	1.118	0.076

Table 4.3 – $\chi_p(117°)$ and $\chi_p(120°)$ values obtained from the 2460 log-normal perturbations added to the power law PSD, given individually by varying height or varying width for both organic- and inorganic-like particles and together as the overall total.

Varying the width instead produced $\chi_p(117°) = 1.046 \pm 0.069$ and $\chi_p(117°) = 1.106 \pm 0.053$, again for inorganic- and organic-like particles respectively. In both cases only organic-like particles produced χ factors compatible with $\chi_{BP}(117°)$; this is likely to be a reflection of the broader compatibility maxima found in the FFPF approximation for the organic case. As a result, when the overall χ factor for all 2460 significant cases is considered for θ = 117° ($\chi_p(117°) = 1.074 \pm 0.062$) (Fig. 4.11a) compatibility with $\chi_{BP}(117°)$ can be rejected at a 5% confidence level using an unequal variances T test ($t = 3.852$, $\nu = 46.77$). When all 2460 significant cases are considered for θ = 120° instead ($\chi_p(120°) = 1.118 \pm 0.076$) (Fig. 4.11b) compatibility with $\chi_{BP}(120°)$ cannot be rejected at a 5% confidence level ($t = 0.276$, $\nu = 47.11$), although compatibility with $\chi_{ST}(120°)$ can. Even then, values for $\chi_p(120°)$ and $\chi_{ST}(120°)$ are found to differ by 1.88%. When θ = 117° is considered, 84% (93%) of all significant cases are found to lie within one (two) standard deviation(s) from the mean, while if θ = 120° is considered, 87.7% (95%) of all significant cases are found to lie within one (two) standard deviation(s) from the mean. Standard deviation values are 5.8% and 6.8% of the mean for θ = 117° and θ = 120° respectively.

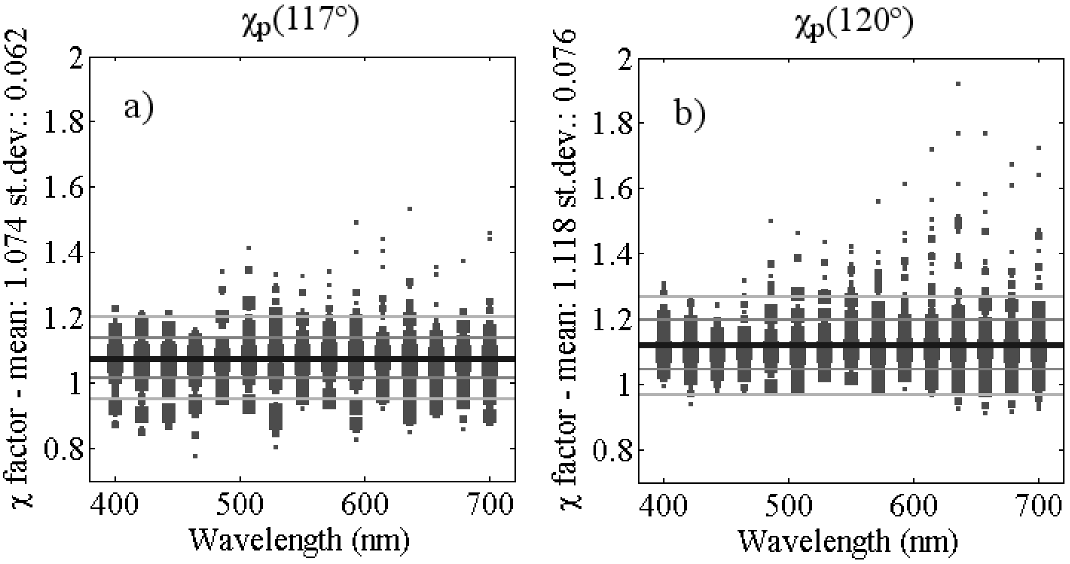


Figure 4.11 – Individual and overall values of (a) $\chi_p(117°)$ and (b) $\chi_p(120°)$. Each marker represents one of the cases, with smaller marks representing more extreme (less likely) deviations. The black, grey and light grey lines represent overall mean, standard deviation and 2x standard deviation of the χ_p values.

4.5 Conclusions

Approximating the suspended particle population to a power law size distribution of spherical, homogeneous particles with a spectrally independent complex refractive index is common practice in optical oceanography, and many of the results on which modern backscattering fraction measurement techniques are based rely on these assumptions. There is conflicting evidence in the literature on spectral dependence or independence of the backscattering ratio (and VSF). The aim of this chapter has been to perform a sensitivity analysis that could reveal potential sources of λ-dependence and circumstances that could explain observations from nature.

While the effect that a spectrally dependent complex refractive index has on the wavelength independence of the backscattering ratio is small, perturbations in the form of local additions of log-normal populations to the power law were found to induce sometimes large λ-dependence in both VSF and backscattering ratio. Of course, as a sensitivity study, the extreme perturbations presented here may be outside of those likely to be encountered in nature. However, they serve a useful purpose in exposing underlying trends and physical mechanisms that could explain observations of variable degrees of λ-dependence. These results demonstrate that the development of a reliable method for the routine determination of the PSDs of marine particle populations remains an important goal in the field of optical oceanography. Marked differences were also found in the results between organic- and inorganic-like particles; while the case of mixed populations was outside the scope of this sensitivity study it is to be expected that a mixed particle population would have a non-negligible effect on VSF and backscattering ratio. Overall, even though technically challenging, efforts towards a method of direct determination of the particle refractive index distribution remain essential to a complete and detailed description of the IOPs of water.

The Mobley *et al*. (2002) parameterisation of the Fournier-Forand phase function was found unable to reproduce all the features generated on the SPF by the PSD perturbations, especially in the forward and narrow backward direction; the angular neighbourhood around $\theta = 120°$ was however confirmed to be one of particular stability for the scattering angular structure. In this area the otherwise poorly performing FFPF was found to approximate the value of the SPF within ±10% of its value in more than 90% of the cases considered.

Conforming to this, the χ factor method for backscattering coefficient determination is found to be a robust one. Results indicate that the refractive index composition of the particle population has important consequences on the choice of the individual angle where the χ factor is measured, and suggest that picking an angle at the upper end of the range of validity suggested by both Boss & Pegau (2001) and Sullivan & Twardowski (2009) (~120°) may improve the resilience of the method. Even for $\theta = 120°$, some individual χ factor values deviate markedly from the average; yet, they generally correspond to log-normal additions to the PSD which are extreme and therefore unlikely. A large majority of the cases here considered produce χ factor values which are close to expectations, with a correspondingly small standard deviation and an overall average $\chi_p(120°)$ value compatible with existing literature.

5. Datasets

5.1 Introduction

This short chapter will present the datasets used to produce the results on which the second half of this thesis work will focus. The goal is to provide a convenient reference to inform the reading and comprehension of those results by collating all information pertaining to the available datasets and to the protocols used to retrieve them within a single, separate chapter. The chapter will be divided in two main parts, corresponding to the two main datasets used in this work: the algal cultures (AC) dataset and the UK coastal waters (UKCW) dataset. For each of these, flow cytometry measurement protocols will be described first. All other relevant data (such as IOP, biogeochemistry and independent PSD measurements) will then be presented along with a brief description of their working principles and their respective measurement protocols.

5.2 Algal cultures

The algal cultures dataset (AC) consists of phytoplankton culture samples selected as part of a separate study on toxic and nuisance species, and includes the following: *Microcystis aeruginosa* (MS), *Synechococcus* sp. (SC), *Heterocapsa* sp. (HS), *Alexandrium minutum* (AM), *Karenia mikimotoi* (KM), *Pseudo-nitzschia seriata* (PS), *Skeletonema marinoi* (SM). All cultures were provided by the Culture Collection of Algae and Protozoa (CCAP) of the Scottish Association for Marine Science (SAMS). The cultures were measured on two separate occasions: 10th-20th of June 2014 and 25th-26th of January 2016; all were allowed to reach stationary phase before measurement, and were in many cases in decline by the time the measurements were made. The highly concentrated culture samples were first diluted in ~ 15 litres of filtered sea water within a large black container to minimize light reflection and diffusion from the container walls and ensure best conditions for attenuation, absorption and backscattering measurements. These were carried out using an ac-9 absorption and attenuation meter and a BB9 backscattering meter

(WET Labs). Subsequently, the diluted samples were analysed using the CytoSense flow cytometer.

5.2.1 *Flow cytometry measurement protocol*

Samples processed in June 2014 (HS, AM, PS, SM) were measured in triplicates at a single side scattering PMT sensitivity setting (56) for 3 minutes at a flow rate of 0.98 µL/s, while samples processed in January 2016 (MS, SC, KM) were measured a single time for each of four sensitivity settings of the side scattering PMT (50, 60, 70, 80) for 6 minutes at a flow rate of 0.96 µL/s. Side scattering was used as the trigger channel in all cases. Additional standard polymer beads measurements necessary for FC method calibration (as will be described in Chapter 6) were taken twice on a weekly basis for the June 2014 sampling period and once at the beginning of the January 2016 sampling period.

5.2.2 *Attenuation and absorption measurements*

5.2.2.1 Absorption and attenuation meter working principle

The WET Labs ac-9 instrument is a submersible, *in situ* absorption and attenuation meter designed to provide real-time measurements over nine individual wavelength channels; a detailed description of the instrument is given by Twardowski *et al*. (1999). In brief, the filters which select for the nine wavelengths are installed on a rotating mount and intercept the light beam generated by the instrument, which is then split into a reference beam and a primary beam. The latter is made to pass through the sample, and the amount of photons lost during transit is then determined from the negative natural logarithm of transmission (given as the ratio of the intensities of the reference and transmitted primary beams) divided by optical pathlength. The ac-9 absorption tube uses total internal reflectance from its internal glass wall to redirect scattered light towards the absorption sensor and minimize scattering losses; however, scattering corrections are still required to account for residual losses. Furthermore, the attenuation sensor may collect photons scattered in a forward direction at angles smaller than the attenuation sensor aperture, which is a lens-pinhole with collection angle θ_c = 0.9°. This artificially lowers attenuation values. A number of different methods for scattering correction of absorption exist, the most commonly used of which has traditionally been the proportional correction (Zaneveld *et al*., 1994). In recent years, two new correction procedures have been proposed, the semi-empirical correction (Röttgers *et al*., 2013) and the iterative correction (McKee *et al*., 2013). The iterative correction in particular corrects for errors in both absorption and attenuation measurements.

5.2.2.2 Absorption and attenuation measurement protocol

The ac-9 attenuation and absorption meter was operated following methodologies established in the user's manual and protocol for this instrument (WET Labs Inc., 2008; WET Labs Inc., 2011); the procedure can also be found detailed in Lefering *et al*. (2016). For this specific set of measurements (wavelengths used: 412, 440, 488, 510, 532, 555, 650, 676 and 715 nm) the diluted samples were poured into a reservoir and then pumped through the instrument using a hand-pump. The out-flowing sample was then collected again within the black container. Attenuation and absorption data were scattering-corrected using the iterative correction proposed by McKee *et al*. (2013), and were further corrected for salinity and temperature dependence following Pegau *et al*. (1997); the final values for absorption and scattering (averaged over the measurement period) are shown in Fig. 5.1.

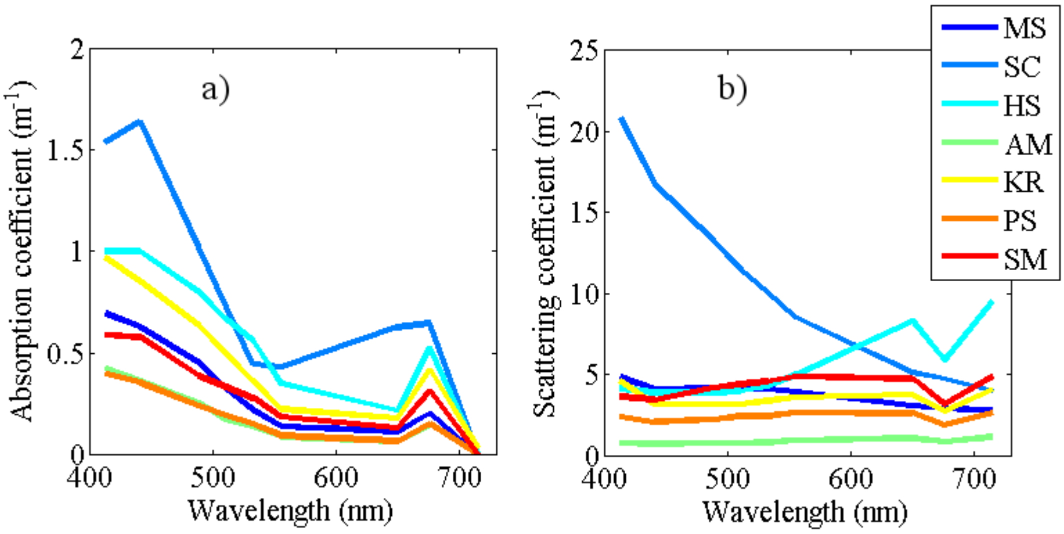


Figure 5.1 – Spectral (a) absorption and (b) scattering coefficients for the AC dataset. Some absorption values resulted negative after scattering correction and are not represented.

5.2.3 Backscattering measurements

5.2.3.1 Backscattering sensor working principle

The WET Labs BB9 instrument is a submersible, *in situ* backscattering sensor which similarly to the ac-9 is designed to provide real-time measurements over nine individual wavelength channels. A thorough review and description of the principles behind backscattering measurements can be found in Sullivan *et al*. (2013). In essence, modern backscattering sensors like the BB9 are VSF meters: they take advantage of a peculiarity of scattering phase functions, which are observed to have low variability in their shape at backwards angles (Sullivan and Twardowski, 2009), leading to the determination of conversion factors capable of providing estimates of backscattering b_b values from single measurements of the VSF at backward angles, known as χ factors (Oishi, 1990; Boss and Pegau, 2001). These are indeed the same χ factors discussed in Chapter 4. Similarly to ac-9 measurements, BB9 backscattering measurements need to be corrected too, this time for absorption over the pathlength of the beam used to illuminate the sample. This makes concurrent ac-9/BB9 measurements important.

5.2.3.2 Backscattering measurement protocol

As with the ac-9, the BB9 was operated following procedures established in the user’s manual for the instrument (WET Labs Inc., 2010). For this specific set of measurements (wavelengths used: 412, 440, 488, 510, 532, 595, 660, 676 and 715 nm) the head of the instrument was plunged in the diluted samples within the container and kept in place to gather a few minutes’ worth of data for each algal species. Backscattering data was then corrected for pathlength absorption using ac-9 absorption data scattering-corrected using the proportional correction (Zaneveld *et al*., 1994) and averaged over the measurement period (Fig. 5.2). Finally, the backscattering values retrieved by the BB9 were interpolated to match the 9 wavelengths used by the ac-9. In turn, the BB9 backscattering data was used to inform the iterative correction procedure used for ac-9 absorption and attenuation, which also explains why pathlength absorption was corrected using the proportional rather than the iterative method.

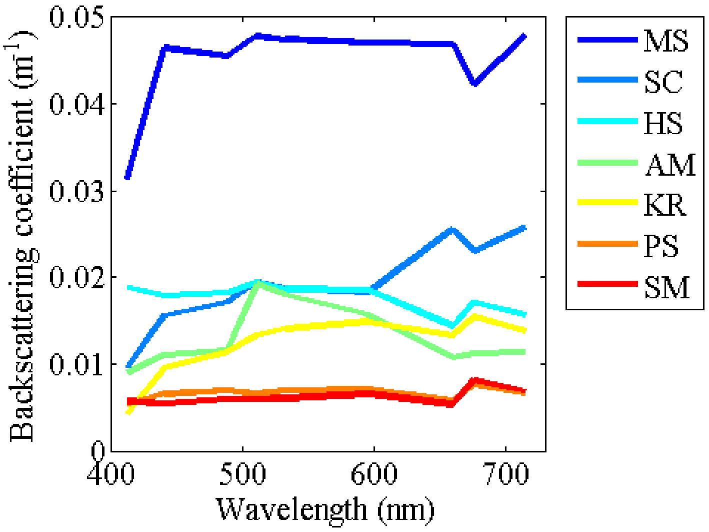


Figure 5.2 – Spectral backscattering coefficients for the AC dataset.

5.3 UK coastal waters

The UKCW dataset consists of natural water samples obtained during the HE442 research cruise in UK waters (4th-21st of April 2015) on board the R/V Heincke. Sixty-two stations were sampled across a variety of case 1 and case 2 waters around the coast of the UK (Fig. 5.3), supplying a total of 50 samples with complete sets of FC data and matching data from other instruments. This includes backscattering, attenuation and absorption values retrieved using an ac-9 absorption and attenuation meter and a BB9 backscattering meter; PSD measurements obtained by two LISST-100x laser diffractometers (Type B, size range: 1.25-250 µm and Type C, size range 2.5-500 µm, both by Sequoia Scientific; LISST-B and LISST-C respectively in the following); and biogeochemistry parameters for suspended matter, organic carbon and chlorophyll-A concentrations.

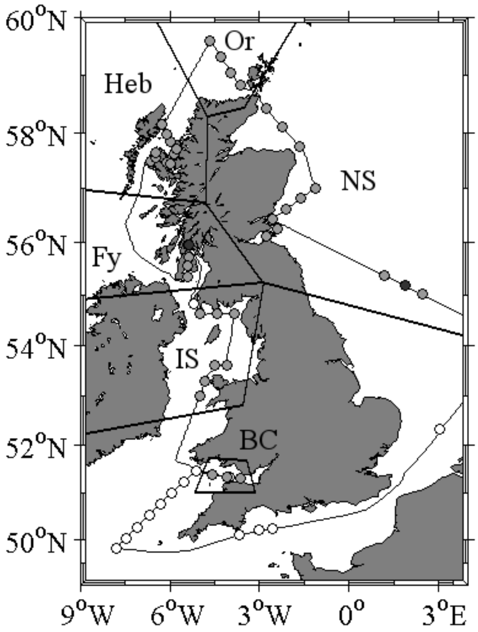


Figure 5.3 – Track of the HE442 research cruise, which took place in April 2015 in UK waters aboard R/V Heincke. Out of the 62 measurement stations visited a total of 50 complete sets of data were retrieved, matching flow cytometric data and ancillary measurements (light grey circles). Dark grey circles denote stations where two samples were taken. The dataset was divided into area groupings to highlight regional behaviour in the metadata and the measured IOPs. In clockwise order: Bristol Channel (BC), Irish Sea (IS), Loch Fyne and Firth of Clyde (Fy), Hebrides and Skye (Heb), North Atlantic - Orkneys (Or) and North Sea (NS).

5.3.1 Depth profiling

Main depth profiling was done via an instrument frame equipped with Niskin bottles for sample retrieval and with both ac-9 and BB9 sensors and one of the two LISST-100x instruments (LISST-C). The frame was lowered into water at each of the stations, kept near surface for a first round of sampling, then lowered to maximum depth to retrieve further samples. BB9, ac-9 and LISST-C instruments logged data throughout the procedure. The samples were taken from the Niskin bottles on the frame as quickly as possible after the frame was back on deck and filled into 10-liter plastic containers. In waters with high turbidity the Niskin bottles were flushed twice to avoid settling out of particulate matter. 48 out of the 50 samples of the UKCW dataset are surface samples (depth: 5-7 m), with two more samples taken from bottom depths instead.

5.3.2 Flow cytometry measurement protocol

All samples were measured by the flow cytometer once for each of four sensitivity settings of the side scattering PMT (50, 60, 70, 80), for 6 minutes and at a flow rate of 0.5 µL/s. Side scattering was used as the trigger channel in all cases. Additional measurements of standard polymer beads for calibration of the FC method were taken daily across the whole sampling period. The low sample flow rate was necessary because of hardware problems which arised in the flow cytometer during the research cruise, causing the sample core to lose stability at higher flow rates.

5.3.3 Attenuation, absorption and backscattering measurements

The ac-9 absorption and attenuation meter and the BB9 backscattering sensor were operated along similar lines to those presented in paragraphs 5.2.2.2 and 5.2.3.2, with the only major difference that the instruments were installed on a submersible frame rather than kept in place within the lab. Attenuation, absorption, scattering and backscattering measurements were averaged over the stationary phase of the profiling, near the surface (or at bottom depth where it applied). As was the case for the AC dataset, ac-9 absorption data corrected using the proportional method was used to correct for pathlength absorption in BB9 backscattering measurements; these were then in turn used to inform the iterative correction of ac-9 absorption and attenuation values. Absorption, scattering and backscattering values for the UKCW dataset are shown in Fig. 5.4-5.6.

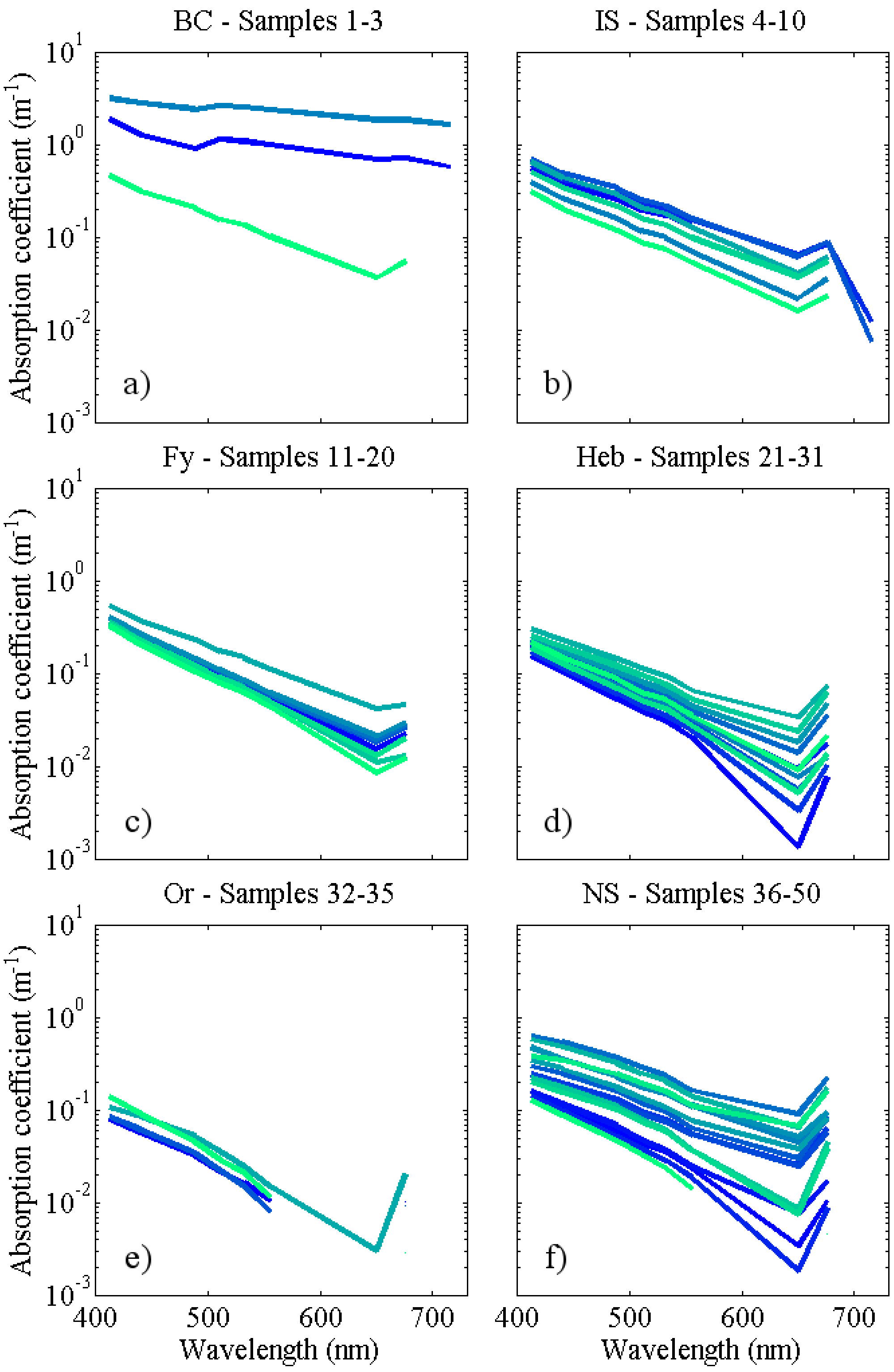


Figure 5.4 – Spectral absorption for (a) BC, (b) IS, (c) Fy, (d) Heb, (e) Or and (f) NS subsets of the UKCW dataset (see Fig. 5.3). Colors represent samples ordinally from dark blue to light green. Some absorption values resulted negative after scattering correction and are not represented.

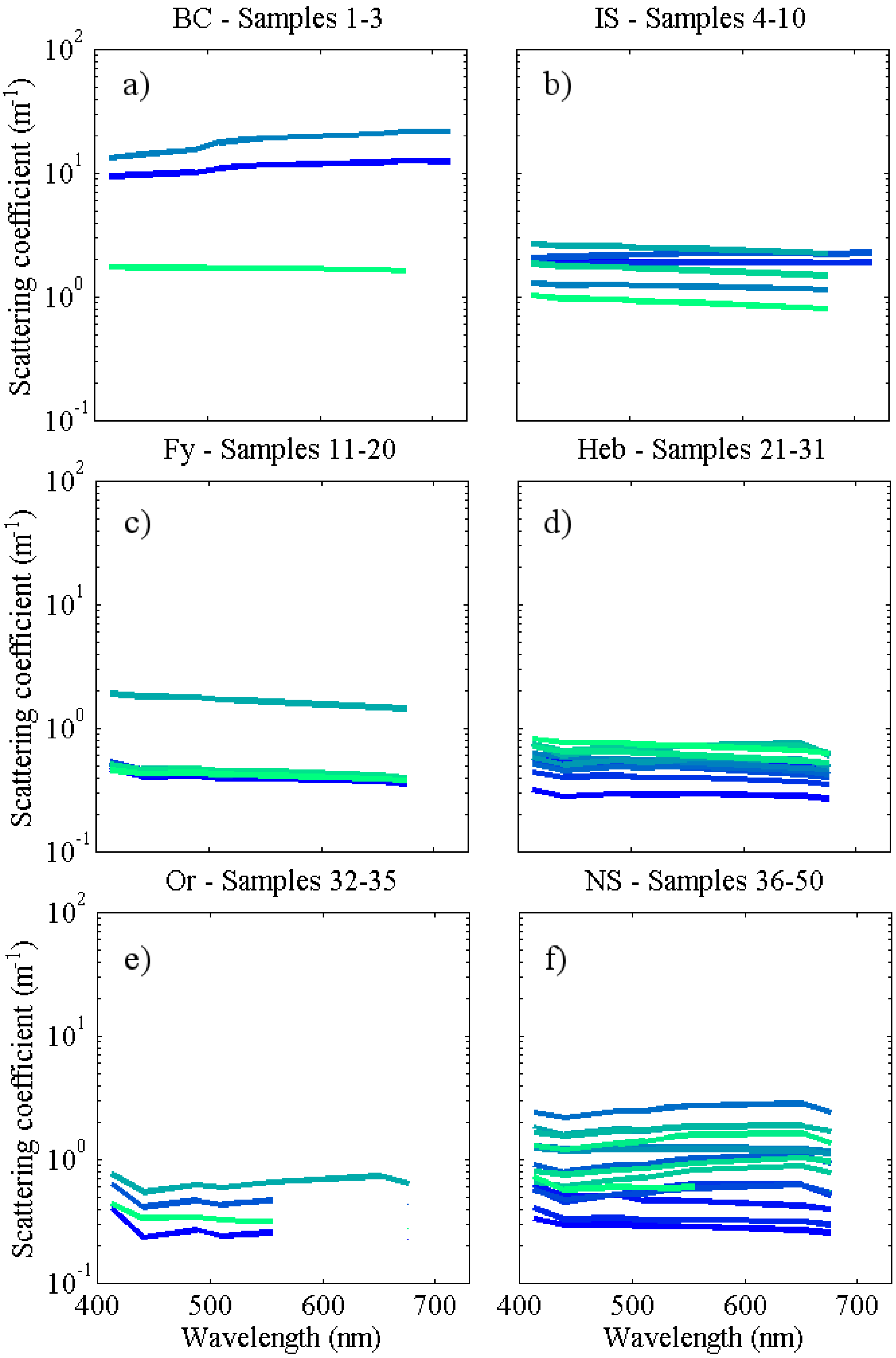


Figure 5.5 – Spectral scattering for (a) BC, (b) IS, (c) Fy, (d) Heb, (e) Or and (f) NS subsets of the UKCW dataset (see Fig. 5.3). Colors represent samples ordinally from dark blue to light green. Some scattering values resulted negative after scattering correction and are not represented.

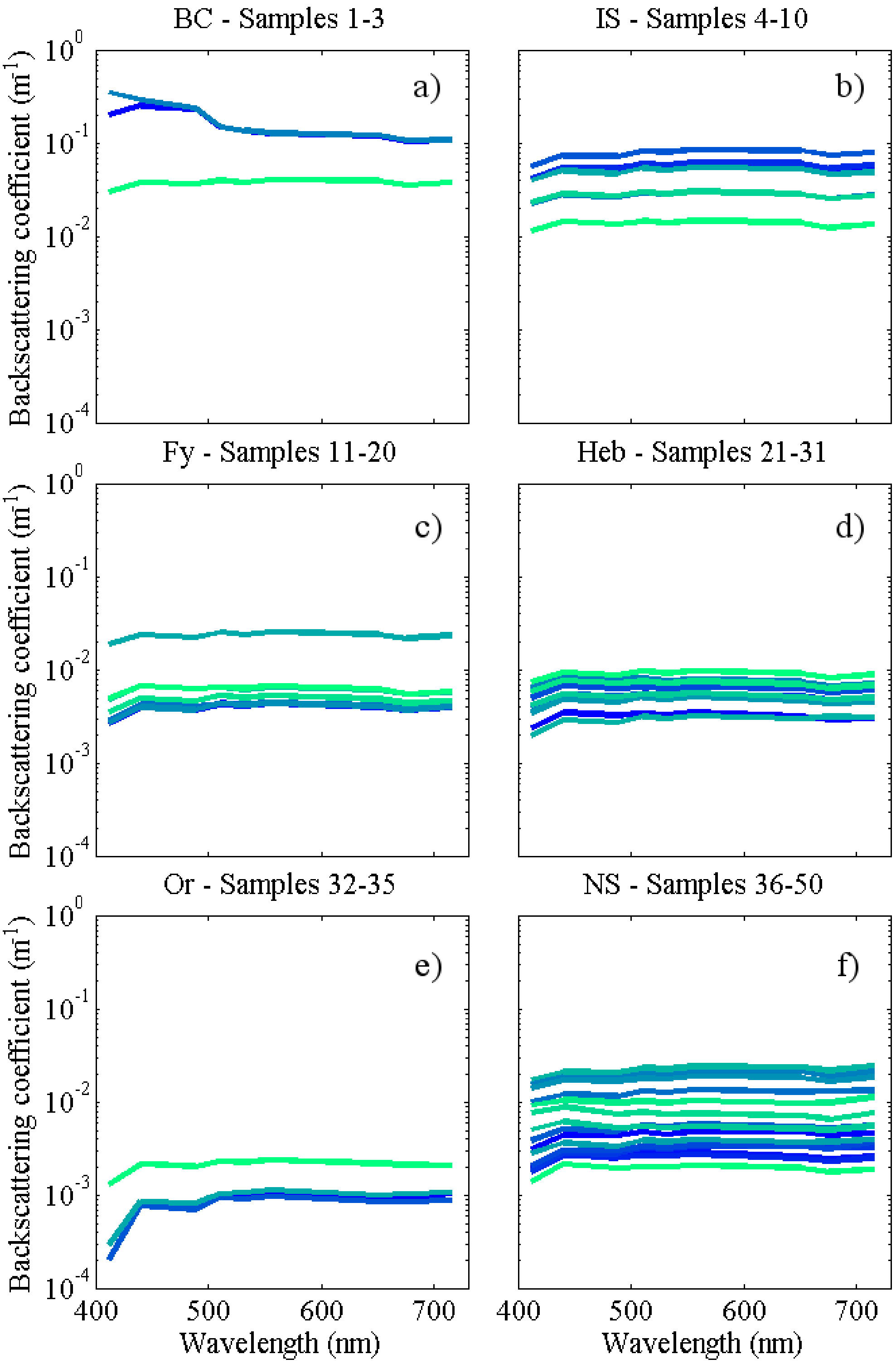


Figure 5.6 – Spectral backscattering for (a) BC, (b) IS, (c) Fy, (d) Heb, (e) Or and (f) NS subsets of the UKCW dataset (see Fig. 5.3). Colors represent samples ordinally from dark blue to light green.

5.3.4 LISST-100x PSD measurements

5.3.4.1 LISST-100x working principle

The LISST-100x is a laser diffractometer which allows for rapid acquisition of bulk PSD profiles (Agrawal & Pottsmith, 1994; Traykovski *et al.*, 1999; Agrawal & Pottsmith, 2000). As the name implies, these instruments take advantage of a principle known as laser diffraction: specifically, Mie theory results show that for particles large compared to the wavelength of the incident light the distribution of light scattered at small forward angles appears identical to the diffraction patterns of light passing through an aperture of equal diameter. As diffracted light does not pass physically through the particle, small angle scattering is largely unaffected by the material and composition of the particle; therefore the method provides results which are chiefly functions of particle size. LISST-100x instruments collect small angle scattering using a series of 32 ring-shaped detectors, with logarithmically increasing radii and covering logarithmically increasing ranges of angles. Scattering from particles peaks at increasingly smaller angles as the particle diameters become larger, so that each detector ring mainly accounts for a specific scattering dominant and therefore for a specific size class (although all detector rings receive some fraction of scattering from the particles). Finally, since total bulk scattering is the result of the sum of scattering from individual particles, the scattering intensity detected in each ring (weighted by the area of the ring itself) is a function of the total particle number within each respective size bin. By inverting the optical power distribution detected in the rings the LISST-100x produces the PSD of the particle population in a sample. Each LISST-100x is further equipped with a transmissometer photodiode, which is used to account and correct for the influence of attenuation on the light collected by the detector rings.

5.3.4.2 LISST-100x PSD measurement protocol

Both LISST-C and LISST-B were maintained and operated following the instructions found in the user’s manual for LISST-100x instruments (Sequoia Scientific Inc., 2013). As stated in paragraph 5.3.1, LISST-C was installed on the instrument frame used for main depth profiling together with the ac-9 and BB9 absorption, attenuation and backscattering meters. Similarly to the latter two instruments, PSD data obtained using the LISST-C was averaged over the stationary phase of the profiling, near the surface or at bottom depth as appropriate to each specific sample (Fig. 5.7a-b). LISST-B was installed on a separate submersible platform, which was deployed and operated independently from the main instrument frame. This second frame housed the Multi-angle Scattering Optical tool, or MASCOT (Sullivan and Twardowski, 2009), an instrument which measures VSFs at a single wavelength at angles between 10° and 170°. Data from the MASCOT is not included in this study.

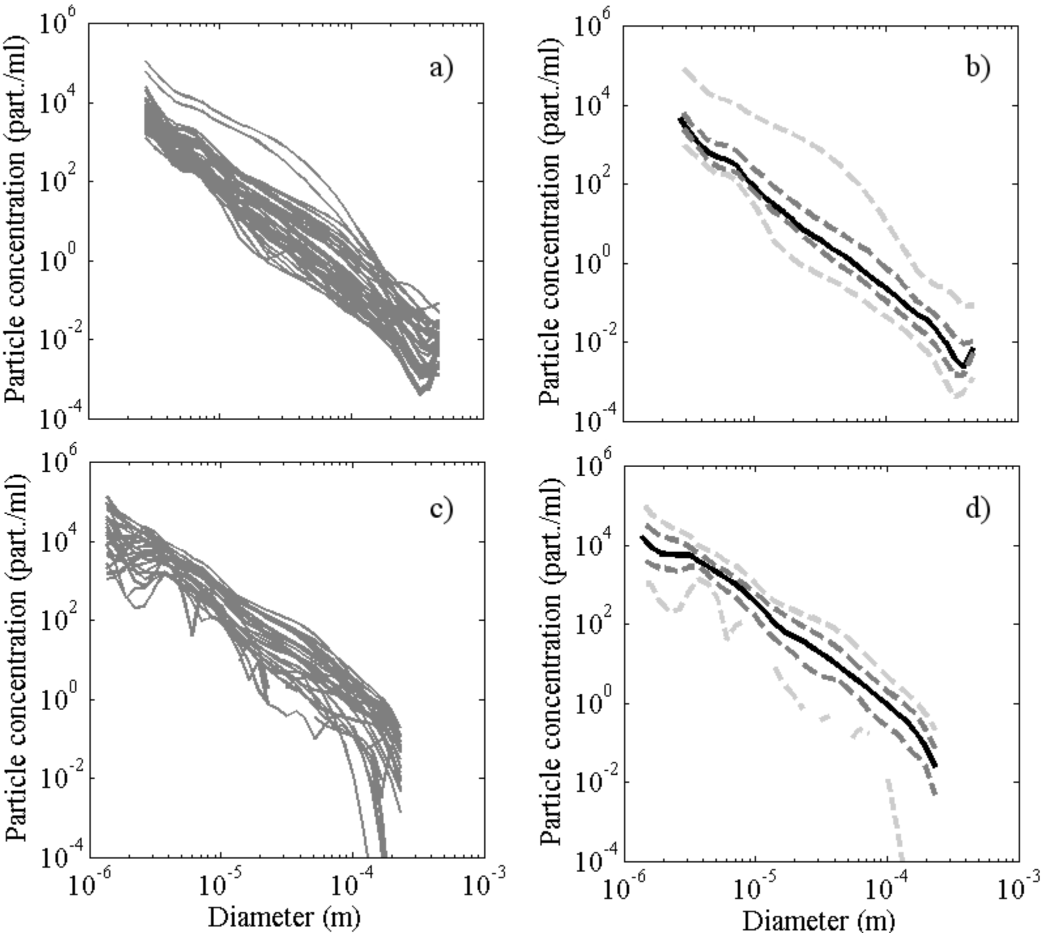


Figure 5.7 – (a) Collective view of all 50 UKCW PSDs obtained by the LISST-C and (b) median PSD for the same set of PSDs (quartiles and max./min. given as dark and light grey dashed lines respectively). In panels (c-d) the same plot is given for the UKCW PSDs obtained by LISST-B. Note that LISST-B data was available for only 28 out of the 50 samples of the UKCW dataset. Interestingly, differences are observed in the PSDs as determined by the two LISST instruments; these will be discussed in Chapter 6.

Depth information from the measurement profiles was compared between the two submersible platforms to extract the LISST-B PSDs matching the data obtained from the other instruments. Two casts were made at each station using the MASCOT frame; the LISST-B data used in the following is the result of averaging between the two casts. LISST-B data was available for only 28 out of the 50 samples of the UKCW dataset (Fig. 5.7c-d).

5.3.5 Suspended matter, organic carbon and chlorophyll-A content

5.3.5.1 Total, inorganic and organic suspended matter

Total suspended matter (TSM, also found in literature as total suspended solids, TSS) was obtained from each sample following procedures detailed by Röttgers *et al*. (2014) (Fig. 5.8a). The sample was run through filter pads under low vacuum, then immediately placed in a petri dish after filtration and put to dry in a vacuum desiccator: TSM values were subsequently obtained by weighing the mass of dried sample and dividing it by the sample volume used. Further separation into organic suspended matter and inorganic suspended matter fractions (OSM and ISM, also found in literature respectively as particulate organic matter and particulate inorganic matter, POM and PIM; Fig. 5.8b) was obtained by volatisation of organics at 500°C: ISM is given as the weight of the sample mass remaining after the combustion procedure divided by the sample volume used, with OSM simply determined as TSM minus ISM.

5.3.5.2 Particulate organic carbon

Particulate organic carbon (POC) was determined using the protocol described by Strickland and Parsons (1972) (Fig. 5.8c). The samples were initially prepared by filtering through a Millipore AA filter (Merck-Millipore) treated with magnesium carbonate to prevent adhesion to the membrane surface. Particulate matter caught in the filter was then washed off and re-filtered through a sintered glass filter disc under vacuum. The glass filter was subsequently removed, placed in a sample beaker with the sample deposit face up and mixed with 1 ml each of phosphoric acid and distilled water; the beaker was then fitted with a coverglass and the sample heated to 110°C for 30 minutes.

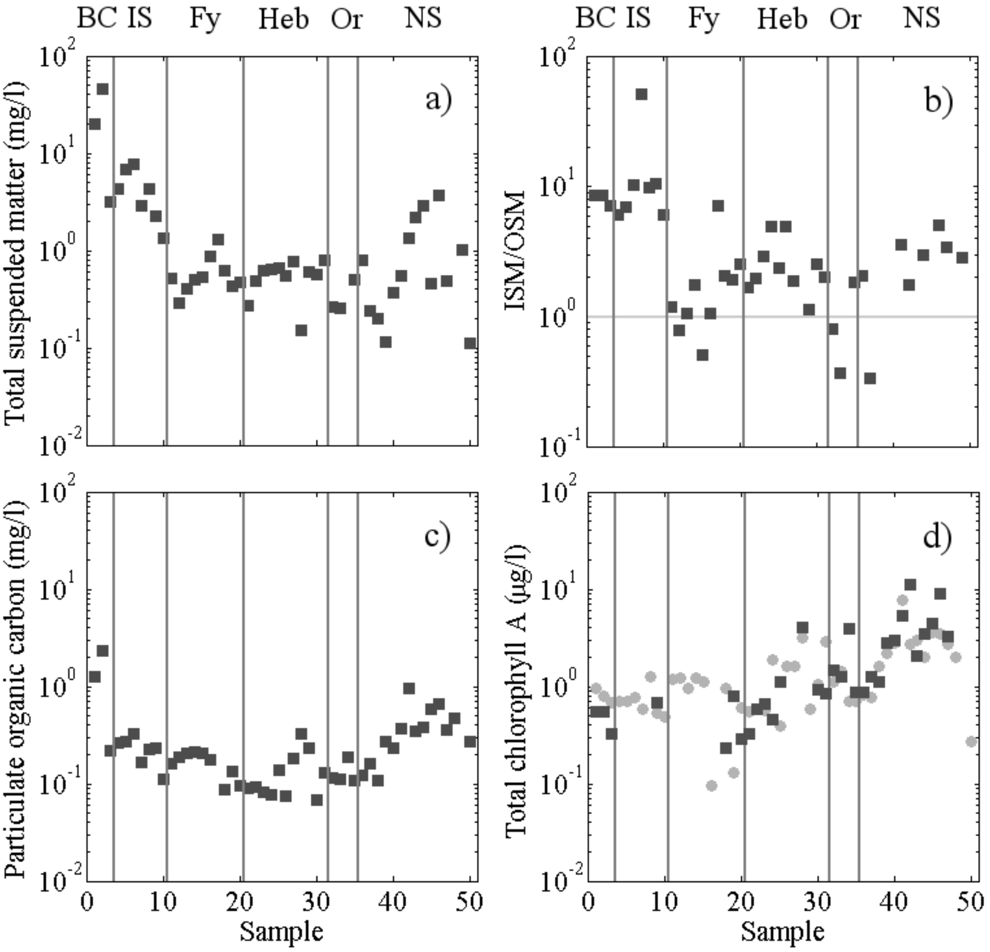


Figure 5.8 - (a) TSM, (b) ISM/OSM balance, (c) POC and (d) ChlA data for the UKCW dataset. In the last panel, grey circles represent HPLC-derived ChlA values and dark grey squares represent ChlA estimates given as 5.5 times the initial fluorescence F_o of the sample. Note that the latter weren't available for all 50 samples of the UKCW dataset.

After 30 minutes, the coverglass was removed to allow further addition of sulphuric acid-dichromate oxidant, which is produced as a mixture of potassium dichromate solution into concentrated sulphuric acid (full details in Strickland and Parsons, 1972). The coverglass was then replaced and the sample heated for a further 60 minutes. After 60 minutes, the mixture was finally left to cool before being thoroughly washed out of the beaker and into a suitably sized graduated cylinder using copious distilled water. After settling, a small volume of the final mixture was centrifuged at 2000 rpm for a few minutes and then transferred to a spectrophotometer: the POC value of the sample is obtained as a function of the measured extinction of a blank solution against that of the mixture (full equation in Strickland and Parsons, 1972).

5.3.5.3 Chlorophyll-A content

Chlorophyll-A content (ChlA) was determined twice using independent methods (Fig. 5.8d). The first method employed high-performance liquid chromatography (HPLC) following Zapata *et al.* (2000): ChlA was determined from fluorescence values using excitation at 440 nm and emission at 650 nm and by comparison with standards of known chlorophyll concentration. The second method produced approximate values of chlorophyll-A content from an empirical power law relationship between chlorophyll-specific algal absorption and chlorophyll content, as described by Bricaud *et al.* (1995). This particular iteration of the technique used F_o as the input, that is the initial fluorescence of the sample as measured by fast repetition rate fluorometry (FRRf) (Connor, 2017); it was found that a simple scaling factor of 5.5 applied to F_o fitted the relationship described by Bricaud *et al.* (1995) well (Connor, 2017). Therefore, $Chl_F = 5.5F_o$ was used as an estimate of ChlA. The decision to utilise chlorophyll data from two separate techniques was taken following Connor (2017), who also worked on the UKCW dataset and found that the relationship between algal absorption and chlorophyll concentration was stronger for F_o-derived concentration values, suggesting that biogeochemistry modelling based on FC PSDs and PRIDs might also compare better with Chl_F estimates of ChlA. The validation of biogeochemistry modelling of ChlA values against both sets of chlorophyll concentration data will be described in Chapter 8.

5.4 Summary

This short chapter presented the two main datasets used in this work, one consisting of phytoplankton culture samples (AC dataset) and one consisting of marine water samples retrieved in UK shelf waters (UKCW dataset). A description of the instruments used during both measurement campaigns was also provided along with relevant principles and protocols. Having established a basis for all data processing and analysis, the rest of the thesis will focus on the presentation of results. The next chapter will describe the core topic of this work: the flow cytometric determination of particle diameters and rRIs, or FC method. Chapters 7 and 8 will then respectively focus on IOP forward modelling and biogeochemistry modelling.

6. Flow cytometric method for particle diameter and real refractive index determination

With the theoretical basis given in Chapters 1-4 for IOPs, flow cytometry and Mie theory, all elements are in place for the development of the flow cytometric method for size and real refractive index determination (*FC method*) and its application to the datasets presented in Chapter 5. The first half of this chapter will present the FC method, including the characterisation of the geometry of sensor apertures in the CytoSense and their respective weighting functions. The second half will discuss the application of the FC method, first to known standards of particle diameter and RI, then to the PSDs and PRIDs of the AC and UKCW datasets.

6.1 Introduction

As explained previously (Section 1.1), flow cytometry was originally developed for biological studies, and is still used in this fashion in a large majority of its applications. An estimation made a few years ago by Dubelaar & Jonker (2000) put the figure for biomedical applications at about 95% of the total, with the rest divided between fields as disparate as pharmaceutical industry, dairy industry, food and water quality control, botany and, of course, marine science; there is little reason to think the figure has changed much in the intervening years. Even within marine science itself, a large number of studies employ flow cytometry in its traditional fashion, using fluorescent dyes and DNA stains to label cells and study their physiology or exploiting in-line imaging to conduct monitoring and taxonomy assessments of marine populations (e.g. Mikulski *et al*., 2005; Sosik & Olson, 2007; Olson & Sosik, 2007; Thyssen *et al*., 2008; Thyssen *et al*., 2011; Brosnahan *et al*., 2013). When it comes to marine optics, however, all components of a population of suspended particles contribute in some degree to the optical behaviour of the water, be they living cells, dead cells, organic detritus or indeed inorganic particles and suspended sediments. Therefore, while fluorescent signals are the most commonly used parameters for triggering particle detection and are certainly well suited when phytoplankton is the focus, they are also insufficient when the entirety of the particle population is considered. The FC method follows in the steps of Ackleson & Spinrad

(1988) and more recently Green *et al*. (2003a, 2003b) by using the scattering signal (specifically side scattering) as trigger instead, and combines Mie theory and flow cytometry to model the output of the sensors of the flow cytometer and assign size and refractive index to each individual particle.

6.2 Flow cytometry method outline

The method for assigning refractive indices and sizes used in this work was first developed by Ackleson & Spinrad (1988) based on algorithms by Dave (1968), and subsequently reprised by Green *et al*. (2003a).

The FC method simulates scattering within the flow cytometer via eqs.

$$\beta(\theta,\lambda) = \frac{1}{2}\frac{1}{k^2 r^2}\frac{(i_1 + i_2)I_i(\lambda)}{E_i(\lambda)} = \frac{1}{2}\frac{(i_1 + i_2)}{k^2} \qquad (2.48)$$

and

$$b_{\theta\prime,\theta\prime\prime}(\lambda) = 2\pi \int_{\theta\prime}^{\theta\prime\prime} \beta(\theta,\lambda)\sin\theta \, d\theta, \qquad (6.1)$$

where θ’ and θ’’ are determined by the aperture of the sensors within the instrument. These Mie theory calculations are carried out for wide ranges of particle diameters and refractive indices and produce a look-up table of simulated forward and side scattering weighted for sensor-specific shape characteristics. After a correspondence between simulated and measured scattering values is established using standards of known diameter and RI, each new measurement pair of forward and side scattering is mapped against the closest node in the look-up table and assigned a corresponding diameter and rRI. The resulting particle data is then binned to produce PSDs and PRIDs. This technique is fundamentally different from traditional flow cytometric methods which directly assess the length of particles from the width of their signal curve (signal baseline length or full-width at half-maximum, FWHM, of the curve), but which cannot discriminate sizes smaller than the width of the laser beam (5 µm in the CytoSense) and do not provide rRI info.

The FC method described here employs an independently developed MATLAB script, using the FASTMie code developed by Slade (2006). The script produces the scattering amplitude functions $i_1(\theta)$ and $i_2(\theta)$, which by virtue of the Mie coefficients a_n and b_n of eqs. (2.24-2.25) are fundamentally dependent on the diameter and RI of each particle. The imaginary refractive index was kept fixed and set at zero: a fixed value for the imaginary component is necessary as the method does not offer information on this term, and was appropriate for the polymer calibration beads which were used for tuning the model. Test simulations for realistic iRI values suggested negligible effect of this assumption on diameter and rRI identification. The quantities $i_1(\theta)$ and $i_2(\theta)$ were used to calculate the single-particle VSF of eq. (2.48) for all combinations of 40 rRI values ranging from 1.335 to 1.725 (relative to vacuum) in increments of 0.01, and 300 log-spaced diameter values ranging from 1e^{-8} to 1e^{-4} metres. The angular resolution for the VSF varied between 0° and 180°, with 0.01° increments between 0° and 1°, 0.1° increments between 1° and 10° and 1° increments between 10° and 180°, for a total of 361 angular values. Finally, $\lambda = 488$ nm was used for the VSF calculations conforming to the wavelength of the laser source within the CytoSense. This wavelength (given for vacuum) was further corrected for transmission through water using the absolute rRI value of water (1.333). All particle rRIs in the following are also given relative to this value unless otherwise stated.

Once obtained, each VSF was integrated to simulate scattering within the flow cytometer, and the results scaled to establish correspondence with measured scattering from the particles of known size and refractive index, thus producing the required look-up table. Having assessed the precision concentration measurement capabilities of the flow cytometer (Chapter 3, Section 3.4), the procedure can be divided in 4 steps:

- Characterization of the sensor geometry within the instrument;
- Scaling of the modelled scattering grid;
- Rescaling and combination of measured data for different PMT sensitivity settings;
- Final binning of particle sizes and refractive indices in PSDs and PRIDs.

FC method calculations were handled using an independently developed MATLAB code which integrates the FASTMie code provided by Slade (2006) to calculate Mie coefficients. All four core sections of the code, each corresponding to one of the 4 steps of the FC method, are presented in Appendix B.

6.2.1 Sensor shape functions

Simulation of the FC scattering sensors requires characterisation of the correct $[\theta', \theta'']$ angles over which eq. (6.1) is to be integrated. Sensor angle ranges were first manually determined by rough approximation. Improved ranges were then refined iteratively until the model grid of look-up table nodes matched scattering data from reference polymer beads of known diameter and refractive index. It is important to remember that the VSF given here is axially symmetric along the direction of the incident light, so that integrating over an angular range accounts for the scattering into ring sections of the radiant sphere. The sensors however only intercept part of the ring sections, so that sensor shape functions are needed to further account for the relative weights.

For the *forward scattering sensor shape function* (polar angle range: ±[2°, 9.7°]), consider the geometrical setup of Fig. 6.1a-b. The length of the arc $n = m - l$ for each point in the range $[\theta', \theta'']$ is to be determined and then used as a correction factor after normalisation to the corresponding arc x of a unit circle. From simple trigonometry,

$$a = \sin l \tag{6.2}$$

and

$$l = \operatorname{asin} a. \tag{6.3}$$

The length of the arc to be determined is (in radians) $n = \frac{\pi}{2} - l$, and can be normalised as

$$x = \frac{\frac{\pi}{2} - l}{\frac{\pi}{2}}. \tag{6.4}$$

The condition of normalisation to a unit circle for a and b is given by $a\!:b = a'\!:1$. It follows that

$$a' = \frac{a}{b} \;\Rightarrow\; x = \frac{\frac{\pi}{2} - \operatorname{asin}\frac{a}{b}}{\frac{\pi}{2}} \tag{6.5}$$

and since from Fig. 6.1b-c it is clear that

$$a = \sin\theta' \tag{6.6}$$

and

$$b = \sin\theta'', \tag{6.7}$$

it finally derives that the forward scattering sensor shape function is expressed as

$$x_{SHAPE,f}(\theta) = \frac{\frac{\pi}{2} - \operatorname{asin}\frac{\sin\theta'}{\sin\theta}}{\frac{\pi}{2}}. \tag{6.8}$$

The shape of the forward scattering sensor shape function is shown in Fig. 6.1d. Note that the weighting function is calculated for just a quarter of the actual area; correct proportionality is ensured by scaling factors.

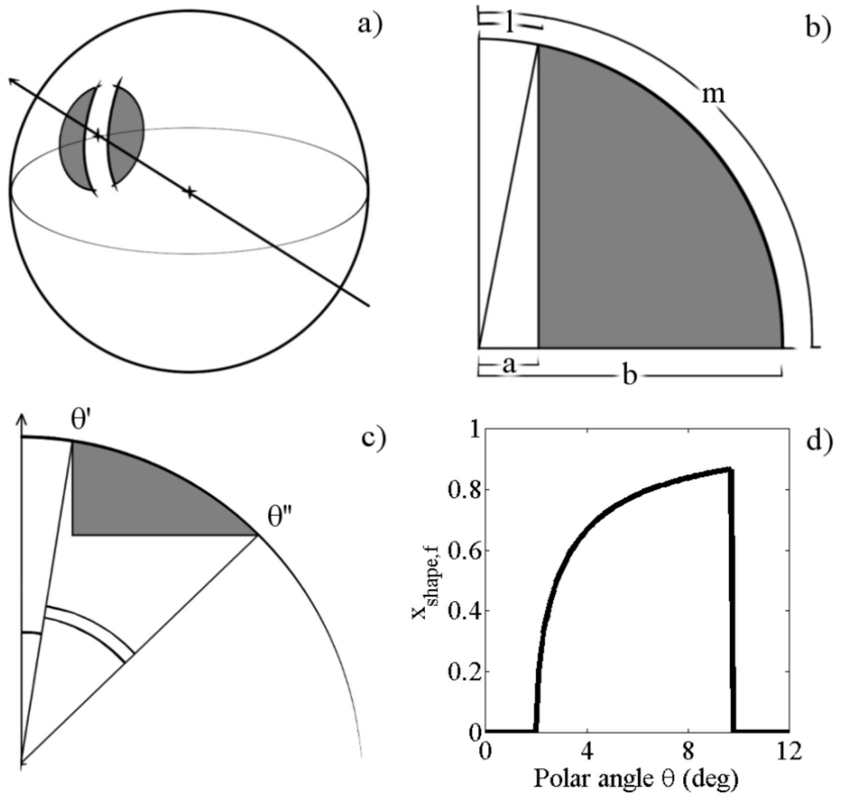


Figure 6.1 – (a) General geometric setup for the forward scattering shape correction calculations and (b) front and (c) top view of the same geometric setup. (d) Shape of the forward scattering weighting function.

For the *side scattering sensor shape function* (polar angle range: [45°, 135°]), consider the geometrical setup of Fig. 6.2a-b. The length of the arc x for each point in the range $[\theta', \theta'']$ centred on $\theta = 90° = \pi/2$ is the parameter to be determined. For each arc x,

$$a = \sin x. \tag{6.9}$$

But since the projection of the relevant area is a circle, the maximum value for *a* will be (in radians)

$$a_{max} = \sin\left(\theta'' - \frac{\pi}{2}\right) \tag{6.10}$$

and a will vary between $[\theta', \theta'']$ like $\sin\xi$ does, given the appropriate mapping between $\theta \in [\theta', \theta'']$ and ξ where $\xi \in [0, \pi]$, i.e.

$$\xi(\theta) = \pi \frac{\theta - \theta'}{\theta'' - \theta'} \tag{6.11}$$

Thus

$$a(\theta) = a_{max} \sin \xi(\theta) = a_{max} \sin \left(\pi \frac{\theta - \theta'}{\theta'' - \theta'} \right). \tag{6.12}$$

By inverting eq. (6.12) and normalising it, it finally derives that the side scattering sensor shape function is expressed as

$$x_{SHAPE,s}(\theta) = \frac{\operatorname{asin} \left[a_{max} \sin \left(\pi \frac{\theta - \theta'}{\theta'' - \theta'} \right) \right]}{\frac{\pi}{2}}. \tag{6.13}$$

The shape of the side scattering sensor shape function is shown in Fig. 6.2d.

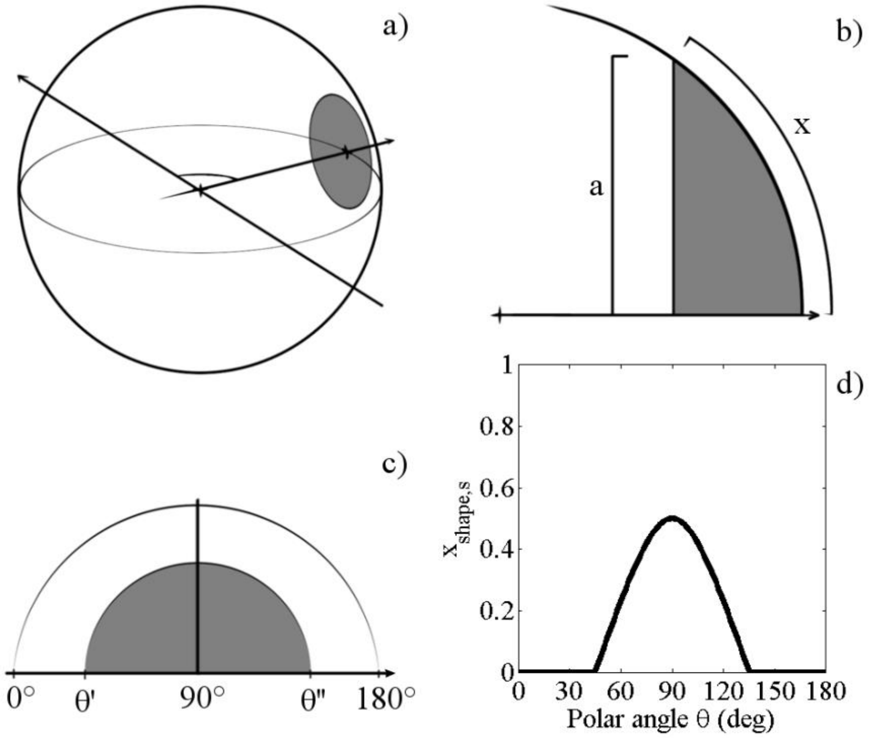


Figure 6.2 – (a) General geometric setup for the side scattering shape correction calculations and (b) forward and (c) side view of the same geometric setup. (d) Shape of the side scattering weighting function.

Similarly to forward scattering, note that the side scattering weighting function is calculated for just half of the actual area; correct proportionality is once again ensured by scaling factors.

Having obtained the expressions for both sensor shape weighting functions, for a single wavelength and axially symmetric scattering the final form of the modelled forward and side scattering within the flow cytometer is given by

$$b^{i}_{grid} = \int \beta_{p}(\theta)\, x_{SHAPE,i}(\theta) d\theta, \tag{6.14}$$

where *i* denotes either forward or side scattering.

6.2.2 Grid scaling

Having established relevant angular ranges and respective weighting functions, simulated particle VSFs were integrated to model the output of the flow cytometer, producing a grid of isolines of diameter and real refractive index between the look-up table nodes (Fig. 6.3). The model was centred and scaled on instrument output data for standard particles of known diameter and refractive index. The average forward and side scattering measurement data for 0.5 µm polymer beads (Duke Particle Counter Size Standards, Thermo Fisher Scientific for these and all other polymer beads used in the remainder of this chapter) was used as reference for the centring of the model grid. The rRI of the beads was 1.197. After scaling the distance between the corresponding 0.5 µm and 1 µm points in the model grid to be the same as the distance between the measured 0.5 µm and 1 µm polymer bead averages, the 0.5 µm point in the scaled model grid was centred on the 0.5 µm polymer bead average, i.e.

$$x^{i}_{GSCALE} = \frac{b^{i}_{avg,1} - b^{i}_{avg,0.5}}{b^{i}_{grid,1} - b^{i}_{grid,0.5}} \tag{6.15}$$

$$b'^{i}_{grid} = \left(b^{i}_{grid} - b^{i}_{grid,0.5}\right) x^{i}_{GSCALE} + b^{i}_{avg,0.5} \tag{6.16}$$

where i denotes either forward or side scattering, b_{grid}^{i} the scattering values of the original model grid, b_{avg}^{i} the average measured scattering values of the polymer beads, and b'^{i}_{grid} the rescaled model grid values.

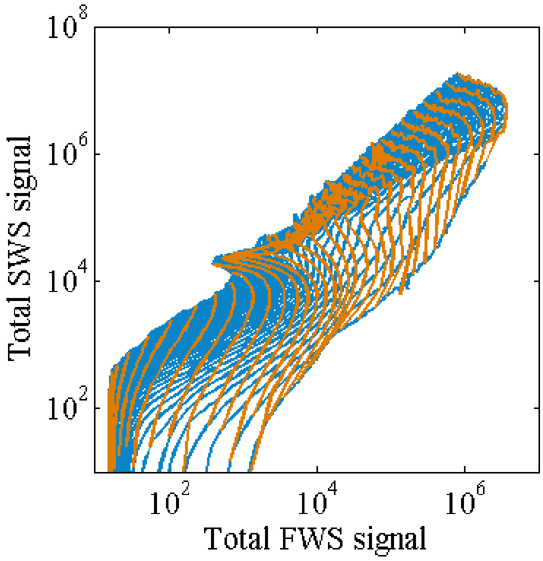


Figure 6.3 – Scattering model grid produced by the FC method, with isolines of diameter (orange lines) and isolines of real refractive index (blue lines).

6.2.3 PMT sensitivity

The range of particle sizes that can be measured with a single PMT sensitivity setting is limited: low sensitivity settings let the instrument detect larger particles without saturating the side scattering PMT, but fail to capture smaller particles; high sensitivity settings let the instrument successfully detect small particles, but saturate the side scattering PMT, hampering the identification of larger particles. To overcome this issue multiple PMT sensitivity settings were used for each sample and results combined together into a single dataset representative of the whole sample.

Care needs to be taken, as each PMT sensitivity setting needs a separate run (*sensitivity run*) and corresponding data ranges partly overlap, so that simple summation of the data matrices will artificially inflate measured concentrations. To deal with this, data from each sensitivity run was therefore cut along user-defined thresholds, producing data subsets that were then merged into a single dataset, working under the necessary assumption that the instrument produces mutually consistent runs for each sub-sample. One sensitivity setting is used as the base setting to which all others are scaled:

$$x^{i}_{SCALE,e} = \frac{b^{i}_{avg,0}}{b^{i}_{avg,e}} \qquad (6.17)$$

where i indicates either forward or side scattering, e the sensitivity setting, $b^{i}_{avg,e}$ the average scattering value of the reference beads for that sensitivity setting, and $b^{i}_{avg,0}$ the average scattering value of the reference beads at the base setting. The choice of PMT base setting in theory doesn't affect the results of the merging; however, a PMT base setting capable of capturing the standard beads necessary for grid scaling without need for interpolation while still keeping PMT saturation in the larger particles at a minimum is preferable in practice. This is suggested as the lowest PMT sensitivity setting capable of clearly measuring both 0.5 µm and 1 µm polymer bead averages. In this study, multiple sensitivity runs for reference beads (1 µm polymer beads) were used to scale the sensitivity run data before cutting and merging the total dataset.

Total analysed volumes for each sensitivity run differ even though the total processed volume is the same across all runs. The discrepancy is due to data transfer overhead – the time during which particle data is acquired and sent to the computer and the instrument is not analysing (but still consuming) the sample (CytoBuoy, private communication). This *does not* produce errors in the estimation of the concentration, but requires further care in the way the final dataset is composed. In this study, the largest total analysed volume among all sensitivity settings was used to scale all other total analysed volumes:

$$x_{VOL,e} = \frac{V_0}{V_e} \qquad (6.18)$$

where e once again marks the individual sensitivity settings and V_0 is the volume used as reference. This factor was used to adjust the concentrations of the corresponding sensitivity run. Data from all sensitivity runs was subsequently merged into the total dataset, and the total number of particles divided by the largest total analysed volume. The resulting dataset is managed as single object, but the resulting concentrations are respectful of the individual contribution of each run:

$$N_{tot} = \sum_e N_e x_{VOL,e} = \sum_e N_e \frac{V_0}{V_e} \Rightarrow C_{tot} = \frac{N_{tot}}{V_0} = \sum_e \frac{N_e}{V_e} = \sum_e C_e \qquad q.e.d. \tag{6.19}$$

Note that while in this implementation of the method the largest total analysed volume was used as reference to scale all other total analysed volumes, the choice is indeed arbitrary; any of the total analysed volumes (corresponding to any of the sensitivity settings) may be used.

6.2.4 Binning

In the last step, the rescaled and merged dataset is mapped on the centred and scaled model grid and binning can take place. Each experimental forward and side scattering data pair is assigned the diameter and refractive index of the closest node in the model grid. The closest node in the grid is identified as the one that minimizes the Manhattan distance

$$l = \left|b^f_{grid,n} - b^f_{exp}\right| + \left|b^s_{grid,n} - b^s_{exp}\right| \tag{6.20}$$

where f and s indicate forward and side scattering respectively, $b_{grid,n}$ is the scattering value for the n-th node in the grid and b_{exp} is the experimental scattering value for the particle. Subsequent binning was carried out using 65 logarithmically spaced diameter bins and 40 linearly spaced rRI bins. Bin values for the diameter (µm) were calculated as $D_i = e^{k_i + 0.16557N}$, with $k_l = -16.87296$, $k_u = -16.70739$ and $k_m = -16.79017$ for lower limit, upper limit and median values respectively and $N \in [0, 64] \subset \mathbb{N}$. The seemingly arbitrary numbers are caused by the adoption of the size bins used in LISST-100x Type-C instruments, the range of which was then extended to cover the 0.05-2000 µm diameter range recommended by Davies *et al*. (2014). Bin values for the (absolute) real refractive index were calculated as $n_i = 0.01N + k_i$, with $k_l = 1.590$, $k_u = 1.600$ and $k_m = 1.595$ for lower limit, upper limit and median values respectively and $N \in [0, 39] \subset \mathbb{N}$.

For each sample, the result of the binning was a $M_{65\times40}$ matrix. Finally, recalling eq. (3.2)

$$x_{PUMP}(Q) = \frac{1.088}{y(Q)} \tag{3.2}$$

where Q is the flow rate and y the value of the general flow rate correction curve and combining all contributions and correction factors together, following eq. (6.19) the concentration for each bin in the matrix is given by

$$C_{D,n}^{tot} = \frac{N_{D,n}^{tot}}{V_0} x_{PUMP} \, , \tag{6.21}$$

where D and n are indices of the size and RI corresponding to each bin. As a final result, summation along the refractive indices or along the diameters respectively provides the PSD and PRID of the sample.

6.3 Flow cytometric method application

The FC method was first applied to standard polymer beads of known diameter and oil suspensions of known refractive index to verify its correct functioning. After this validation, the method was applied to the datasets described in Chapter 5: the AC dataset of algal cultures and the UKCW dataset of marine water samples obtained during the HE442 research cruise in UK shelf waters. Resulting PSDs were compared with known algal parameters for the AC dataset and LISST-100x data for the UKCW dataset. PRIDs were compared with expected refractive index values and (for the AC dataset) with values reconstructed from FC PSDs and attenuation and absorption data using the *anomalous diffraction method* (ADM), which will be briefly described in the following.

6.3.1 *Diameter and rRI standards*

6.3.1.1 Diameter retrieval

To evaluate the overall reliability of diameter retrieval, the FC method was tested on a mixture of polymer beads of several different diameters and a known refractive index of 1.197 (Tab. 6.1).

Polymer bead diameter (µm)	
Mean	St.Dev.
0.498	0.009
0.994	0.015
4.993	0.040
10.12	0.06
50.2	1.0
100.0	1.5

Table 6.1 – Nominal diameter values for the set of reference polymer beads used in the validation of the diameter retrieval capabilities of the FC method.

The method was found to correctly identify small particle diameters down to the minimum detection limit of 0.5 µm (Fig. 6.4). Particles larger than 5 µm however presented increasingly anomalous side scattering profiles (Fig. 6.4c). This is possibly due to the breakdown of the Mie theory assumption of an incident plane wave, which becomes increasingly less true as particles become larger than the laser beam width (5 µm), a problem that was already recognized by Ackleson & Spinrad in their work (1988). This leads to substantial discrepancies between modelled and measured scattering, causing increasingly marked diameter underestimation (Fig. 6.4d).

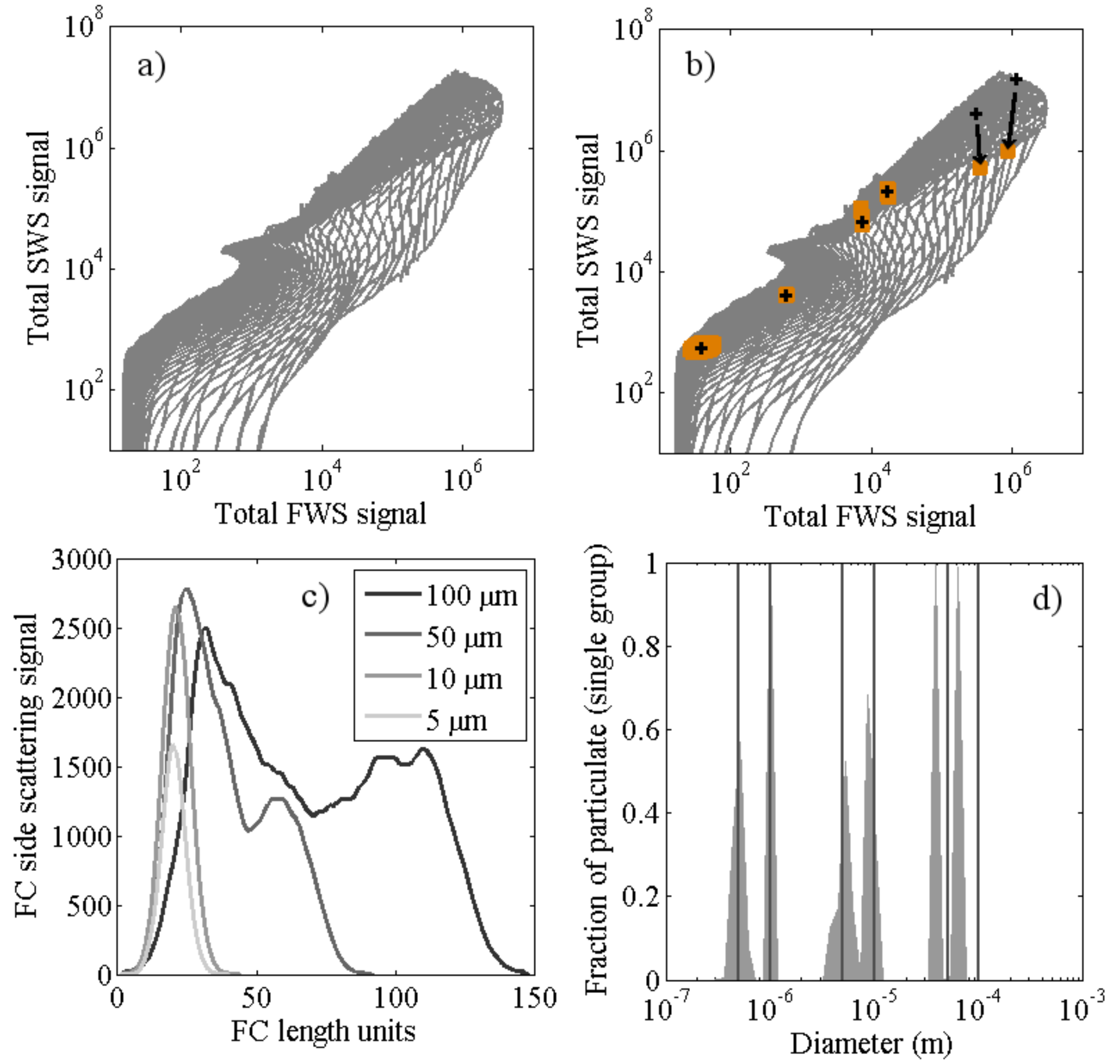


Figure 6.4 – (a) The FC model grid of Fig. 6.3 is replicated here. (b) When data for the set of polymer beads of Tab. 6.1a is superimposed (orange squares) large particles (upper right area) show a marked discrepancy between measured and expected values. (c) This is due to the shape of the scattering profiles increasingly deviating from a Gaussian distribution as diameter increases, possibly due to the breakdown of Mie theory assumptions i.e. particles becoming substantially larger than the laser beam width (5 µm). (d) This is reflected in the resulting size distributions, with the FC method increasingly underestimating the diameter of the larger particles (d > 10 µm). Target values (Tab. 6.1) are represented by vertical lines.

6.3.1.2 Real refractive index retrieval

Polydisperse suspensions of oils with known refractive indices (Tab. 6.2, Fig. 6.5) were additionally used to test the retrieval of rRIs. The method was found to correctly identify low, organic-like rRIs as well as high rRIs when particle diameters are smaller than 5 µm (Fig. 6.5b). Natural bundling of rRI isolines at high rRIs impeded unequivocal determination of a rRI value for particles ~5 µm, providing a generic high-rRI signifier for rRIs ≥ 1.15 as opposed to precise values (note that all rRI values in this study are given relative to water).

Real refractive index		
Substance	Absolute	Relative
Polymer beads	1.595	1.197
Olive Oil	1.469	1.102
Dodecane	1.421	1.066
Nonane	1.405	1.054

Table 6.2 – Nominal rRI values for the set of oil suspensions used in the validation of the rRI retrieval capabilities of the FC method.

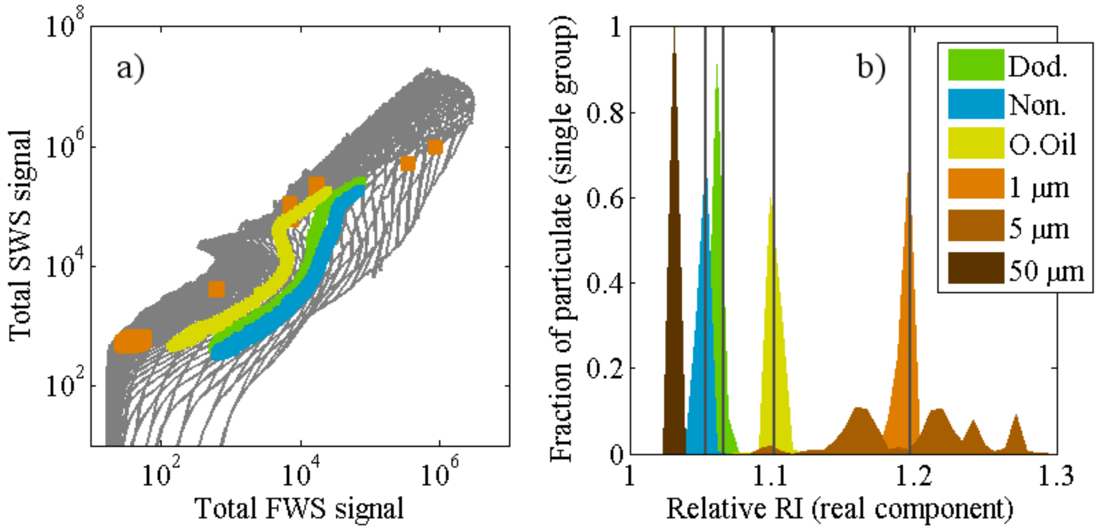


Figure 6.5 – (a) Data points for the polymer beads and three oil suspensions and (b) the resulting rRI distributions. Target values are represented by vertical lines. The rRI detection works well at low refractive indices and/or high refractive indices when particles are small, but fails for high rRIs and larger diameters due to the overlap of the isolines of the model grid for these parameters. Real refractive index detection in very large particles produces values which grossly underestimate the expected ones.

Finally, anomalous side scattering profiles for particles larger than 5 μm leads to gross underestimation of the real refractive index in high-rRI particles (Fig. 6.5b). Overall the FC method was seen to be effective in determining the size of particles in the 0.5-10 µm range and rRIs up to ~1.15; real refractive index values above 1.15 are not precise, but still positively indicate high refractive indices.

6.3.2 AC dataset

Before delving into the application of the FC method to the algal culture samples of the AC dataset, a brief description is given of the anomalous diffraction method, which will be used to model RI values with which to compare those retrieved by the FC method.

6.3.2.1 Anomalous diffraction method

The real and imaginary bulk refractive indices of the algal cultures were estimated using the anomalous diffraction method (ADM) (Bricaud & Morel, 1986; Stramski & Mobley, 1997). A thorough presentation of this method is beyond the scope of this study; in brief, the method compares the experimental bulk values of the absorption and attenuation efficiencies with their expression as derived from the anomalous diffraction approximation (van de Hulst, 1957), which is valid for relative real refractive indices n_r close to unity and imaginary refractive indices n_i close to zero. By minimising the difference between the two the ADM identifies an appropriate approximate complex refractive index for the sample. Assuming spherical particles, following eqs. (1.26) and (1.29-1.30) the bulk absorption efficiency is given by

$$Q_a(\lambda) = \frac{a(\lambda)}{\sum N(D)G} = a(\lambda)\sum \frac{4}{N(D)\pi D^2}, \qquad (6.22)$$

where $a(\lambda)$ is the bulk absorption coefficient at wavelength λ, $N(D)$ the PSD within the sample and G is the geometric cross-section. The bulk experimental attenuation efficiency is similarly defined replacing $a(\lambda)$ with $c(\lambda)$. Under the anomalous diffraction approximation, the corresponding expressions for absorption and attenuation efficiencies are

$$Q_a(\rho') = 1 + 2\frac{e^{-\rho'}}{\rho'} + 2\frac{e^{-\rho'} - 1}{\rho'^2} \tag{6.23}$$

and

$$\begin{aligned} Q_c(\rho) = 2 - 4e^{-\rho\tan\xi}\Bigg[\frac{\cos\xi}{\rho}\sin(\rho - \xi) \\ + \left(\frac{\cos\xi}{\rho}\right)^2\cos(\rho - 2\xi)\Bigg] + 4\left(\frac{\cos\xi}{\rho}\right)^2\cos 2\xi, \end{aligned} \tag{6.24}$$

where $\rho' = 4\alpha n_i$, $\alpha = \pi D/\lambda$, $\rho = 2\alpha(n_r - 1)$ and $\xi = \tan^{-1}\left(\frac{n_i}{n_r - 1}\right)$. Both were calculated as weighted averages over the PSDs.

To determine the bulk complex refractive indices, firstly an appropriate value for n_i is found that minimises the difference between the bulk experimental and anomalous diffraction approximation values of Q_a; the value for n_i thus found is then fed into the anomalous diffraction approximation for Q_c and the procedure repeated to find an appropriate value for n_r. For the AC dataset, two arrays of iRI and rRI values were used in the procedure: 10^4 values ranging from 10^{-5} to 10^{-1} in increments of 10^{-5} for the former and 10^4 values ranging from $1 + 10^{-4}$ to 2 in increments of 10^{-4} for the latter. Values were iteratively extracted from the first array and fed into eq. (6.23) until an optimal value n_i was found and applied to eq. (6.24); the iterative procedure was then repeated with the second array and eq. (6.24) until n_r was similarly determined.

6.3.2.2 AC diameter retrieval

Figure 6.6 shows AC PSDs retrieved using the FC method. These were characterised by a power law background distribution with a superimposed log-normal component corresponding to the main phytoplankton population, suggesting that cultures had significant levels of detritus and possibly bacteria. After selecting for fluorescent particles only, size identification results for the FC method were mixed (Tab. 6.3).

	Particle size (µm)		
Species	FC method	Signal length (20% peak height)	Expected
MS	2-7	n/a	2.5-5.5
SC	1-1.5	n/a	0.5-1.5
HS	3-10	3-40	13-20
AM	10-15	15-30	15-29
KM	3-10 (1-20) *	20-40	15-40
PS	10-20	40-60	61-160 **
SM	5-15	5-50 ***	2-12

Table 6.3 – Sizes as determined by the FC method, as given by particle signal length at 20% of its peak height and as expected for the phytoplankton species of the AC dataset. *Values given for the overall peak and the broader distribution. **Size range includes formae *seriata* and *obtusa*. ***The wide range corresponds to chains of various lengths. References: Tomas *et al*. (1997) (HS, AM, PS); Daugbjerg *et al*. (2000) (KM); Sarno *et al*. (2005); John *et al*. (2002) (MS); Waterbury *et al*. (1979) (SC).

MS and SC size peaks were located at 2-7 and 1-1.5 µm respectively, although *Synechococcus* fluorescence couldn't be isolated from the non-fluorescent background and peak data was taken from its overall size distribution. Both are compatible with expected dimensions for *Synechococcus* and *M. aeruginosa*. HS showed its size peak at 3-10 µm, smaller than the expected dimensions for the smallest species of Heterocapsa such as *H. rotundata* (10-12 µm). AM size peak was located between 10-15 µm, smaller than expected dimensions for *A. minutum* (15-30 µm), while KR showed a broad distribution between 1-20 µm and an overall peak between 3-10 µm, smaller than expected dimensions for *K. mikimotoi* (20-40 µm). PS size peak was located between 10 and 20 µm, substantially less than the expected length but larger than the expected diameter of a single *P. seriata* cell, while SM showed its size peak at 5-15 µm, seemingly encompassing both expected diameter (2-4 µm) and length (10-12 µm) of a single *S. marinoi* cell.

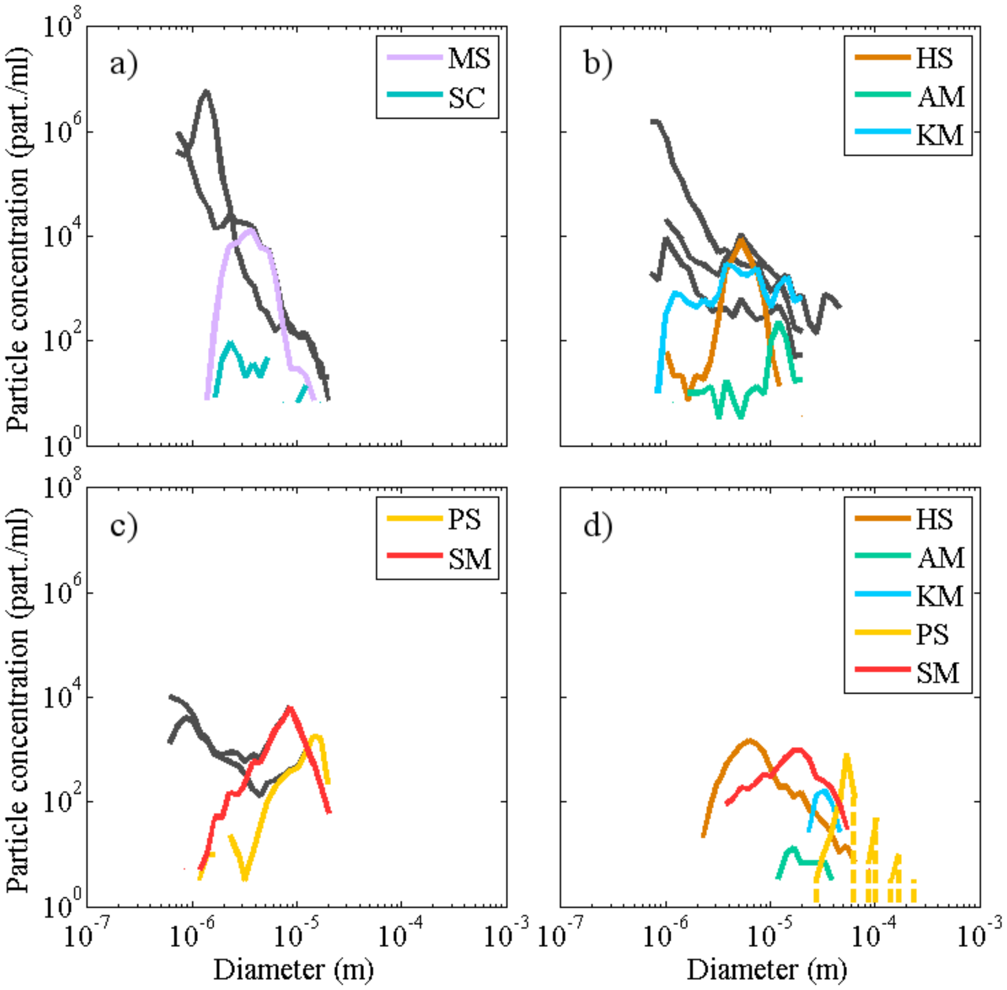


Figure 6.6 – PSDs retrieved by the FC method for (a) small spherical or quasi-spherical species (*M. aeruginosa*, *Synechococcus sp*.), (b) large spherical or quasi-spherical species (*Heterocapsa sp., A. minutum, K. mikimotoi)* and (c) chain-forming species (*P. seriata, S. marinoi)* within the AC dataset: whole PSDs are dark grey, with the fluorescent part of the particle population highlighted in colour. Note that the fluorescent component of the *Synechococcus* sample couldn't be extracted from the background. In the last panel (d), PSDs retrieved by the length method for the same species (MS and SC are too small for the method to apply).

Size identification results obtained from particle signal length taken at 20% of the signal peak height were generally closer to literature values for each species (Tab. 6.3), although only available for the largest species (HS, AM, KM, PS, SM) as per the limitations of the method (Fig. 6.6d). HS showed a wide size peak between 3-40 µm, possibly indicative of a population of detrital particles (e.g. thecae fragments) associated with the main *Heterocapsa* population. AM, KM, SM size peaks were located between 15-30, 20-40 and 5-50 µm respectively, all compatible with expected cell dimensions (and chains of various lengths in the case of *S. marinoi*). PS showed discrete peaks at 40-60, 90-100, ~150 and ~ 210 µm, corresponding to single *P. seriata* cells and two-, three- and four-celled chains respectively but still falling short of reported usual lengths for *P. seriata* cells (Tomas *et al.*, 1997). Size values obtained from particle signal length were further substantiated by direct imaging of the phytoplankton cells (Fig. 6.7). Interestingly, available pictures of *P. seriata* support the finding that cells within the PS sample were indeed shorter than typical cell lengths reported in the literature, possibly indicating that the culture environment (as opposed to a natural one) had influenced the growth of the cells (Fig. 6.7e).

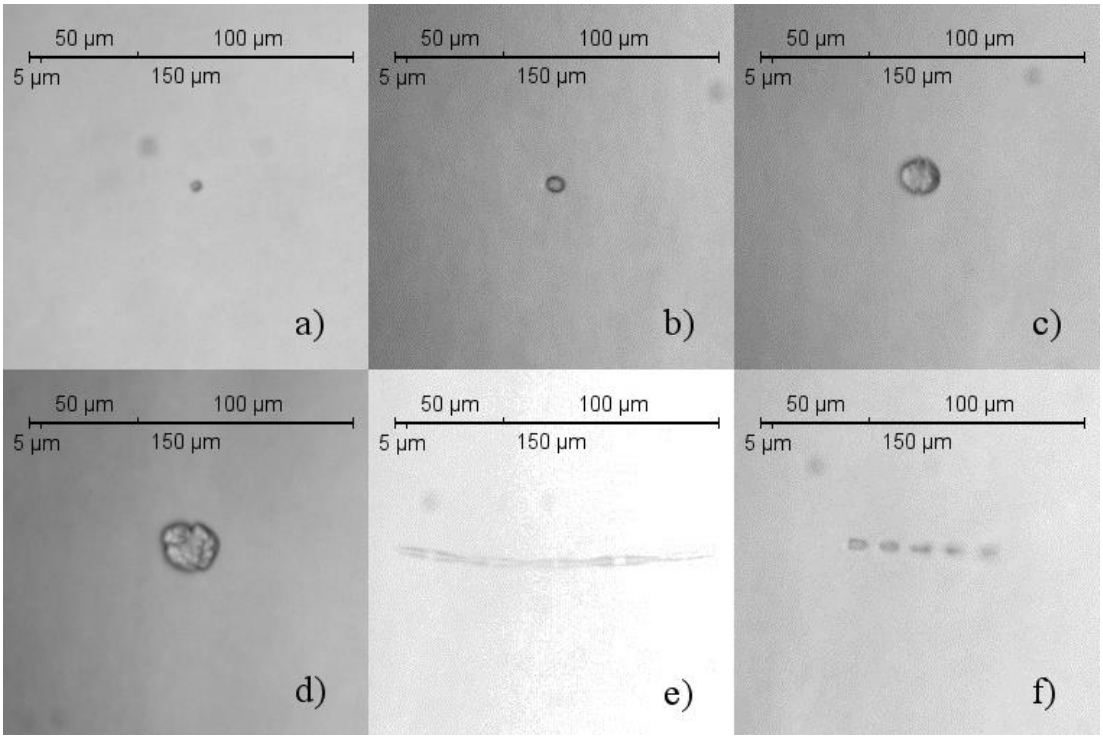


Figure 6.7 – Direct imaging of phytoplankton as retrieved by the CytoSense in-flow camera: (a) *M. aeruginosa*; (b) *Heterocapsa sp.*; (c) *A. minutum*; (d) *K. mikimotoi*; (e) *P. seriata*; (f) *S. marinoi*. *Synechococcus* cells were too small to image successfully. Note that images support the finding that *P. seriata* cells were smaller than typical cell lengths reported in the literature.

6.3.2.3 AC rRI retrieval

Real refractive index determination results were found to follow two distinct trends (Fig. 6.8). The 2014 part of the dataset (HS, AM, PS, SM) displayed a background trend of relatively flat or increasing concentrations towards higher rRIs, with superimposed log-normal distributions in the case of the culture samples. The 2016 part of the dataset (MS, SC, KM) also displayed superimposed log-normal distributions, but on a background trend of decreasing concentrations towards higher rRIs.

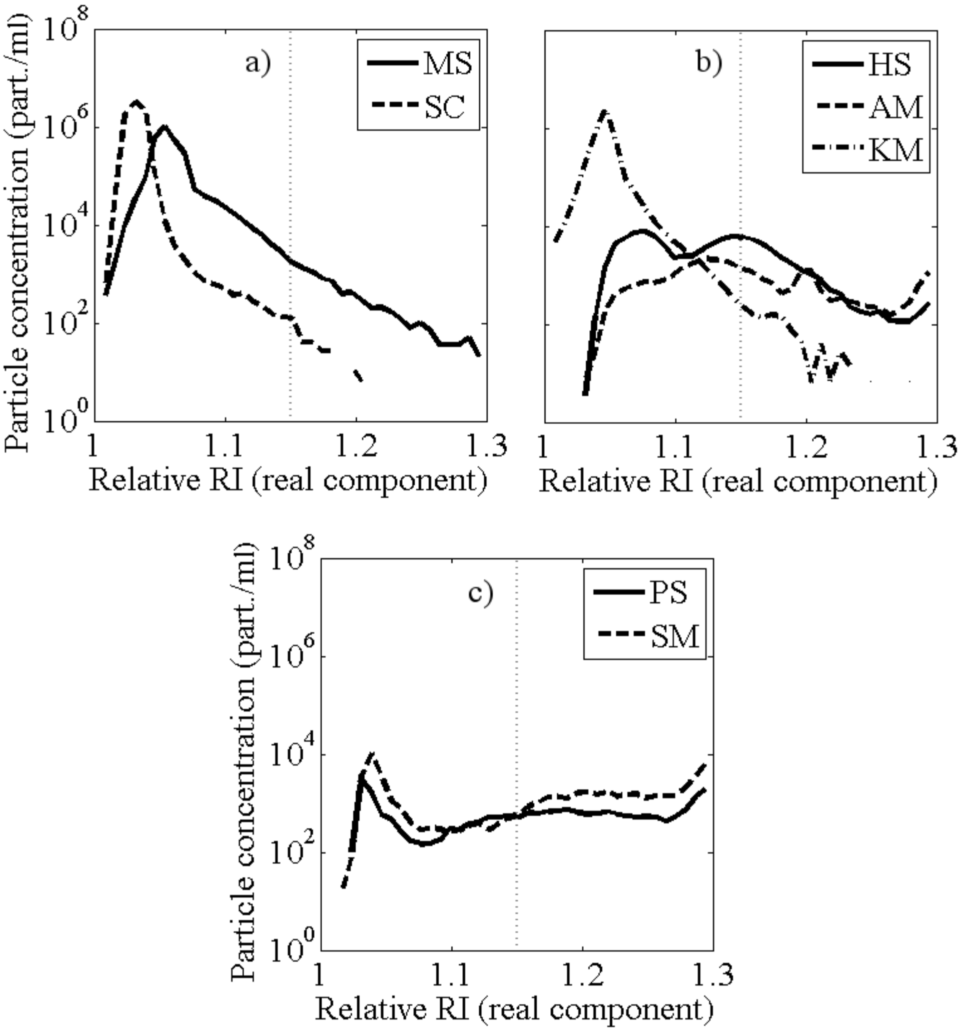


Figure 6.8 – PRIDs retrieved by the FC method for (a) small spherical or quasi-spherical species (*M. aeruginosa*, *Synechococcus sp*.), (b) large spherical or quasi-spherical species (*Heterocapsa sp., A. minutum, K. mikimotoi)* and (c) chain-forming species (*P. seriata, S. marinoi)*. Real refractive index values above 1.15 are not precise, but still indicate high refractive indices.

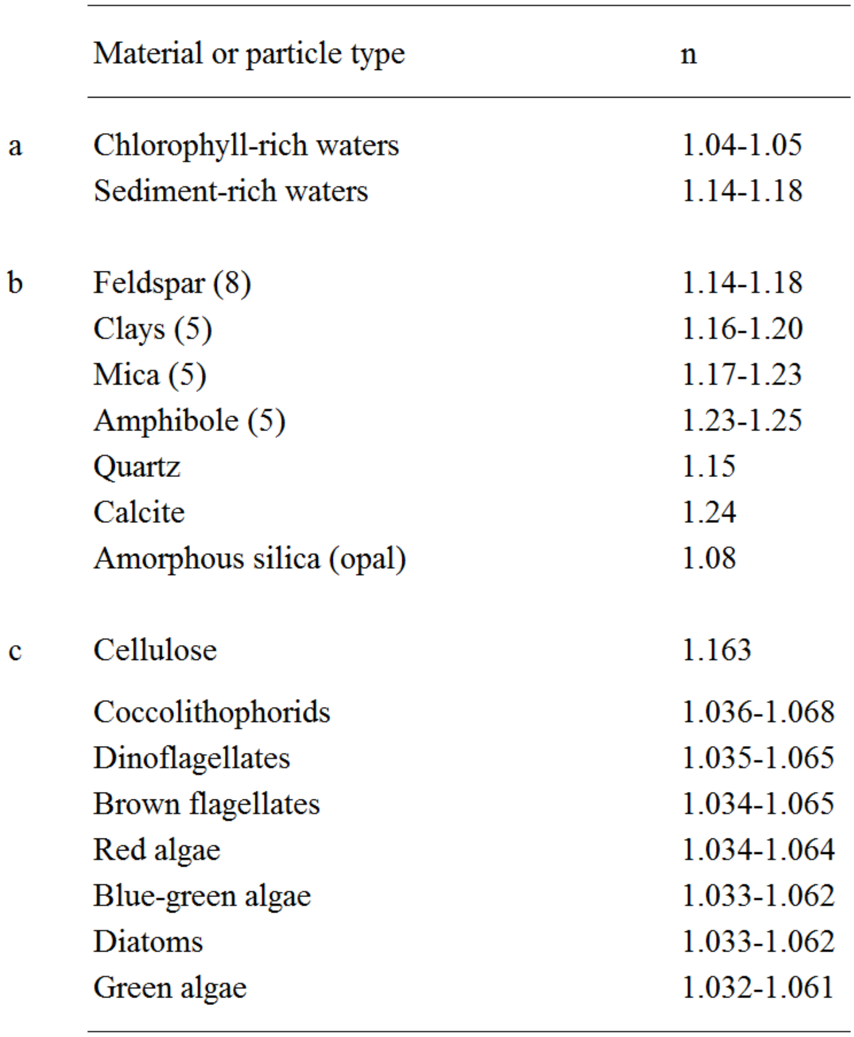
HS showed bi-modal rRI peaks, the first centred between 1.05-1.09 and the second between 1.13-1.16. This lends support to the concurrent presence of two populations, one of *Heterocapsa* cells and one of detritus derived from cellulose thecae. AM showed a single, wide rRI peak between 1.10-1.15, which is higher than expected values for phytoplankton but falls short of values compatible with thecae detritus, while KM rRI peak was located at 1.04-1.06, which is within expected phytoplankton values. All other species considered (MS, SC, PS, SM) were within expected phytoplankton rRI values, with peaks at 1.04-1.07, 1.02-1.04, 1.02-1.04 and 1.03-1.05 respectively. As noted, samples from the 2014 part of the dataset (HS, AM, PS, SM) all displayed large concentrations of particles with rRI values above 1.15. While the validation of the rRI detection capabilities of the method showed that high rRIs values are to be interpreted as generic rather than specific, the concentration spike visible at ~1.3 in particular is made up of particles above and outside the scattering model grid. Under the current implementation of the method these particles are defaulted at the highest rRI considered, i.e. 1.3.

6.3.2.4 ADM bulk rRI comparison

The FC method estimates of peak rRI were generally rather consistent with estimates for bulk rRI derived using the ADM technique (Tab. 6.4) and broadly consistent with values in the literature (Tab 6.5), with some disagreement in the case of SC and AM. The procedure produced an ADM rRI for MS (1.801) which is highly unrealistic; when this outlier was removed, comparison between average FC rRIs and ADM rRIs produced a RMS%E value of 6.627%.

Species	Grid method rRI	ADM rRI
MS	1.04-1.07	1.801
SC	1.02-1.04	1.074
HS	1.05-1.09 1.13-1.16	1.052
AM	1.10-1.15	1.195
KM	1.04-1.06	1.051
PS	1.02-1.04	1.026
SM	1.03-1.05	1.038

Table 6.4 – Real refractive indices as determined by the FC method for the phytoplankton species of the AC dataset and comparison with ADM rRIs. Compare the values of Tab. 6.5.



	Material or particle type	n
a	Chlorophyll-rich waters	1.04-1.05
	Sediment-rich waters	1.14-1.18
b	Feldspar (8)	1.14-1.18
	Clays (5)	1.16-1.20
	Mica (5)	1.17-1.23
	Amphibole (5)	1.23-1.25
	Quartz	1.15
	Calcite	1.24
	Amorphous silica (opal)	1.08
c	Cellulose	1.163
	Coccolithophorids	1.036-1.068
	Dinoflagellates	1.035-1.065
	Brown flagellates	1.034-1.065
	Red algae	1.034-1.064
	Blue-green algae	1.033-1.062
	Diatoms	1.033-1.062
	Green algae	1.032-1.061

Table 6.5 – Tab. 1.1 is recalled here for the reader's convenience, presenting values of the real part of the complex refractive index for some of the main constituents of marine particle populations given relative to water. Reference: a) Twardowski *et al*. (2001); b) from Lide (1997) via Twardowski *et al*. (2001). The number of mineral types included within each class is indicated within parentheses; c) from Aas (1996).

6.3.3 UKCW dataset

6.3.3.1 UKCW diameter retrieval

Fig. 6.9a shows a typical set of FC data for a natural water sample. In all UKCW samples, the vast majority (>99%) of points were found to lie within the area of the grid predicted by Mie theory for reasonable estimates of rRI. Points lying outside the grid exhibit properties that diverge from the assumptions underpinning Mie theory, with likely candidates being morphological characteristics such as strong non-sphericity and presence of facets and vertices.

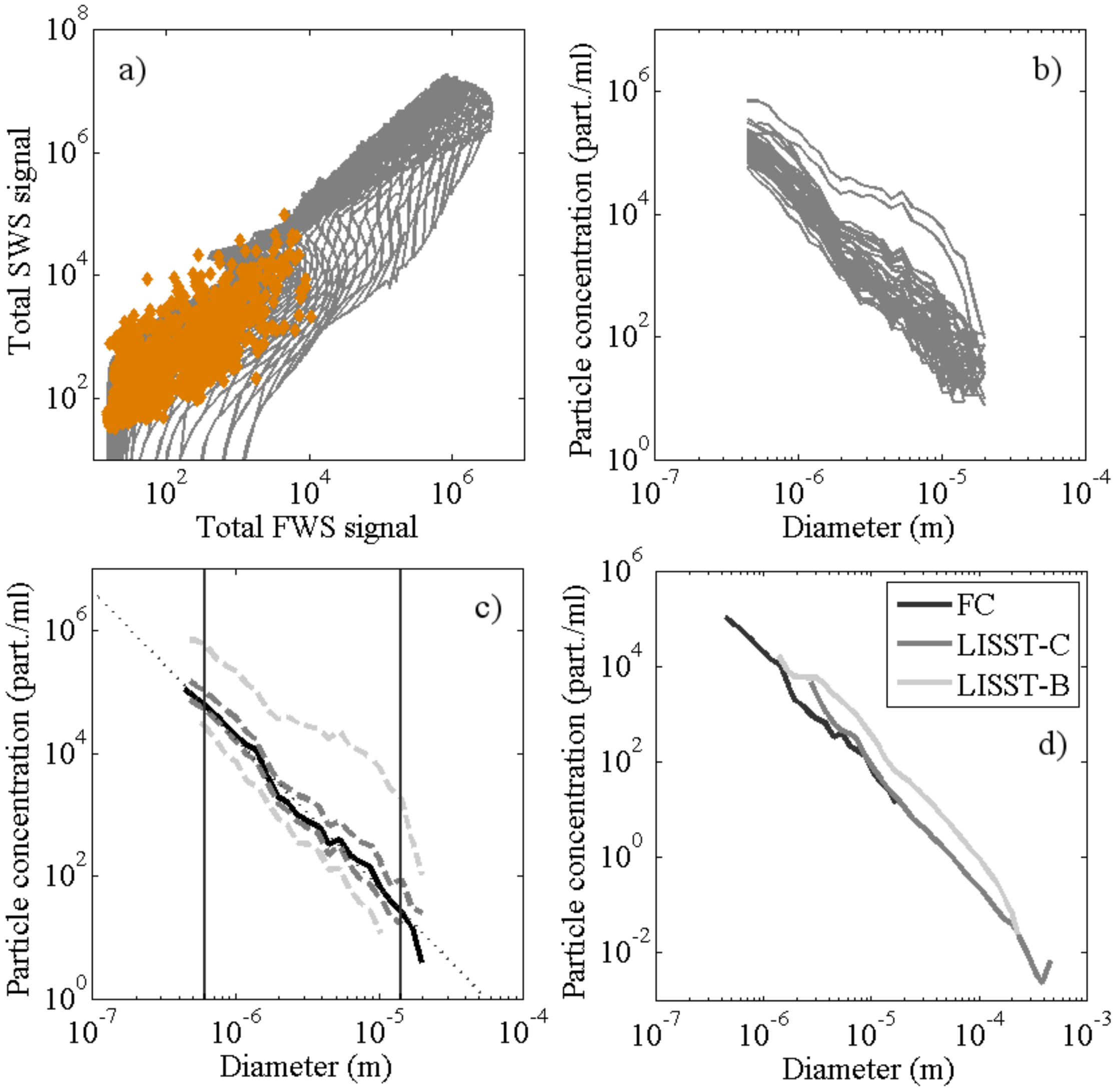


Figure 6.9 – (a) Example of a typical sample from the UKCW dataset superimposed on the FC model grid and (b) collective view of all 50 UKCW PSDs produced by the FC method. (c) The slope for each PSD was calculated with a power law fit over the available range (minus the first and last two data points to avoid possible boundary effects, as indicated by vertical lines), as demonstrated on the median UKCW PSD (quartiles and max./min. given as dark and light grey dashed lines respectively). In the last panel (d), a comparison between the median FC and LISST PSDs.

PSDs determined by the FC method were found to broadly follow power law distributions (Fig. 6.9b). PSD slope values were thus obtained through least squares best fit of power law distributions (cf. Paragraph 1.2.4) as defined by

$$N'(D) = \frac{N(D)}{dD} = kD^{-\gamma} \tag{6.25}$$

after excluding both the first two and last two points in the PSD range to avoid possible boundary effects (Fig. 6.9c). The form given in eq. (6.25) is necessary because while the FC PSDs have bin-like nature, PSD slopes in literature are generally given as those of the underlying density function. Values of the distribution slopes for all 50 samples were found between 2.587-3.813 (mean slope: 3.361 ± 0.250). All PSDs were observed to rapidly lose statistical significance above ~20 μm due to few particle counts and particle concentrations reaching a floor corresponding to bins containing a single particle. PSD data above this size threshold was therefore considered unreliable and subsequently ignored. Overall consistency in the UKCW PSDs was demonstrated when the dataset was reduced to its overall median, quartiles and max/min PSDs (Fig. 6.9c), with the exception of two obvious outlier samples that were collected in the turbid, mineral-rich waters of the Bristol Channel (cf. Fig. 6.9b). These show a clear secondary feature centred on ~ 8 μm that could be associated with strong tidal resuspension of mineral particles in this area.

Fig. 6.9d, showing median PSDs for the entire UKCW dataset for both FC and LISST instruments, demonstrates broad consistency between FC and LISST results. However, a closer look at ratios of corresponding PSD pairs between the three instruments revealed variability both between FC and LISST data and (perhaps more surprisingly) between the two sets of LISST data (Fig. 6.10). LISST-B PSDs were found to be similar to LISST-C PSDs in shape, slope and features, as evinced from the mostly flat profile of their median ratio (Fig. 6.10a), as were the FC PSDs and LISST PSDs when compared with each other (Fig. 6.10b-c). Similarity between FC and LISST PSD slopes was further demonstrated in the shape of the respective distributions of slope values (Fig. 6.11). As with FC PSD slopes, LISST PSD slopes were determined through least squares best fit of power law distributions after neglecting the extremes of the PSD-LISST overlapping size ranges, which were found to produce boundary effects in the PSD ratios (Fig. 6.10b-c). Unlike PSD shape and slopes, particle concentrations were instead found to differ between the FC and the two LISST instruments, with lower overall concentrations for the flow cytometer and median FC/LISST-C and FC/LISST-B values of ~77% and ~32% respectively across the plateau (Fig. 6.10b-c). Surprisingly, particle concentrations were also found to differ markedly between the two LISST-100x instruments, with higher overall concentrations for LISST-B and a median LISST-C/LISST-B ratio value of ~48% across the plateau (Fig. 6.10a).

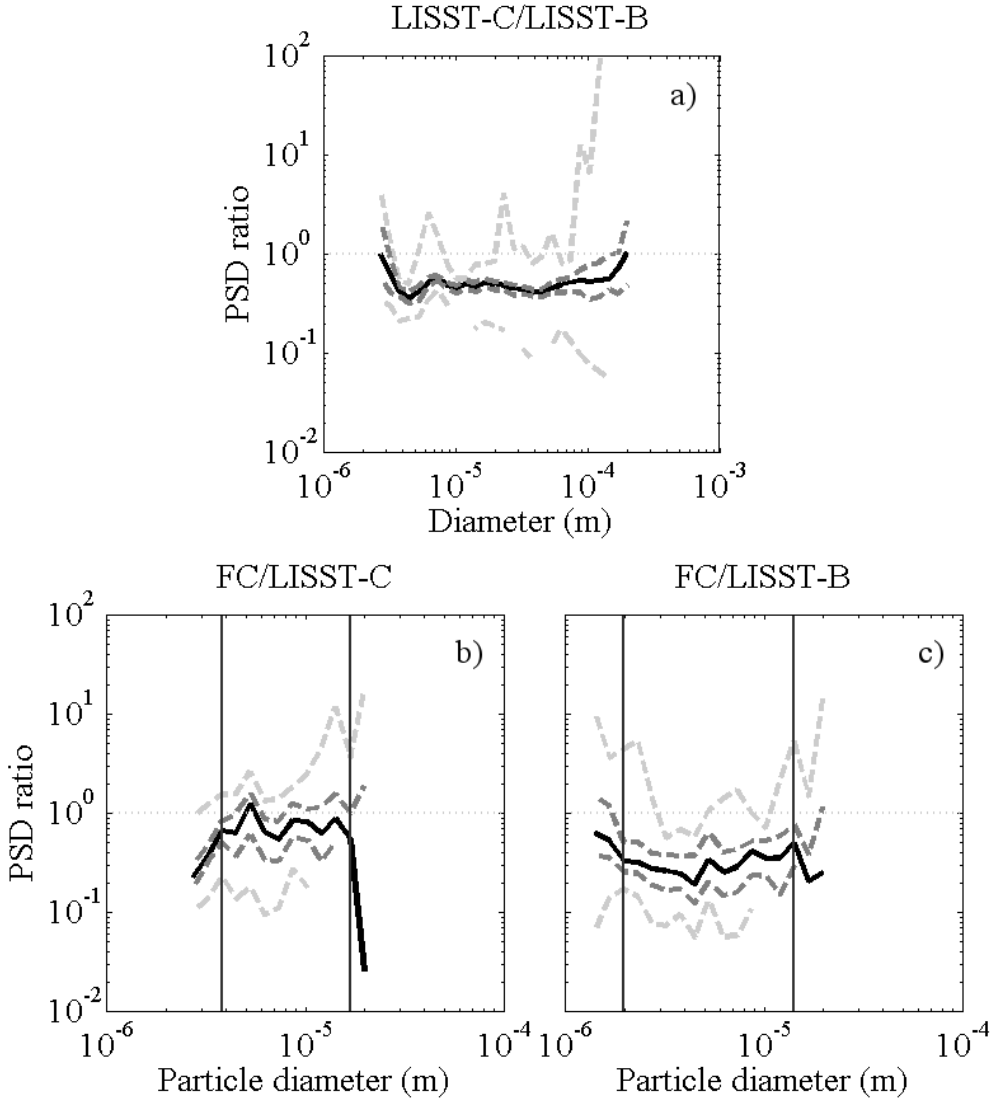


Figure 6.10 – Comparison between PSDs produced by the FC method and the two LISST-100x instruments used in this study, given as the median of all PSD/PSD ratios for the (a) LISST-C/LISST-B, (b) FC/LISST-C and (c) FC/LISST-B relationships (quartiles and max./min. given as dark and light grey dashed lines respectively). The vertical lines indicate the size range over which slopes for the LISST PSDs were calculated.

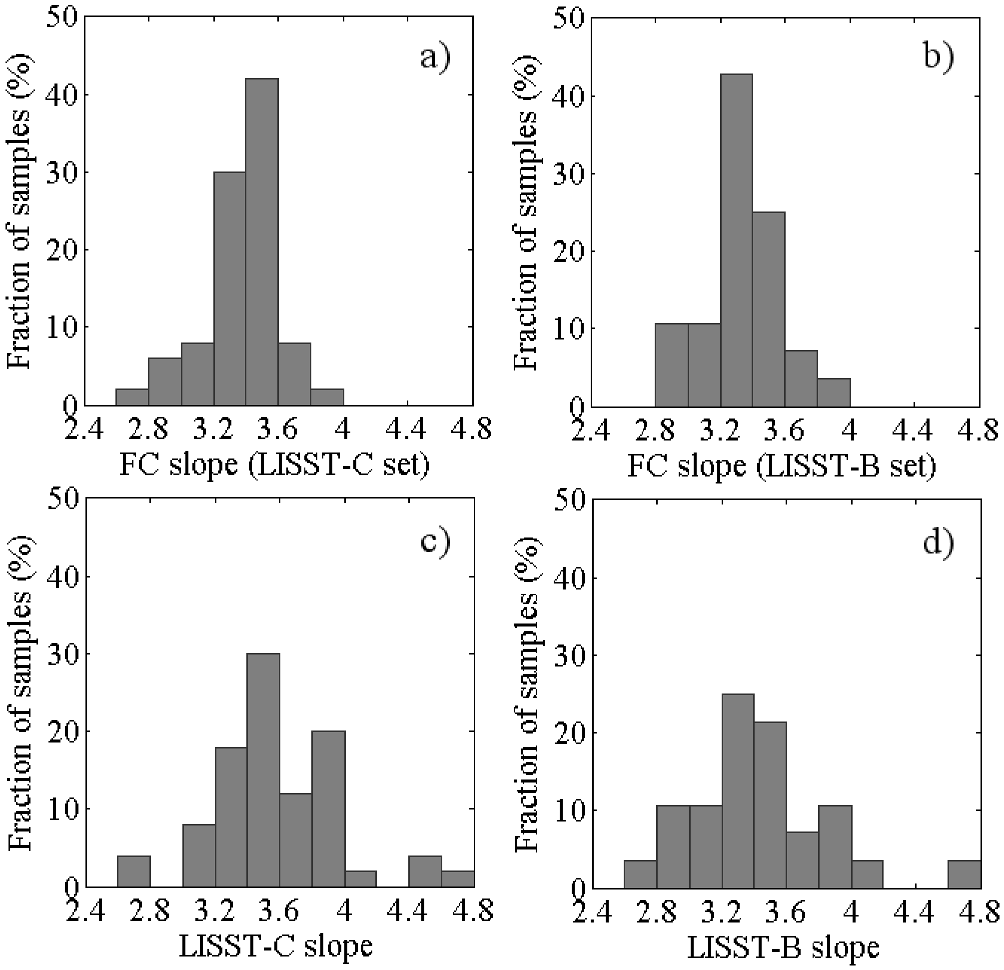


Figure 6.11 – Histograms of FC and LISST PSD slope distributions, paired as FC-LISST-C (a-c) and FC-LISST-B (b-d). Data is presented as two separate subsets because while LISST-C PSDs were available for all 50 samples of the UKCW dataset, LISST-B PSDs were available for only 28 out of those 50.

6.3.3.2 UKCW rRI retrieval

Real refractive index distributions in the UKCW dataset were found to be fairly homogeneous in shape across all samples (Fig. 6.12). Highest median concentrations were found in a broad feature between 1.05-1.15, with the overall peak at 1.07-1.09. This is in accordance with literature values of the rRI of the most common components of marine particle populations (Tab. 6.5). A tail of particles with rRI > 1.15 is also observed in all samples, with median concentrations 5 to 10 times lower than peak ones. As discussed previously, real refractive index values above 1.15 are not to be interpreted as precise, but still positively indicate high refractive indices, and the concentration spike visible at ~1.3 in particular is mostly made up of the particles above and outside the scattering model grid discussed in Paragraph 6.3.3.1. Under the current implementation of the method these particles are defaulted at the highest rRI considered, i.e. 1.3, forming the observed feature (cf. Paragraph 6.3.2.3). As mentioned, these particles were found to account for less than 1% of the total particle population in all samples of the UKCW dataset, and are possibly small, hard inorganic particles with facets and edges for which Mie theory breaks down.

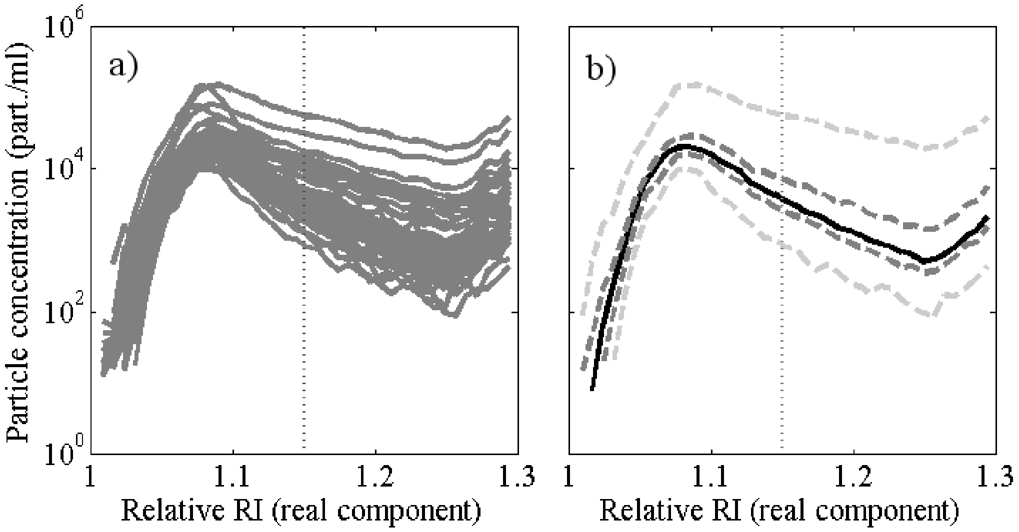


Figure 6.12 – (a) PRIDs retrieved by the FC method for the 50 samples of the UKCW dataset and (b) median of all 50 samples of the dataset (quartiles and max./min. given as dark and light grey dashed lines respectively). Real refractive index values above 1.15 are not precise, but still indicate high refractive indices.

The PRIDs determined by the FC method offer the chance to identify and compare individual subpopulations of particles within the total PSD, and to explore the individual contributions of these fractions to the optical properties of the particle population; these will be described in greater detail in Chapter 7. At the number densities level, separating the particle population of a sample at the rRI = 1.1 threshold (taken as the midpoint between 1.05 and 1.15 as typical organic and inorganic rRIs; see Tab. 6.5) allowed for broad assessment of population dynamics (Fig. 6.13a). Samples collected in the Bristol Channel and the Irish Sea showed a prevalence of inorganic particles, while samples collected west of Orkney and in the North Sea were dominated by organics and samples from Loch Fyne and the

Hebrides displayed a roughly balanced mixtures of organics and inorganics (Fig. 6.13b). These results too will be discussed in further detail in Chapter 7 within the context of individual particle subpopulation contributions to the IOPs.

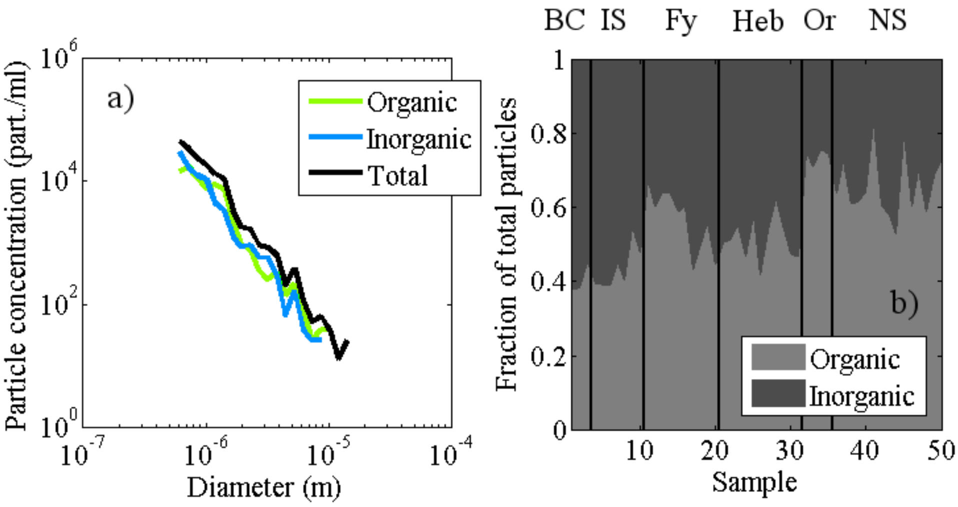


Figure 6.13 – (a) Total, organic (rRI < 1.1) and inorganic (rRI ≥ 1.1) PSDs corresponding to the sample shown in Fig. 6.9a. (b) Organics vs. inorganics balance across the UKCW dataset. The vertical lines reflect the regional groupings of Fig 5.3.

6.4 Discussion

The FC method was found to be capable of identifying diameter and real refractive index within a range of diameters and rRIs in a variety of standard reference samples of polymer beads and oils suspensions (Fig. 6.4-6.5). Diameters and real refractive indices were successfully identified in a range from ~0.5 to ~10 µm for diameter and up to 1.15 for rRIs. Higher rRI values couldn't be identified reliably, but still provided a generic high-rRI signifier as opposed to the more specific low rRI values.

Results obtained using particle signal length at 20% of its peak height on the phytoplankton culture samples of the AC dataset showed that traditional flow cytometry techniques can easily identify the length of particles above the 10 µm threshold (Tab. 6.3); at the same time, the FC method was able to identify diameters below the width of the laser beam, where using signal length becomes unreliable. Furthermore, the FC method allows for the retrieval of refractive index information along with particle size information, and is proposed as perhaps unique in its capability of offering both PSD and PRID determination within a single procedure.

Size determination in phytoplankton culture samples using the FC method was seen to be strongly influenced by shape and likely structure of the phytoplankton cells (Tab. 6.3), as is particularly evident in the extreme case of *P. seriata*. Rounded, smaller cells offer the best results as it is to be expected in the context of Mie theory. Real refractive index detection was found to be much less affected than size detection by the aspect ratio of the particles (Tab. 6.4), even though the tendency for the FC method of underestimating the rRI of the largest particles (as in the case of very large polymer beads) may have biased the rRI results for phytoplankton towards organic-like values.

When applied to natural water samples, the FC method produced PSDs which compared well with the features and slope of independent LISST-100x PSD measurements (Fig. 6.9-6.11). However, there was a lack of consensus on particle concentration between the two LISSTs and the FC (Fig. 6.11). Independent calibration of the latter against particle concentration standards adds weight to the FC data (see Section 3.4.2), but further work is required to fully establish the reliability of concentration data for natural samples. Natural FC PRIDs were found to be largely homogeneous with a broad maximum between 1.05-1.15 and peak values at 1.07-1.09, consistent with a mix of organic and inorganic components (Fig. 6.12). The range of values accounts for most of the common components of marine suspended particulate matter (Tab. 6.5).

A tail of particles with rRI > 1.15 is observed in all UKCW samples, with median concentrations 5 to 10 times lower than peak ones (Fig. 6.12). A concentration spike visible at ~1.3 in particular is found to be made up of particles outside the scattering model grid, which default to the highest rRI considered, i.e. 1.3. These were found to account for less than 1% of the total particle population in all samples of the UKCW dataset, and interpreted as small, faceted, high-rRI particles incompatible with Mie theory assumptions. Due to their relatively small number the influence of particles outside the scattering model grid is negligible. This is also true for particles larger than 10 µm, which were found to be exceedingly few at least in the UKCW samples, although this is likely to be an effect of the measurement protocol described in Chapter 5, which selected for smaller particles. Finding statistically significant numbers of large particles will necessitate a revision of the measurement protocol or at the very least longer acquisition times.

6.5 Conclusion

The results presented in this chapter demonstrate the potential of a FC method capable of simultaneous particle size and refractive index identification, both in controlled samples with monotypic spherical particles (polymer beads, oil suspensions) and in natural water samples containing mixed populations. This implementation of a Mie theory inversion scheme, based on Ackleson & Spinrad (1988), not only enables estimation of PSDs and PRIDs for both algal cultures and natural seawater samples: the resulting ability to partition the PSD by particle type offers new capability to resolve changes in particle population dynamics in natural waters as well. Furthermore, it offers the chance to investigate the links of PSDs and PRIDs to optical properties and the individual contributions of particle subpopulations to the bulk IOPs, as the size range covered by the method (0.5-10 µm) is particularly relevant for optical studies. These topics will be the focus of the next chapter, which is dedicated to Mie-based IOP forward modelling.

7. Forward modelling of inherent optical properties from flow cytometry estimates of particle size and refractive index.

Just as Mie theory was employed in Chapter 6 to retrieve particle diameters and refractive indices from the scattering properties of the particle population, so it can be employed in forward calculations to reconstruct bulk IOPs from the physical properties of the particles. In this chapter, a Mie-based forward modelling procedure is described and used to reconstruct bulk IOPs from the PSDs and PRIDs retrieved by the FC method for the particle suspensions of the AC and UKCW datasets. The first part of the chapter will summarise the procedure, the theory of which has been discussed in Chapter 2. A discussion of the refractive index approximations used in the PSD extrapolations which account for optically significant particles beyond the limits of FC detection is also included. The rest of the chapter will focus on the results of the IOP calculations, discuss their compatibility with the ac-9 and BB9 IOP values presented in Chapter 5 and explore the contributions of individual particulate subpopulations to the total IOPs.

7.1 Introduction

In Chapter 6, sets of rRI and diameter values were used as inputs for Mie theory to simulate scattering within the CytoSense and produce a look-up table of scattering values which allowed for the detection of size and refractive index distributions of particle samples. Much in the same fashion, the PSDs and PRIDs produced by the FC method can be used as inputs for Mie forward optical modelling. Given that the FC-determined PSDs only cover part of the optically relevant diameter range (0.05-2000 μm; Davies *et al*., 2014), extrapolations are needed to account for the particle fraction outside the detection limits of the FC method. Complex refractive indices are then formed by combining the directly estimated PRIDs with imaginary refractive indices adapted from literature for both organic and inorganic components (cf. Paragraph 4.3.1). Once the appropriate PSDs and complex refractive indices are established for the entire optically relevant particle population, the Mie forward model produces volume scattering functions which are integrated to produce scattering and backscattering coefficients, and absorption efficiencies which are used to calculate an absorption coefficient. Modelled IOPs provide insight into the effects of the particle population composition on the optical properties of water and can be validated against independent absorption, scattering and backscattering measurements. The chief focus of the forward modelling procedure is to confirm that the FC method produces PSDs and PRIDs which are compatible with the IOPs of their respective samples, and not to provide alternative means for determining IOPs: therefore, agreement between FC and ancillary IOPs serves as further validation of the FC method as a useful technique for concurrent diameter and rRI retrieval. Nevertheless, such agreement would also represent a step towards optical closure, i.e. the successful agreement of modelled and/or measured parameters between independent measuring techniques. Furthermore, the information contained in the PRIDs determined by the FC method offers the chance to explore the individual contributions of various fractions of the total particle population to the IOPs of a sample, highlighting population dynamics and possibly geographical trends in a dataset.

7.2 Forward optical modelling

As described in Chapter 2, the foundation of Mie forward modelling rests on the additive nature of the inherent optical properties: all IOPs are the result of the sum of the single contributions from each of the individual constituents of a water volume. Section 2.2 provided a detailed treatment of the concepts and equations at the basis of forward optical modelling; a few of those concepts are recalled here and expanded upon in light of the results obtained in Chapter 6. As with the FC method, forward optical modelling calculations were handled using an independently developed MATLAB code which integrates the FASTMie code (Slade, 2006) to calculate Mie coefficients. The core sections of the forward modelling code are presented in Appendix C.

7.2.1 Total particulate IOPs

7.2.1.1 Scattering and backscattering

In Chapter 2 it was shown that for a given density function of the PSD, the axially symmetrical total VSF is expressed as

$$\beta(\theta,\lambda) = \beta_w(\theta,\lambda) + \iint \beta_{D,n}(\theta,\lambda) N'(D,n)\, dD dn\,, \tag{2.44}$$

where $\beta_w(\theta,\lambda)$ is the VSF contribution of pure water, $\beta_{D,n}(\theta,\lambda)$ is the VSF contribution of a single particle of diameter D and complex refractive index n and $N'(D,n)$ the corresponding value of the density function. Given that the ac-9 instruments used in the AC and UKCW datasets are capable of isolating the IOP contribution of the particles from the pure water background, in practice the term $\beta_w(\theta,\lambda)$ can be neglected. Furthermore, it was shown in Chapter 6 that the PSDs and PRIDs determined by the FC method exist not as density functions but as discrete collections of bins. Consequently, for this application eq. (2.44) is more properly expressed as the axially symmetrical total particulate VSF

$$\beta_p(\theta,\lambda) = \sum_{D,n} \beta_{D,n}(\theta,\lambda) N_{D,n}\,, \tag{7.1}$$

where $N_{D,n}$ is the total number of particles within the PSD × PRID bin corresponding to diameter D and complex refractive index n (see the $M_{65\times40}$ matrix described in paragraph 6.2.4). By way of eq. (2.48), eq. (7.1) is finally rewritten as

$$\beta_p(\theta,\lambda) = \sum_{D,n} \frac{1}{2} \frac{i_{1_{D,n}} + i_{2_{D,n}}}{k^2} N_{D,n}\,. \tag{7.2}$$

Total particulate scattering and backscattering coefficients are calculated by integrating this total VSF over the appropriate angle ranges, following eqs. (1.21-1.22). The spectral backscattering ratio is a derived quantity simply defined as the ratio between the backscattering and scattering coefficients, as per eq. (1.23).

7.2.1.2 Absorption

The Mie coefficients a_n and b_n can be used to calculate the efficiency factors for attenuation and scattering following eqs. (2.36-2.37), which in turn define the absorption efficiency factor via eq. (2.38). The absorption efficiency factor can be then combined with eqs. (1.29-1.30) to define the spectral absorption coefficient of a single spherical particle as

$$a_i(\lambda) = \frac{\pi {D_i}^2}{4} Q_{a,i}(\lambda), \tag{7.3}$$

where D_i is the diameter of the particle and $Q_{a,i}(\lambda)$ the absorption efficiency of the particle. Analogously to eq. (7.1), the total particulate absorption is now simply defined as the sum of contributions from all particles,

$$a_p(\lambda) = \sum_{D,n} a_{D,n}(\lambda) N_{D,n} = \sum_{D,n} \frac{\pi {D_i}^2}{4} Q_{a,i}^{D,n}(\lambda)\, N_{D,n}\,. \tag{7.4}$$

7.2.2 PSD extrapolations

Any kind of forward modelling will require the entire optically relevant particle fraction to be included as the input, or the output will not be comparable with any independently measured bulk IOP. In Chapter 6, the CytoSense was found to reliably retrieve diameters between ~0.5-10 µm using the FC method. This range covers a large fraction of the contribution to scattering and backscattering, but the whole optically relevant range spans from tens of nanometres to a few millimetres (Davies *et al.*, 2014). The undetectable fraction of the particle population has to be accounted for using an approximation of the PSD to extend the range of the distribution over the whole relevant range. As discussed in Chapter 1, ever since pioneering work in the ‘60s and ‘70s found that the number of particles suspended in the ocean increased continuously and monotonically towards smaller scales (Bader, 1970; Sheldon *et al.*, 1972), power law distributions of the type used by Junge (1963) for aerosols have been the most common form of approximation for natural seawater particle populations (e.g. Stramski & Kiefer, 1991; Ulloa *et al.*, 1994). The UKCW PSDs determined by the FC method in Chapter 6 broadly conformed to this model (with notable exceptions in the samples retrieved in the waters of the Bristol Channel), and a least squares best fit of power law distributions was used to determine the slopes of the UKCW PSDs. Consequently, following in the steps of Green *et al*. (2003a), a least squares best fit of the measured PSDs through power law distributions as defined by

$$N(D) = N'(D)dD = kD^{-\gamma}dD \qquad (7.5)$$

was used to extend measured PSDs as well. The form given in eq. (7.5) is necessary because the FC PSDs have bin-like nature; accordingly, the extrapolations need to be bin-like as well. This is in contrast with eq. (6.25), where the underlying density function was needed to define a PSD slope.

7.2.3 RI approximations

Values for the rRI in the Junge extensions must also be accounted for using some approximation of the PRID to extend the range of known refractive indices; various assumptions may be used to do so. The most conservative approach (and the one used in this study) is to assign a fixed value to the rRI, either the same on both arms of the extension or a separate one for each. This can be done in a variety of ways (Fig. 7.1), e.g. by using the average rRI of the entire measured fraction of PSD on both arms of the extension (mode A); by using the values of the rRI at the extremes of the measured fraction of PSD on the respective arms of the extension (mode B); or by using literature-derived rRIs based on an hypothesis of probable particle composition in the extension (mode C). The results obtained using each of these different approaches will be detailed in the following. Since the FC method doesn’t offer any information on the imaginary part of the refractive indices, iRI values are unknown both in the available FC PSDs and in their extrapolations. Following the same approach used in Chapter 4, typical values for the imaginary component of the relative refractive index were estimated from Fig. 8 of Babin *et al*. (2003), for both organic and inorganic particles. These were then assigned to the particles according to the value of the real refractive index of each bin, both directly determined by the

FC method and extrapolated. Accordingly, with this approach the values of the imaginary part of the refractive index in the PSD extensions will be ultimately dependent on the rRI approximation used.

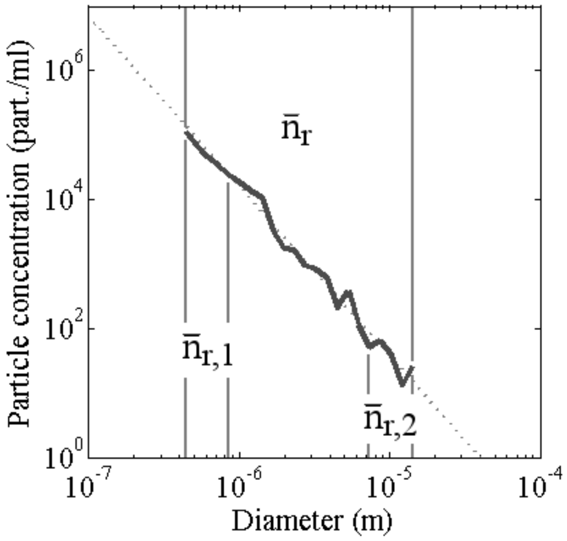


Figure 7.1 – Power law best fit and PSD extension rRI approximations in a typical natural particle population sample. A single rRI value averaged across the PSD range ($\bar{n}_r$) may be used on both upper and lower ends of the extension (mode A), or independent rRI values averaged over the extremes of the PSD ($\bar{n}_{r,1}$, $\bar{n}_{r,2}$) may be used on the respective ends of the extension (mode B). Alternatively, rRI values derived from literature based on the probable composition of the particle population in the extensions may be used (mode C).

7.3 Forward optical modelling application

Mie-based forward optical modelling was applied to the PSDs and PRIDs determined by the FC method for the datasets described in Chapter 5: the AC dataset of algal cultures and the UKCW dataset of natural seawater samples. The resulting IOPs were compared with corresponding absorption, scattering and backscattering coefficients determined by ac-9 and BB9 instruments. In all cases, 532 nm was used as the wavelength of choice for IOP calculations, as it is a green wavelength near the centre of the visible spectrum and one shared by both ac-9 and BB9 instruments without any need for interpolation. In the UKCW dataset, resulting IOPs were then further separated into individual contributions from fluorescent/organic/inorganic fractions and from different size classes.

7.3.1 AC dataset results

7.3.1.1 PSD extrapolation and rRI approximation details

The PSDs of the AC dataset as determined using the FC method were found in Chapter 6 to be heterogeneous, with power law distributions underlying superimposed log-normal distributions corresponding to the population peaks of each cultured species (Fig. 7.2a). Two separate trends were observed, corresponding to the separation between 2014 and 2016 samples. These were most likely caused by differences in the quality of the culture medium, and produced the clear difference in slopes observable in Fig. 7.2a.

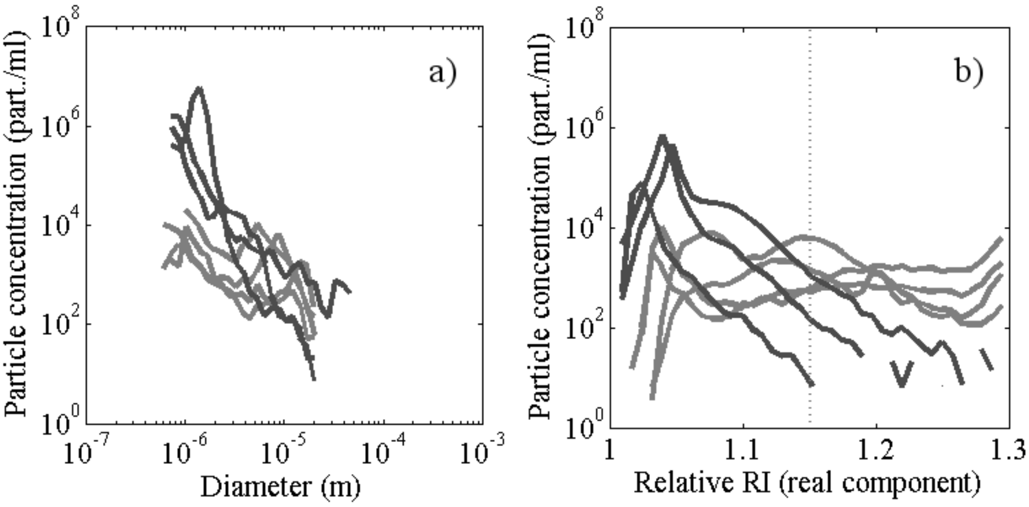


Figure 7.2 - Collective view of (a) AC PSDs and (b) AC PRIDs produced by the FC method. Lighter grey lines correspond to 2014 samples; darker grey lines correspond to 2016 samples. Note that real refractive index values above 1.15 are not precise, but still indicate high refractive indices.

PSDs for the AC dataset were extended using power law extrapolations after an effort to isolate the underlying power law trend by subtraction of the fluorescent component of the PSDs. The extrapolations were carried out between 0.05 and 2000 μm, following the recommendations of Davies *et al.* (2014) to account for distribution slope values as low as 3.5, the lowest value considered in their study. Slope values for the 2014 cultures were however found to be even lower (mean slope value: 2.62); the mean slope value for the 2016 cultures was found to be 4.18. The PRIDs of the AC dataset were similarly found to follow two distinct trends corresponding to the separation between 2014 and 2016 samples, with the latter more strongly peaking towards organic relative rRIs and possibly hinting at a lower overall quality of the culture medium in the 2014 samples (Fig. 7.2b). All three methods of approximation of the rRI in the PSD extensions described in Paragraph 7.2.3 were followed; in particular, when literature values were used, 1.05 was chosen as the relative rRI on both arms of the PSD extension.

7.3.1.2 Forward modelling results

Tab. 7.1 presents the results of the comparison between IOPs as determined by Mie forward modelling for the AC dataset and measured IOP values. Across all three modes of rRI approximation, large species with spherical or quasi-spherical aspect ratio (LS: *Heterocapsa sp.*, *A. minutum* and *K. mikimotoi*) were found to differ greatly from measured values, while small species with spherical or quasi-spherical aspect ratio (SS: *M. aeruginosa* and *Synechocccus sp*.) and (perhaps surprisingly) chain-forming species (Ch: *P. seriata* and *S. marinoi*) generally performed better. Mode B was found as the best performing approximation: corresponding results are shown in Fig. 7.3. The cause of the disagreement is identified in the large overestimation of measured IOP values for species *Heterocapsa sp.* and *A. minutum*. Interestingly, when the PSD extrapolations are taken out of the forward modelling calculations, the values of modelled IOPs for these two species are found to be much closer to measured IOP values with a corresponding improvement in the degree of agreement (Tab. 7.1).

		RMS%E (%)			
		Small Spheres	Large Spheres	Chains	Total
Mode A	a	55.4	1.20E+06	50.5	7.70E+05
	b	69.2	1.20E+05	35.7	7.90E+04
	b_b	50.9	1.30E+04	63.5	8.60E+03
	$\tilde{b}_b$	562.2	65	44.2	304.4
Mode B	a	55.7	1.20E+06	50.5	7.70E+05
	b	71.1	1.20E+05	34.9	8.00E+04
	b_b	74	1.80E+04	50.5	1.20E+04
	$\tilde{b}_b$	189	59.4	29.9	109.4
Mode C	a	55.6	1.20E+06	51	7.60E+05
	b	71	1.20E+05	35.4	8.00E+04
	b_b	72.5	4.10E+04	57.7	2.70E+04
	$\tilde{b}_b$	190	52.4	36	109
No ext.	a	76.8	50.7	39.8	56.9
	b	72.9	34.7	37.1	49.3
	b_b	92.7	72.9	64.9	77
	$\tilde{b}_b$	74.1	63.3	45.1	62.2

Table 7.1 – Results of the comparison between modelled and measured IOPs for the AC dataset. Values are displayed for both the rRI approximation modes A, B and C described in Paragraph 7.2.3 (see also Fig. 7.1) and for the non-extended PSD case.

Care should be taken in the interpretation of RMS%E values in the context of quality of agreement between the modelled and measured IOPs, as although RMS%E values for the non-extended PSDs are lower, inspection of Fig. 7.3 shows general underestimation of the measured IOPs as is to be expected when optically significant particles are excluded from calculations.

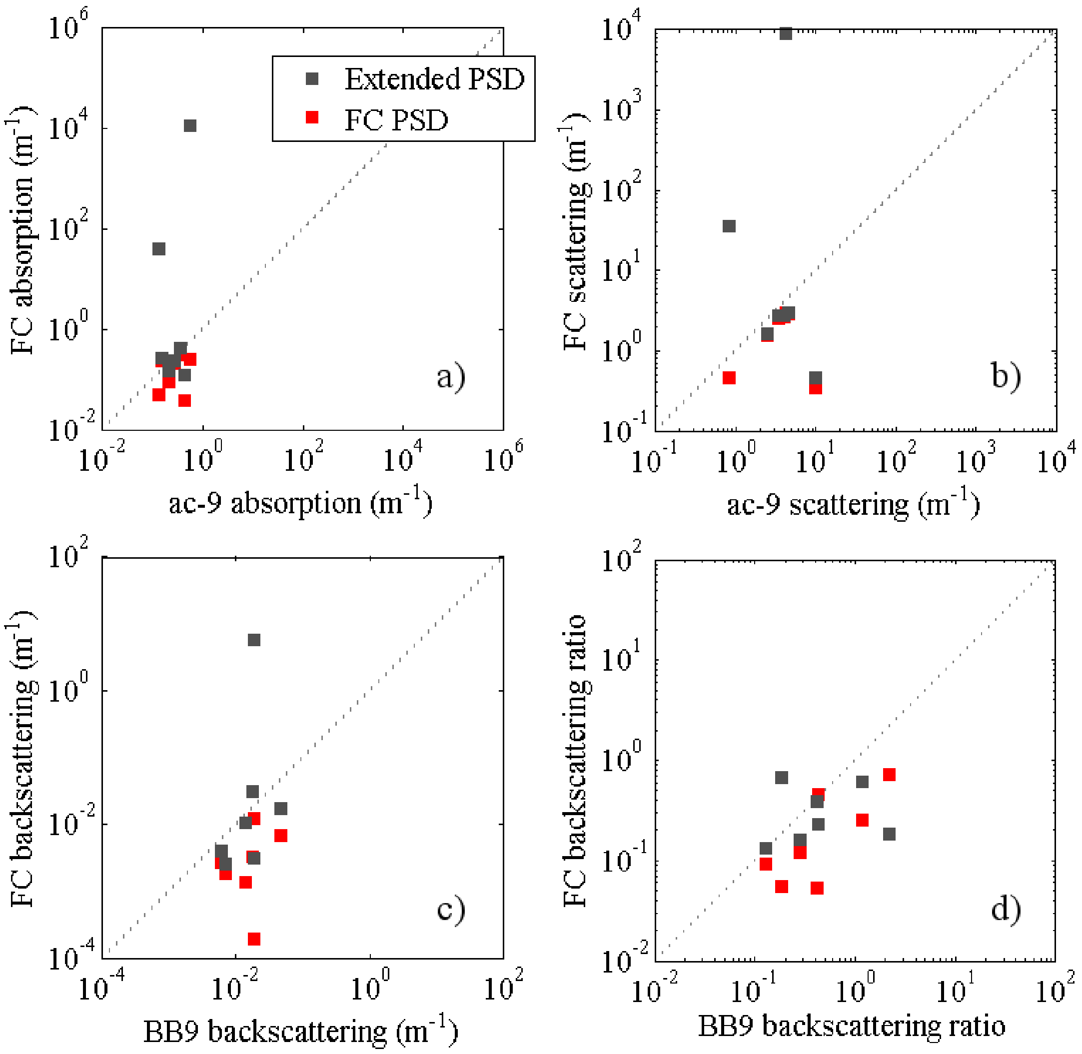


Figure 7.3 – Optical closure analysis of the AC dataset for non-extended PSDs (red squares) and for Mode B of rRI approximation in the Junge extensions (grey squares) for (a) absorption, (b) scattering, (c) backscattering and (d) backscattering ratio. RMS%E values for the closure analysis are given in Tab. 7.1.

Nonetheless, the results for *Heterocapsa sp.* and *A. minutum* indicate that in some cases the PSD extrapolations add more particles than are likely to be actually present. On one hand, it could be argued that such a result should be obvious in culture samples where the largest particles present should logically be the cultured phytoplankton cells; on the other hand, Mie forward modelling for chain-forming species from the same sampling period is shown to perform much better using similar extrapolations. Furthermore, there was reason to doubt the quality of the culture medium, so that the presence of foreign particles in unknown quantities couldn't be ruled out. Overall, the results of Mie forward optical modelling for the AC dataset highlight the risks of using power law extrapolations on samples which, despite seeming to follow power law-like trends, do not represent natural particle populations.

7.3.2 UKCW dataset

7.3.2.1 PSD extrapolation and rRI approximation details

The PSDs of the UKCW dataset were found in Chapter 6 to broadly follow power law distributions, with the main difference between stations being the overall concentration of the particle population (Fig. 7.4a). Two obvious outliers are present, corresponding to samples from the turbid waters of the Bristol channel; close inspection reveals structures that may be closer in nature to models such as the double gamma distribution proposed by Risović (1993), and that indeed may be identified to a lesser degree in the other samples as well. Nonetheless, the power law approach remains a reasonable approximation for a large majority of the dataset, and was used for the PSD extrapolations accordingly. PSDs for the UKCW dataset were extended between 0.05 and 2000 μm following Davies *et al*. (2014); the mean slope value of the UKCW dataset was found to be 3.35. Particle refractive index distributions were found to be fairly homogeneous across all samples (Fig. 7.4b), with distribution peaks found between 1.05-1.15 and within expectations for the rRI of the most common components of marine particle populations. All three methods of approximation of the rRI in the PSD extensions described in Paragraph 7.2.3 were followed; in particular, when literature values were used, 1.15 was chosen as the rRI of particles smaller than 1 μm and 1.05 as the rRI representative of particles larger than 10 μm. A few rare instances occurred of gaps being present in some PSDs at ~10 μm when overall particle concentrations were low. This caused the loss of statistical significance caused by low particle counts to occur earlier and sometimes produce empty size bins. Where these gaps were present in the PSDs, a value of 1.1 was used for the power law extrapolation between 1-10 μm.

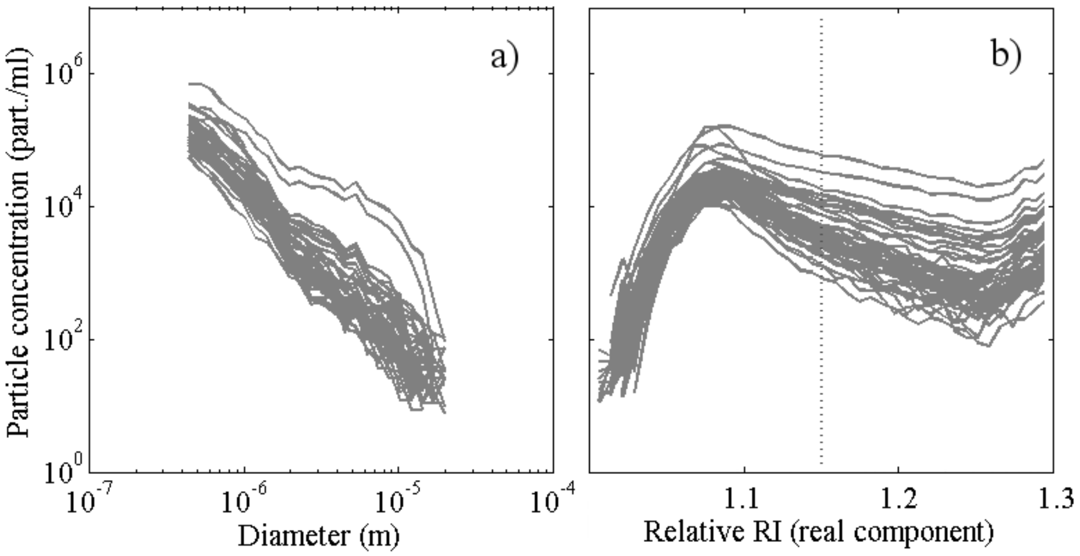


Figure 7.4 – The collective view of all 50 (a) UKCW PSDs and (b) UKCW PRIDs produced by the FC method is recalled here from Fig. 6.9 and 6.12. Note that real refractive index values above 1.15 are not precise, but still indicate high refractive indices.

7.3.2.2 Forward modelling results

Tab. 7.2 presents the results of the comparison between IOPs as determined by Mie forward modelling for the UKCW dataset and measured IOP values. Mode B was found to provide the best overall retrieval of IOPs, and corresponding results are shown in Fig. 7.5.

	RMS%E (%)		
	Mode A	Mode B	Mode C
a	115.5	118.1	120
b	34.3	35.3	35
b_b	60.1	44.5	48.9
$\tilde{b}_b$	112.7	77	94.3

Table 7.2 – Results of the comparison between modelled and measured IOPs for the UKCW dataset.

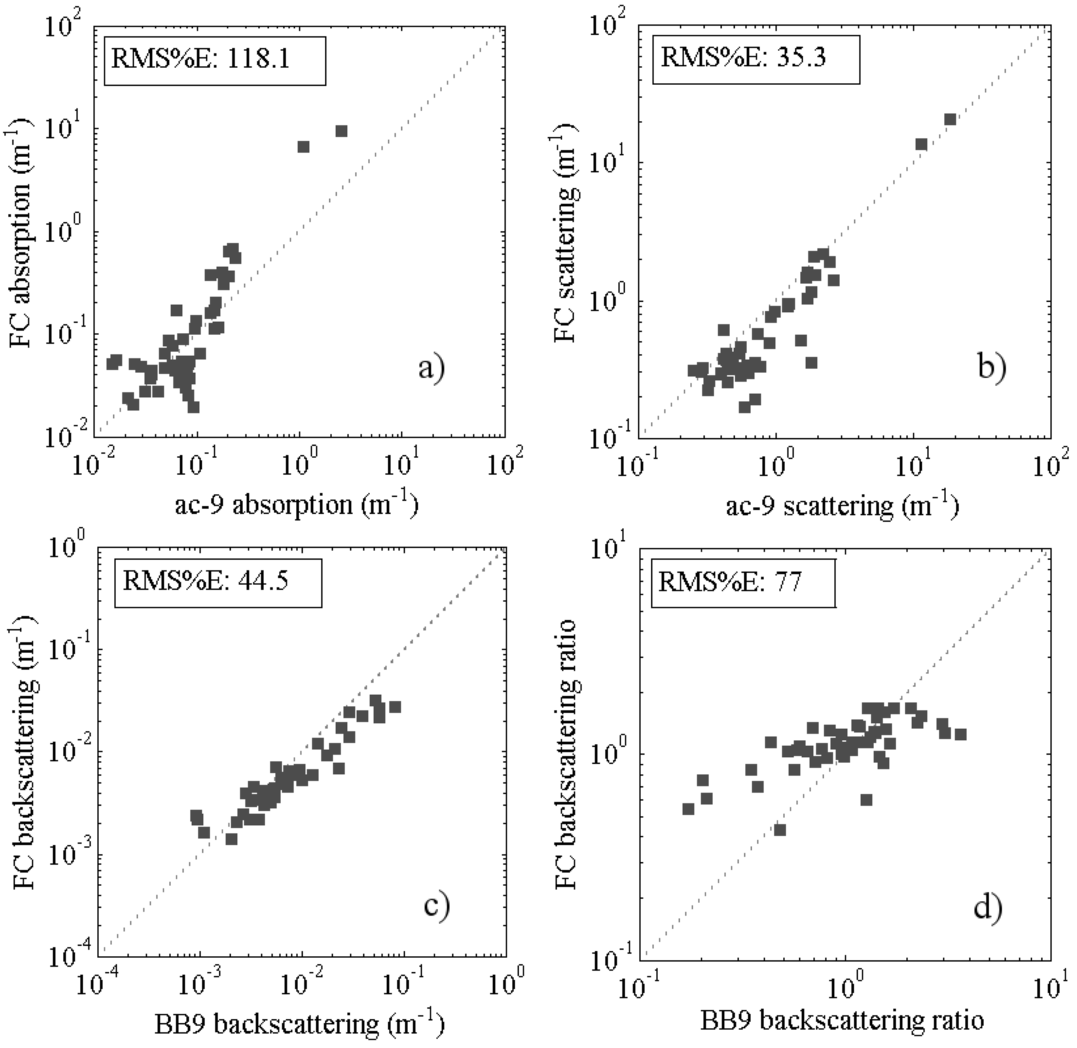


Figure 7.5 – Optical closure analysis of the UKCW dataset for Mode B of rRI approximation in the Junge extensions for (a) absorption, (b) scattering, (c) backscattering and (d) backscattering ratio.

Retrieval of IOPs was found to be variable, with trends to either underestimate or overestimate at higher values for each IOPs. Nonetheless, overall agreement is rather good for absorption, scattering and backscattering, with the backscattering ratio demonstrating the negative impact of compounding errors by taking ratios. When considering the sample-by-sample ratio of modelled FC IOPs vs. measured IOPs, different behaviours can be observed for each of the IOPs. FC absorption displayed large variance across the entire dataset (Fig. 7.6a), while FC scattering was generally lower than ac-9 scattering, with disagreement progressively increasing in later samples which displayed relatively high chlorophyll-A content (see Chapter 5) (Fig. 7.6b). This could represent a breakdown in the performance of the ac-9 iterative correction used to correct ac-9 data (Paragraph 5.3.3) under algal bloom conditions, possibly as a consequence of deviations from the Fournier-Forand phase function that underpins the method (McKee *et al.*, 2013).

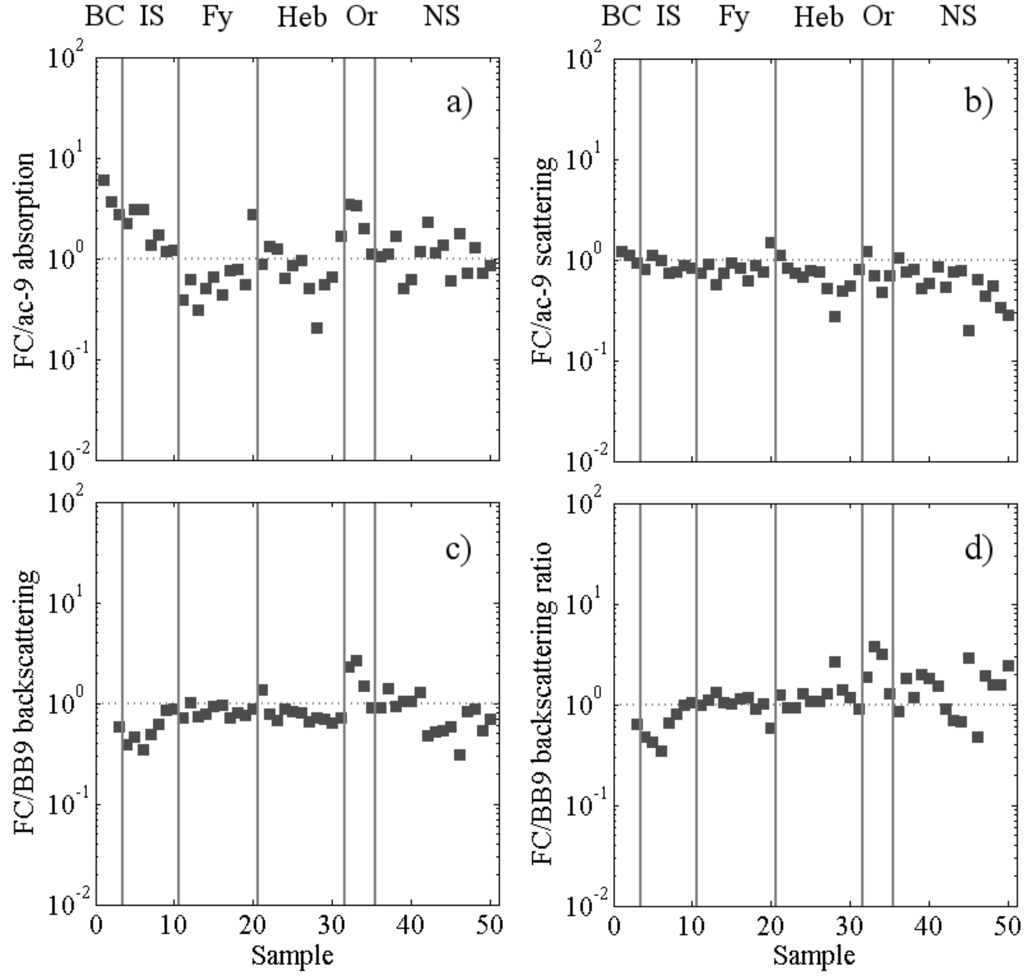


Figure 7.6 – FC vs. ac-9-derived IOP ratios for (a) the absorption coefficient and (b) the scattering coefficient and FC vs. BB9-derived IOP ratios for (c) the backscattering coefficient and (d) the backscattering ratio. The vertical lines reflect the regional groupings presented in Fig. 2. Note that the first two data points for BB9 backscattering and backscattering ratio are missing because of saturation in the BB9 backscattering meter.

FC backscattering was generally lower than BB9 backscattering, with larger disagreement in the Irish Sea, Orkneys and North Sea samples (Fig. 7.6c). FC backscattering ratios (Fig. 7.6d) generally replicated the pattern displayed by the backscattering data (Fig. 7.6c). For b, b_b and $\tilde{b}_b$, samples from the Firth of Clyde and the Hebrides were seen to produce results which were consistently closer to the validation IOPs.

The capability of the FC method to determine the size and real refractive index of each particle, combined with the ability to detect fluorescence, enabled the analysis of the individual contribution of fluorescent and non-fluorescent organic and inorganic fractions of the particle population to the IOPs, although limited to the effective size detection range of the method. Across the 50 samples in the UKCW dataset, inorganic particles were seen to account for particle population fractions ranging from 16.6% to 62.2% of the total, with an average value of 43% (Fig. 7.7a). Samples from the Bristol Channel and the Irish Sea displayed the largest inorganic content, while later samples collected west of Orkney and from the North Sea were generally dominated by organic particles. Fluorescent organic particles were found to account for a small fraction of the particle population (0.1-15.6%), lower than 5% in most samples. The organic fraction was seen to have a strong influence on both absorption and scattering, with fraction contributions between 18.9-93.7% and 23-82.6% of the total respectively across the UKCW dataset (Fig. 7.7b-c). Lowest organic contributions came from the Irish Sea and (particularly) from the Bristol Channel. Fluorescent organic content was also observed to have significant impact on a_p and b_p despite its small population fraction, with contribution values as high as 73.9% and 53.8% respectively and generally higher than 20% in a large number of samples. Once again, the Irish Sea and Bristol Channel samples (where the fluorescent content was found to be numerically negligible) displayed the lowest fluorescent fraction contribution to both IOPs. In contrast, backscattering was found to be largely dominated by the inorganic fraction of the particle population, with values ranging from 63.3% to 93.1% of the contribution (Fig. 7.7d).

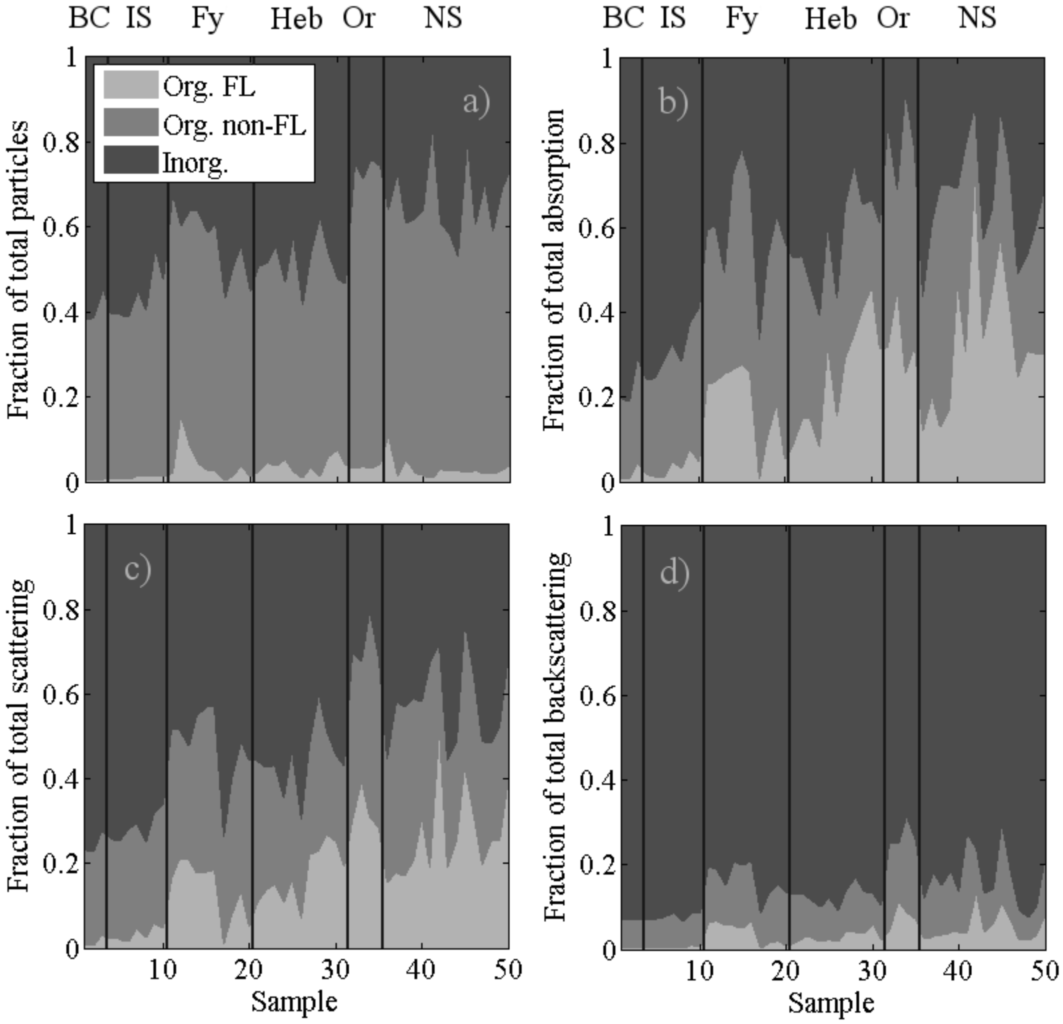


Figure 7.7 – (a) Relative abundance of the inorganic and organic (fluorescent and non-fluorescent) fractions of the total particle populations, and contribution of each of these population fractions to (b) total absorption, (c) scattering and (d) backscattering coefficients.

The reason behind the different observed fractional IOP contributions rests in the different response that the IOPs have to particles of the same size, as demonstrated when the individual IOP contribution of the measured FC fraction of the PSD is compared to those induced by the upper and lower PSD extensions (i.e. large and small particles respectively). Absorption *a* (Fig. 7.8a) was found to be dominated by large particles above ~ 10 μm (27.1-98.1%), with only minor contribution from particles smaller than ~0.5 μm (0.1-12.1%). Scattering *b* (Fig. 7.8b) was found to be overwhelmingly influenced by particles larger than ~0.5 μm (92.1-99.9% of the contribution), with a non-negligible contribution from particles larger than ~10 μm

(3.3-80.6%). Finally, backscattering b_b (Fig. 7.8c) was found to be largely influenced by particles smaller than ~10 μm (53.4-99.4% of the contribution), with a sizeable influence from particles smaller than ~0.5 μm (1.4-49.9%).

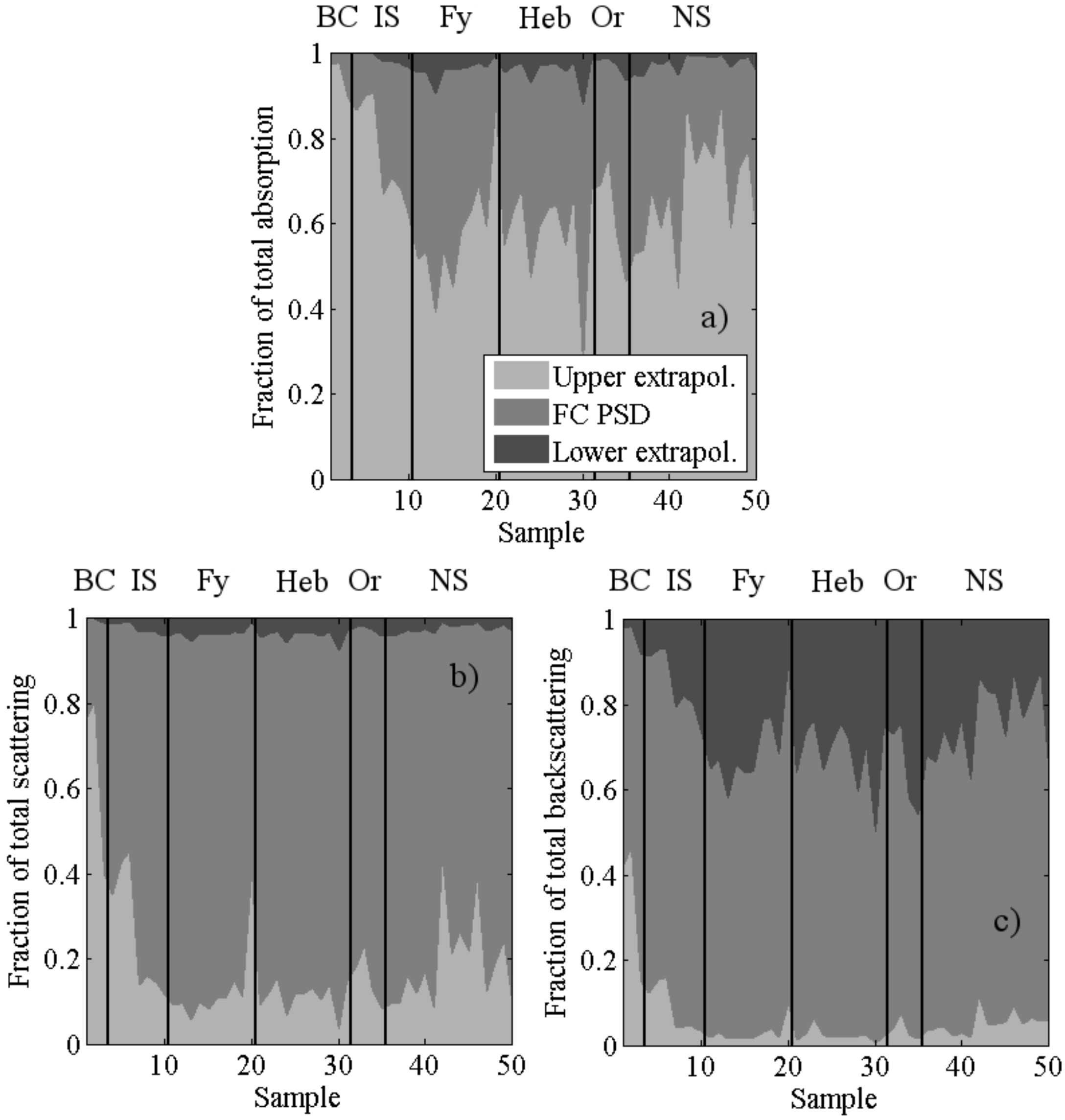


Figure 7.8 – Comparison of the relative contributions of the FC PSD and of its upper and lower power law extensions (i.e. large and small particles respectively) to (a) total absorption, (b) scattering and (c) backscattering coefficients.

Considering these results, the large inorganic contribution to the total backscattering shown in Fig. 7.7d may be interpreted to suggest that small particles in the UKCW dataset were mostly of inorganic nature; it should be noted however that the lower diameter detection limit of the FC method is slightly higher for organic particles than it is for inorganic particles, because given equal diameters a higher refractive index produces a higher side scattering signal. Therefore, organic number densities fall off slightly earlier than inorganic ones as diameters approach the detection limit. Although all care was taken to minimise this spurious effect, it is likely that some part of the b_b fraction of Fig. 7.7d was caused by it.

7.4 Discussion

The results of closure analysis between modelled and measured IOPs in the AC dataset were found to vary greatly between algal species, with large, roughly spherical species *Heterocapsa sp.* and *A. minutum* found to largely overestimate corresponding measured IOPs while small-rounded and chain-forming species achieved good agreement across all IOPs considered (Tab. 7.1). The disagreement was diminished greatly when the PSD extrapolations were excluded from IOP calculations, suggesting that the power law extensions introduced more particles than were actually present in the samples; at the same time, not only other species achieved better agreement using similar PSD extrapolations, but removing the contribution of undetected particles from the IOP calculations caused general (if not large) underestimation of the measured IOPs (Fig. 7.3). A definitive interpretation of the nature of the problem is difficult, although it appears clear that the application of a PSD approximation designed for natural particle populations to culture samples is risky, and produces unpredictable results.

The best results for the closure analysis between measured and modelled IOPs for the UKCW dataset were achieved for scattering and backscattering coefficients. In the latter case, more refined approximations for the rRIs used in the Junge extensions of the PSDs (i.e. rRI averages based on available data as opposed to fixed values adapted from literature) markedly improved the quality of the agreement, an effect consequently also reflected in backscattering ratio values. Agreement between FC and ac-9 absorption coefficients remained the least successful across all three rRI approximation used, perhaps as expected for a quantity which is chiefly influenced by imaginary refractive indices. The FC method of flow cytometric determination of size and refractive index used in this study offers no information on the imaginary component of the refractive index, so that no better solution than adapting literature values to the organic/inorganic fractions of the particle population could be used.

Given the relatively narrow size range that the FC method addresses, the need to extrapolate both size and RI values, and the absence of direct observations of iRI, the degree to which forward modelled IOPs match measured values is surprisingly good. Variation in the quality of match-ups across the dataset can be attributed to a number of interwoven factors. For example, selection bias in favour of small particles potentially affects absorption retrieval more than scattering and backscattering, as Fig. 7.8a demonstrates the relatively strong contribution from the upper extrapolation. Fig. 7.6 indicates that there may be regional factors at play, probably corresponding to changes in the composition of the particle population. It is worth noting that it is quite likely that such changes might also influence the performance of scattering correction procedures for ac-9 absorption and attenuation measurements, and therefore influence the degree of compatibility between the measured IOPs and IOPs derived from Mie forward modelling. There is scope to extend the FC sampling procedure to capture a wider range of particle sizes at the high end and to improve statistical significance throughout. However, it is clear that the current iteration of the FC method as it was presented in Chapter 6 already provides a useful insight into the contribution of different particle components to the formation of bulk particulate IOPs.

The overall broad agreement between measured and modelled IOPs provides useful validation of the PSDs and PRIDs generated by the FC method. In doing so, it also suggests that Mie theory is sufficient to provide useful estimates of bulk IOPs for natural particle populations. This is not to say that Mie theory is capable of accurately predicting optical properties for all marine particles; indeed, there are many well-established situations where more complex optical models are required (see Paragraph 1.1.2). Rather, results indicate that Mie theory, which is analytical, conservative and computationally fast, has the ability to reproduce bulk IOPs for randomly orientated, mixed populations of naturally occurring marine particles. In this context, it is interesting to evaluate the results of the optical closure analysis for the AC dataset, in particular for the chain-forming species *P. seriata* and *S. marinoi*. Surprisingly, these species were found to produce the best agreement between modelled and measured IOPs. While this may be easier to justify in the case of *S. marinoi*, the cells of which are structurally complex but have ellipsoidal aspect ratio, it is a remarkable result for *P. seriata*, which has spindle-shaped individual cells possessing rather extreme aspect ratios. This is compounded by the dynamics of FC measurements, which, being carried out within a flow system, tend to arrange long particles along the direction of the sample flow. A possible interpretation of such results would be to suggest that the nature of the FC method makes it so that the diameter and rRI values produced by it correspond to spheres optically equivalent to the particles processed by the flow cytometer; the modelled IOPs therefore would be compatible with bulk IOP measurements taken by instruments which observe ensembles of randomly orientated particles. However, more data is required to substantiate this interpretation.

Finally, it is important to consider that *in situ* IOPs used to validate the forward modelling are themselves subject to variable degrees of error. Forward modelled scattering values, resolved down to 0.01° in the forward direction, provide useful validation of (or at least demonstrate mutual consistency with) *in situ* values generated using iterative correction of the scattering error in the ac-9 attenuation measurements. Similar degrees of closure with uncorrected *in situ* b_p data is possible by restricting calculations to angles beyond the transmissometer collection angle but are no better in quality. This is good circumstantial evidence that the iterative scattering correction is a useful development. On the other hand, as mentioned previously, the iterative correction relies on selection of appropriate scattering phase functions and there is evidence (Fig. 7.6b) that there is possibly some degree of breakdown in performance as waters become increasingly dominated by large phytoplankton.

7.5 Conclusions

PSDs and PRIDs determined using the FC method for both algal cultures and natural seawater samples were used as inputs to Mie forward optical modelling. With the selection of appropriate PSD extrapolations and rRI approximations, which accounted for undetected particles, the calculations produced IOP values which were then compared with corresponding measured IOPs. AC dataset results were mixed, highlighting the unreliability of using PSD extrapolation models developed for natural samples on non-natural samples like algal cultures. The good agreement reached with modelled IOPs of chain-forming algae may also indicate that the FC method produced PSD and PRID results actually representing equivalent spheres. UKCW dataset IOP match-up results were found to be variable, reflecting geographic variability and possibly limitations in the quality of *in situ* IOPs used for validation. However, given the relatively narrow FC method size detection range, the need to extrapolate PSDs and approximate rRIs to account for undetected particles, and an inherent lack of direct iRI measurements, the degree of agreement between modelled and measured IOPs is surprisingly good. These results validate the PSDs and PRIDs produced by the FC method, suggest a reassessment of the value of Mie theory in the context of predicting bulk IOPs for natural particle populations and underline the requirement to consider the role of measurement uncertainties, particularly systematic errors, when using *in situ* IOPs for validation of forward optical modelling. Finally, PSDs and PRIDs as determined by the FC method offered the chance to gain useful insight into the individual contributions of different particle subpopulations to the bulk IOPs of seawater samples; in the next and final chapter, a similar approach will be taken to investigate to what degree the same PSDs and PRIDs can be used to successfully inform models of particle density and of organic carbon and chlorophyll content, and to explore the biogeochemistry of the particle populations of the UKCW dataset.

8. Biogeochemistry

8.1 Introduction

Knowledge of the size and refractive index distributions within a particle population not only allows for IOP forward modelling and for an assessment of the relative contribution of the particle fractions to the optical behaviour of water; coupled with models of particle density and carbon content, PSDs and PRIDs offer the chance to explore the biogeochemical properties of a particle population as well. In this chapter, models of organic and inorganic particle density are adapted from literature together with information on carbon and chlorophyll-A cell volume scaling to produce modelled values of total suspended matter (TSM), organic suspended matter (OSM), inorganic suspended matter (ISM), particulate organic carbon (POC) and chlorophyll-A content (ChlA) from the particle data of the UKCW dataset. The modelled values thus obtained are then compared against the results of actual biogeochemical measurements, and the parameters used to assess carbon and chlorophyll content are optimized on a dataset-wide basis to explore the physiology of the cells encountered during the HE442 research cruise.

8.2 Data inputs and UKCW comparison data

It is convenient to recall here the biogeochemical data presented in Chapter 5 for the UKCW dataset (Fig. 8.1); TSM, OSM, ISM, POC and ChlA measurements taken during the HE442 research cruise will be used in the following sections as comparison for the results of the modelling. Accordingly, PSDs and PRIDs as determined in Chapter 6 by the FC method for the UKCW samples will be used as inputs for the biogeochemical models.

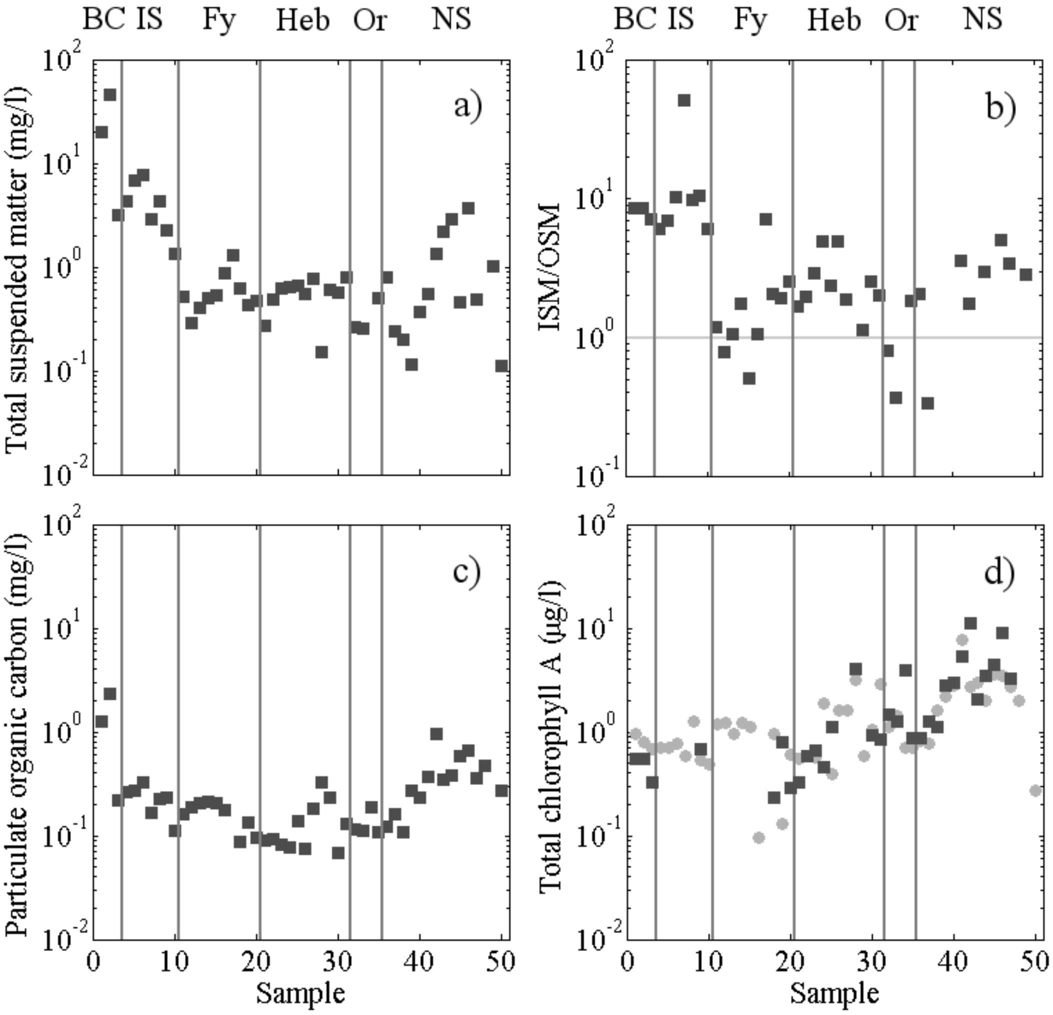


Figure 8.1 – (a) TSM, (b) ISM/OSM balance, (c) POC and (d) ChlA data for the UKCW dataset. In the last panel, grey circles represent HPLC-derived ChlA values and dark grey squares represent ChlA estimates given as 5.5 times the initial fluorescence F_o of the sample. The vertical lines reflect the regional groupings presented in Chapter 5: Bristol Channel (BC), Irish Sea (IS), Loch Fyne and Firth of Clyde (Fy), Hebrides and Skye (Heb), North Atlantic - Orkneys (Or) and North Sea (NS).

By exploiting the particle composition information given by the PRIDs, the total PSD of each UKCW sample is further separated into an inorganic PSD, an organic PSD and a fluorescent fraction, which is itself a sub-fraction of the organic PSD (Fig. 8.2a). These will be respectively used to calculate OSM & POC, ISM and ChlA, while TSM will be calculated both from the total PSD and as the sum of the model ISM and OSM. To account for the fraction of the particle population undetected by the FC method, the total, organic and inorganic PSDs must be extended following the methodology used in Chapter 7 (Fig. 8.2b); note that since the PSD extrapolations only approximate the particle populations outside the FC method detection range, the sum of the extended organic and inorganic PSDs generally is not exactly equal to the extended total PSD, making the dual calculation of TSM from both the total PSD and as the sum of model ISM and OSM values a useful check of the error thus introduced.

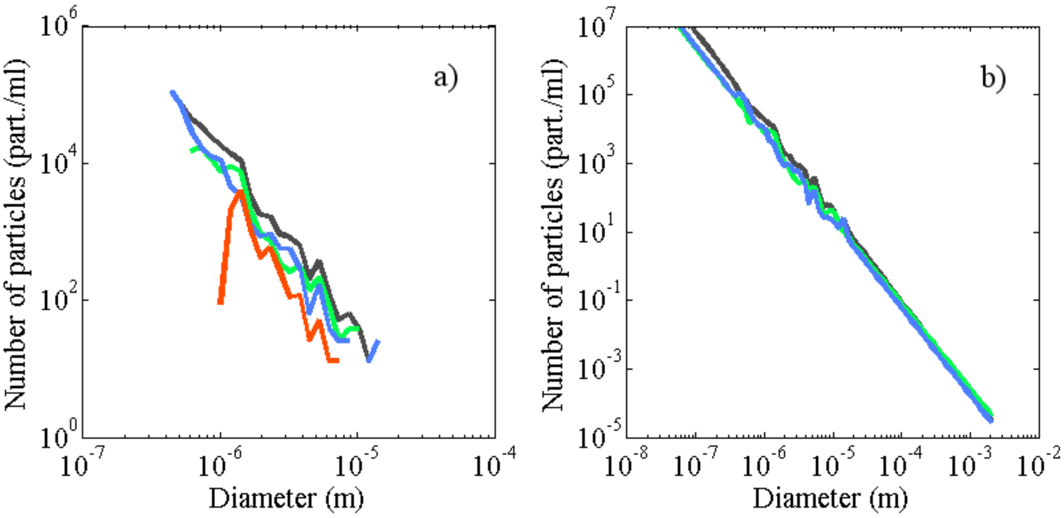


Figure 8.2 – (a) Total (black), organic (green), inorganic (blue) and fluorescent (red) PSDs for a typical sample of the UKCW dataset and (b) power law extension of the total, organic and inorganic PSDs. Note that the extended organic and inorganic PSDs intersect the extended total PSD; therefore the sum of the extended organic and inorganic PSDs is not exactly equal to the extended total TSM. To evaluate the error thus introduced, TSM values are modelled both from the total PSD and by summing model ISM and OSM values.

8.3 Particle mass modelling from apparent density of hydrated matter

Calculations for modelled values of TSM, OSM and ISM were made following the technique presented by Zhang *et al*. (2014). Building on the approach presented by Morel and Ahn (1990) and Babin *et al*. (2003) the technique estimates an apparent density for the particulate matter which is dependent on the real part of the refractive index of the particles and is designed to account for their water content. The equation takes the form

$$\rho = \rho_o V_o = \rho_o \frac{n-1}{n_o - 1}, \tag{8.1}$$

where n is the apparent real refractive index of the whole particle and n_o, ρ_o and V_o are respectively the real refractive index, density and fractional volume of the dry matter fraction of the particle. Values for the $\rho_o/(n_o - 1)$ ratio were defined following those employed by Zhang *et al.* (2014). For organic particles (defined as the fraction of the particle population with $n < 1.1$) the mean value of the ratio was set at $(8.56 \pm 1.1)\times10^6$ g/m^3; these are particles with a high water content, as high as ~80 ± 10% for algal cells (Aas, 1996). For mineral particles (defined as the fraction of the particle population with $n \geq 1.1$) the mean value of the ratio was instead set at $(15.52 \pm 1.84)\times10^6$ g/m^3. These particles have low water content; when the fractional volume of dry matter reaches unity (i.e. water content within the particles is zero) the apparent density of the particle becomes equal to the density of the dry mineral matter and can be calculated accordingly. Zhang *et al.* (2014) find $n = 1.16$ as the threshold above which $V_o = 1$, and give $\rho = [(6.42 \pm 0.85)n - (4.86 \pm 0.99)]\times10^6$ g/m^3 as the corresponding density based on a linear regression of literature values of density and refractive index for a number of mineral species. Overall, the final expression of eq. (8.1) used in practice was

$$\rho = \begin{array}{ll} 8.56 \times 10^6 (n-1) & n < 1.1 \\ 15.52 \times 10^6 (n-1) & 1.1 \leq n < 1.16 \\ 6.42 \times 10^6 n - 4.86 \times 10^6 & n \geq 1.16 \end{array} \tag{8.2}$$

with all density values given as g/m^3. Keeping with the assumption of particle sphericity used in the models employed thus far, a first attempt of total particle mass calculation for the total, organic and inorganic fractions of the particle population was made as a simple bin-by-bin summation of spherical masses, i.e.

$$m_{tot} = \sum_{D,n} \rho_n V_D N_{D,n} \tag{8.3}$$

where V_D is the volume of a sphere of diameter D, D and n are the diameter and rRI corresponding to each bin and $N_{D,n}$ is the number of particles within each bin. Given eq. (8.3) TSM, ISM and OSM can then obtained respectively by summing over the entire range of rRIs or by limiting the summation to real refractive index values above or below the n = 1.1 threshold. The model TSM, ISM and OSM values produced using this simple particle volume model however grossly overestimated the corresponding UKCW measurements, in certain cases by over two orders of magnitude (Fig. 8.3).

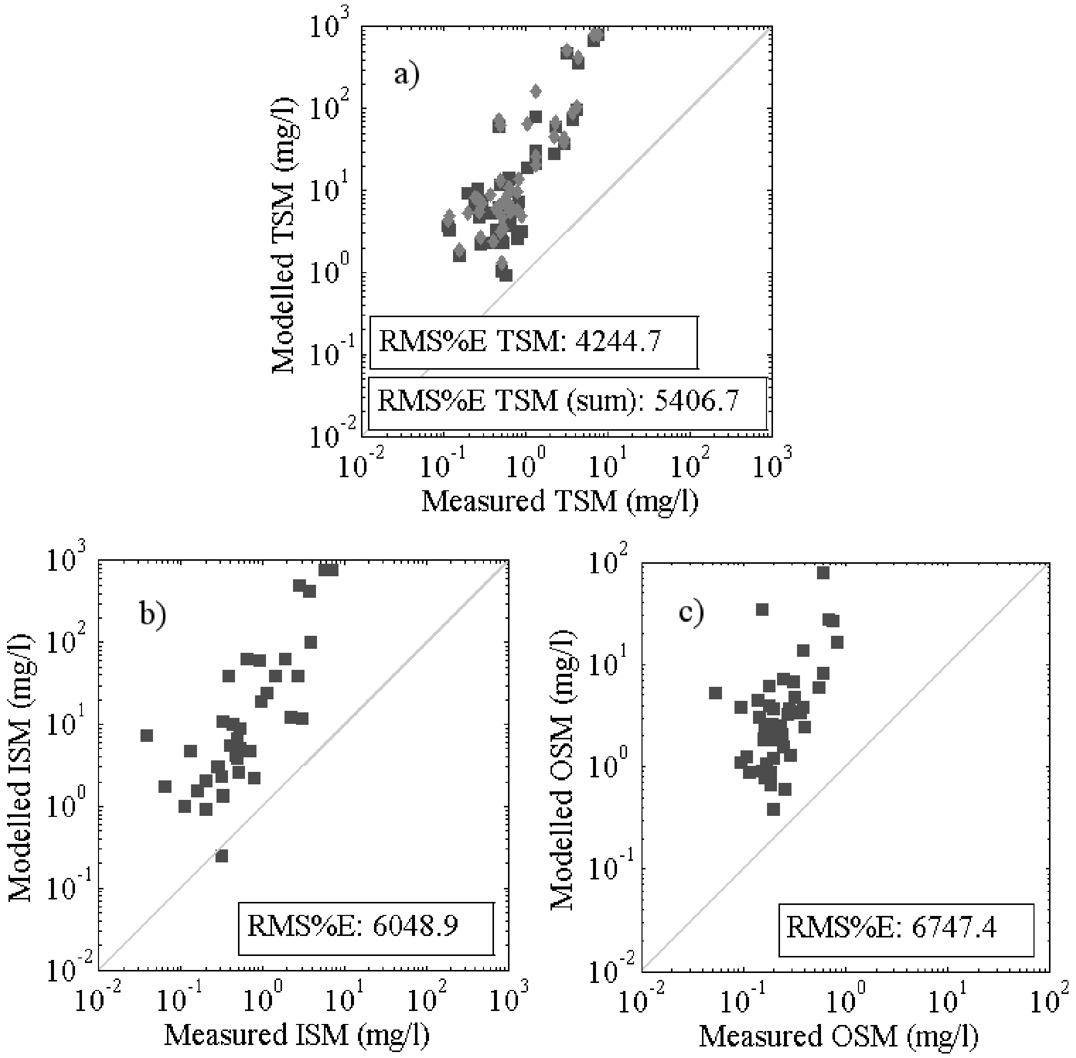


Figure 8.3 – Comparison of modelled vs. measured (a) TSM, (b) ISM and (c) OSM values for a simple spherical volume model. TSM values derived from the total PSD are represented as dark grey squares, while TSM values calculated as the sum of ISM and OSM are represented as light grey diamonds.

A second calculation attempt was therefore carried out using a slightly modified version of the particle volume and total mass model employed in Zhang *et al*. (2014). This model is designed to account for the fractal nature of some marine particles, which can exist as aggregates of smaller units rather than as individuals exclusively, and takes the form

$$m_{tot} = \sum\nolimits_{r,n} \frac{4\pi}{3}\left(\frac{r}{r_o}\right)^{F(r)} r_o^3 \rho_n N_{r,n} \tag{8.4}$$

where

$$F(r) = 3\left(\frac{r}{r_o}\right)^{\beta} \tag{8.5}$$

is the fractal dimension of the particles, r_o is the radius of the primary particles of which the aggregates are typically made up of and r is the effective radius of the particles.

The value of r_o and of exponent β are given as 0.5 µm and -0.0533 respectively (Khelifa and Hill, 2006, via Zhang *et al*., 2014), and for $r < r_o$ the value of F is fixed at 3. Crucially, the implementation of the model used here substitutes summation for the original integration to reflect the nature of the FC PSDs, extends the original range of 0.25-1000 µm to the 0.05-2000 µm one suggested by Davies *et al*. (2014) and used throughout this work, and implements the effective radius r directly as half the particle diameter D determined by the FC method (rather than as the geometric formulation $4/3 \times V/A$ used by Zhang *et al*., where A is the average projected area). This usage of Mie-derived parameters within a fractal model of mass highlights the dual nature of such modelling procedure: Mie theory is initially employed within the FC method to retrieve an optical size for the particles; the fractal model then reconciles this value with their physical size. Over the extended size range the value of F was found to vary between 3 (its maximum possible value) and 1.998. The TSM, ISM and OSM values produced using this fractal procedure were found to model the corresponding UKCW measurements much better than those produced using the simple spherical model, with RMS%E values 57.4%, 148.5% and 83.1% for TSM, ISM and OSM respectively (Fig. 8.4).

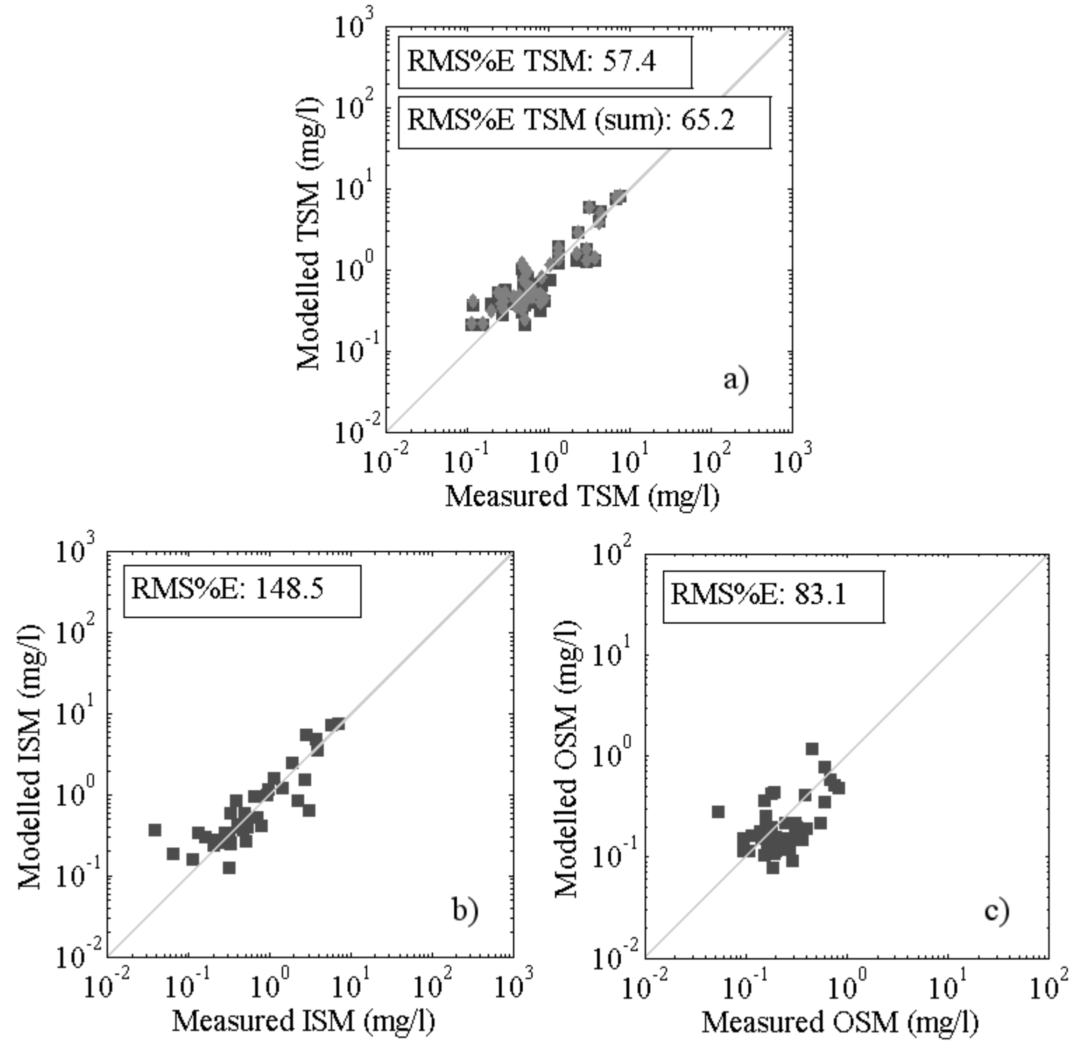


Figure 8.4 - Comparison of modelled vs. measured (a) TSM, (b) ISM and (c) OSM values for the fractal volume model. TSM values derived from the total PSD are represented as dark grey squares, while TSM values calculated as the sum of ISM and OSM are represented as light grey diamonds.

TSM values obtained as the sum of ISM and OSM were found to be close to those derived from the total PSD (RMS%E value 65.2%), indicating that the error introduced by the PSD extension is small. The first two samples of the UKCW dataset (corresponding to stations with large particle load in the Bristol Channel) produced clear outliers on all three accounts and were not included in the analysis. Median cumulative distributions of TSM, ISM and OSM were also produced, showing that in a majority of samples 90% of the contribution to all three parameters is done by particles between 0.2 and 200 µm (Fig. 8.5).

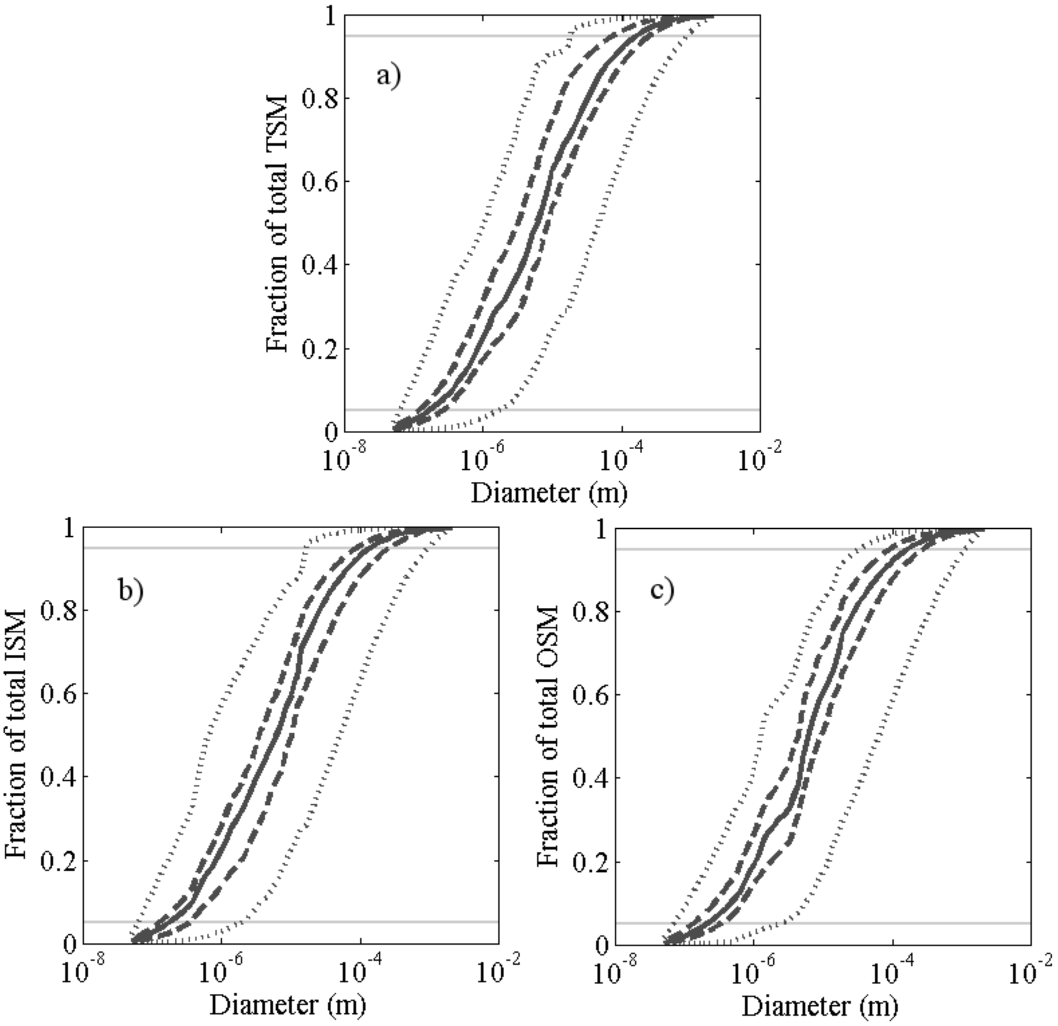


Figure 8.5 – Cumulative distributions of modelled (a) TSM, (b) ISM and (c) OSM values for the fractal volume model. The TSM curves refer to TSM values calculated from the total PSD. Solid, dashed and dotted lines represent median, upper/lower quartiles and maximum/minimum values respectively.

8.4 Particulate organic carbon and chlorophyll-A content modelling

8.4.1 Literature-derived models

POC and chlorophyll-A content within a cell are not linear functions of the cell volume (also defined as biovolume by some authors); total per cell content of either of the two quantities is instead derived using empirical relationships defined by volume scaling exponents, which can be then summed over organic and fluorescent PSDs to obtain POC and ChlA values respectively, i.e.

$$C = \sum_{r} y(r) N_r \, , \qquad (8.6)$$

where C represents either POC or ChlA, y is the corresponding total carbon or chlorophyll content per cell and N_r is the number of particles within each size bin. A number of these empirical conversion relationships can be found in the literature for the modelling described here: 4 sets of parameters for carbon (Verity *et al*., 1992; Montagnes *et al*., 1994; Menden-Deuer & Lessard, 2000) and 2 sets of parameters for chlorophyll-A (Montagnes *et al*., 1994, Álvarez *et al*., 2017) were employed. These are for carbon

$y = 0.433\, V^{0.863}$ (Verity *et al*, 1992)

$y = 0.109\, V^{0.991}$ (Montagnes *et al*., 1994)

$y = 0.288\, V^{0.811}$ (Menden-Deuer & Lessard, 2000 – Diatoms)

$y = 0.216\, V^{0.939}$ (Menden-Deuer & Lessard, 2000 – Non-diatom mixed protists)

and for chlorophyll-A

$y = 0.00429\, V^{0.917}$ (Montagnes *et al*., 1994)

$y = 0.0398\, V^{0.863}$. (Álvarez *et al*., 2017)

Here the particle volume V was defined using a simple spherical model. Of the four models used to calculate POC values, only the diatom model given by Menden-Deuer and Lessard produced results compatible with POC measurements (RMS%E: 92.9%, Fig. 8.6a). This possibly reflects the taxonomical composition of the algal populations encountered during the HE442 research cruise being mainly composed of diatom species typically associated with the spring bloom. The median cumulative distribution of POC for the diatom model shows an almost linear contribution from all size classes in a majority of samples, although results are shown to range widely from cases where the contribution is dominated by small particles to cases where, oppositely, the largest particles contributed the most (Fig. 8.6b). This is likely the result of the interaction between the model parameters and the slope of the PSDs, and may also indicate that the parameters of the model work well for a majority, but not all of the samples.

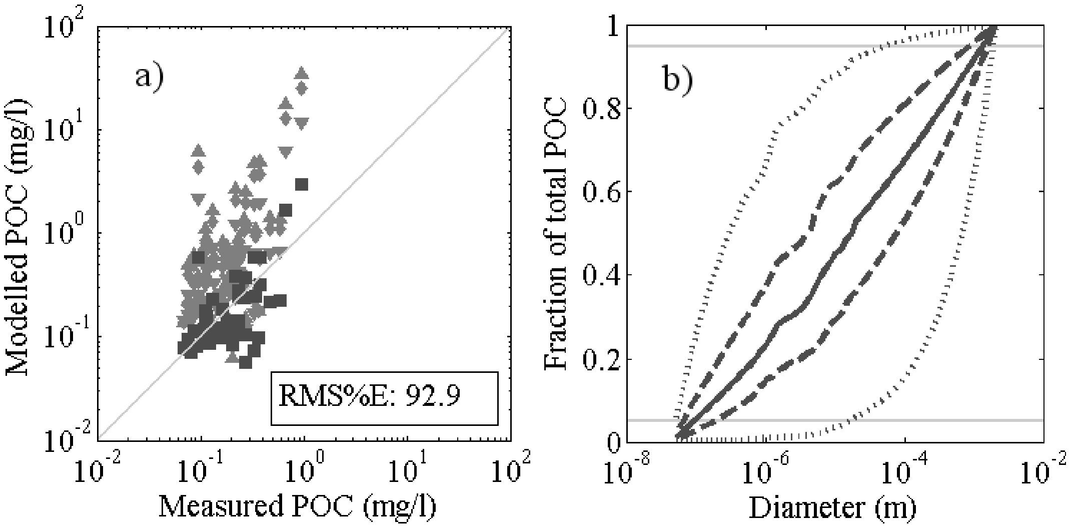


Figure 8.6 – (a) Comparison of modelled vs. measured POC. POC values calculated using the diatom model (Menden-Deuer & Lessard, 2000) are represented by dark grey squares; the RMS%E value refers to these. POC values calculated using the other three models are represented by light grey diamonds and triangles. (b) Cumulative distribution of modelled POC for the diatom model. Solid, dashed and dotted lines represent median, upper/lower quartiles and max./min. values respectively.

The two chlorophyll-A cell content models were applied to both HPLC-derived and fluorescence-estimated ChlA values (see Paragraph 5.3.5.3). In both cases, results were found to be mostly incompatible with ChlA measurements, but ChlA values obtained from initial fluorescence F_o (RMS%E of the combined dataset: 177.9%; Fig. 8.7a) performed notably better than the HPLC-derived ones (RMS%E of the combined dataset: 389.5%; Fig. 8.7b). This was found in agreement with results presented by Connor (2017) for the same UKCW dataset, which showed fluorescence-estimated ChlA values to perform slightly better than HPLC-derived ones when compared against algal absorption. The fluorescence-estimated values were accordingly selected as the best representation of total chlorophyll-A content; ChlA values in the following will solely refer to these.

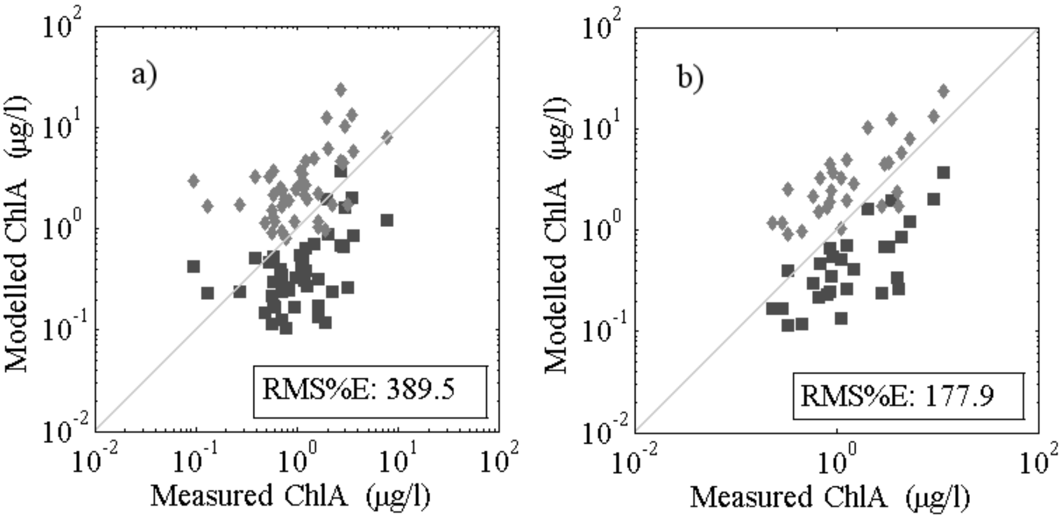


Figure 8.7 - Comparison of modelled vs. measured ChlA for (a) HPLC-derived and (b) fluorescence-estimated total chlorophyll-A content. ChlA values calculated using the Montagnes *et al*. (1994) model and the Álvarez *et al*. (2017) model are represented by dark grey squares and light grey diamonds respectively; the RMS%E value refers to a combined dataset of both.

8.4.2 POC and Chlorophyll-A content model optimization

A simple inversion of the procedure used to calculate POC and ChlA values allows for the empirical optimization of the parameter pairs used in the models. Maintaining the general form

$$y = kV^{h}, \tag{8.7}$$

arrays of values for the parameters k and h can be generated, combined and substituted in eq. (8.7), and the results compared and fitted against available measurements to identify the best parameter combinations, respectively (h_1, k_1) and (h_2, k_2) for POC and chlorophyll-A content. The optimization was initially applied to the whole UKCW dataset. Three parameter arrays were generated, one shared by exponents h_1 and h_2 plus one each for factors k_1 and k_2. The ranges were designed to encompass the parameter values of the models used thus far: specifically, 201 linearly spaced values for exponents h_1 and h_2 in a 0.6-1 range, 301 linearly spaced values for factor k_1 in a 0.05-0.65 range and 461 linearly spaced values for factor k_2 in a 0.004-0.05 range. Each (h_1, k_1) and (h_2, k_2) combination was then applied dataset-wide, compared against measured POC and ChlA and evaluated using the goodness-of-fit of a forced linear fit of the 1:1 line. The best parameter combinations were selected as those that concurrently maximized the adjusted R-square and minimized the RMSE values of the forced fit, i.e. as those that minimized a combined $R_{gof} = RMSE - R^2$ value. The best POC model for the UKCW dataset (Fig. 8.8a) was found as

$$y = 0.446\, V^{0.728}, \tag{8.8}$$

while the best chlorophyll-A content model (Fig. 8.8b) was found as

$$y = 0.036\, V^{0.762}. \tag{8.9}$$

Comparisons of the optimised model results with the measured POC and ChlA values of the UKCW dataset are shown in Fig. 8.9. RMS%E values for the comparisons were found to be 47.3% and 98.4% for POC and ChlA respectively.

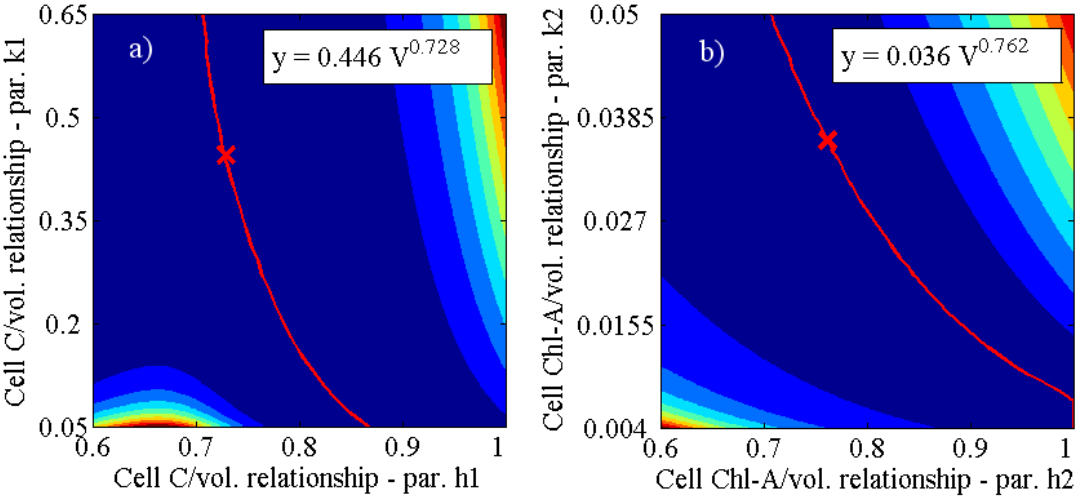


Figure 8.8 – Optimal model parameters identification for (a) POC and (b) chlorophyll-A volume scaling relationships, calculated over the whole UKCW dataset. The colours represent a combined $R_{gof} = RMSE - R^2$ value, rescaled to a range between 0 and 1 (blue to dark red). The red lines represent a continuum of minima separating the two poles of maximum R_{gof} values. The best combinations (red cross) are the ones corresponding to the respective global R_{gof} minima.

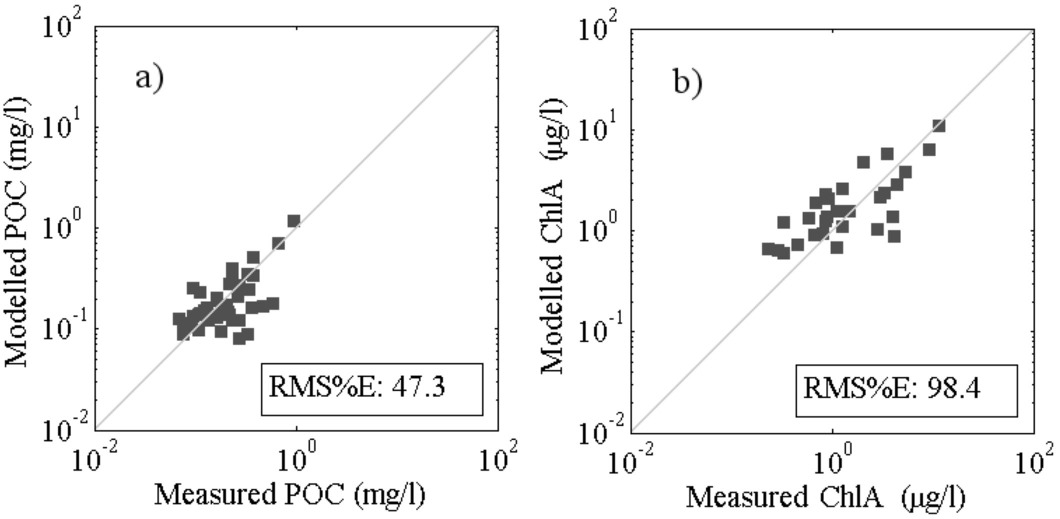


Figure 8.9 – Comparison of (a) POC and (b) ChlA values as determined by the optimised models of eqs. (8.8-8.9) vs. their respective measured values.

8.5 Discussion

The modelling of suspended matter parameters for the samples of the UKCW dataset produced an interesting result: PSDs and PRIDs generated using a sphere-based, Mie-derived methodology failed to produce TSM, ISM and OSM values which compared successfully with their corresponding measured parameters when paired with an equally simple spherical volume model (Fig. 8.3); however, the same PSDs and PRIDs produced comparable results when used directly as inputs for a fractal model of particle volume instead (Fig. 8.4). On one hand, this result offers further proof that the particle diameters found by the FC method are equivalent diameters rather than a measure of the physical dimension of the particles strictly (cf. Section 7.4); the former will tend to the latter the closer a particle gets to an ideal homogeneous sphere. On the other hand, this result may also shed light on the outcome seen in Chapter 6 for the FC-LISST comparison. The LISST reconstructs PSDs using measurements of the bulk VSF of a sample, and its measurement procedure is not mechanically intensive: flocs and aggregates are left relatively intact by the instrument as it descends through the water column. By contrast, the FC method reconstructs PSDs by composing together thousands of individual particle measurements, and flow cytometry itself is a mechanically intensive method which is quite likely to affect the integrity of flocs and aggregates as they are processed by the machine. However, the apparent paradox of having consistently similar PSD slopes and features between the two techniques may be solved if the FC method also calculates equivalent sphere diameters even in the presence of scale invariant fractal structures, and if the LISST "sees" not only scattering from whole flocs, but also from the primary units within them, as suggested by Graham et al. (2012).

POC and ChlA modelling had some success in reproducing measured quantities from literature-derived parameters (Fig. 8.6a-8.7) using a spherical model of particle volume, although the highly variable rate of success between models mirrors the lack of consensus on a single model to accurately represent size-physiology relationships in marine phytoplankton. Indeed, a single set of model parameters is likely to be simply insufficient to adequately represent the metabolic complexity of all algal organisms (Marañón *et al.*, 2007). Nevertheless, the further application of an empirical optimisation procedure to select POC and ChlA model parameters produced good match-ups between modelled and measured POC and ChlA values across the UKCW dataset (Fig. 8.9). Given the results obtained for TSM, ISM and OSM using a fractal model of particle volume, at a first glance the application of POC and ChlA modelling to simple spherical volumes appears incongruous. However, close inspection of eqs. (8.4-8.5) and (8.7) quickly reveals that the POC and ChlA empirical volume scaling relationships hide proportionalities which are very close to those of the TSM, ISM and OSM models. Specifically, from eqs. (8.4) and (8.7) and for a single particle:

$$\begin{aligned} m = \; & r_o^{3-F(r)} \rho \left(\frac{4\pi}{3}\right)^{1-\frac{F(r)}{3}} \left(\frac{4}{3}\pi r^3\right)^{\frac{F(r)}{3}} = \\ & = u(r) V^{\frac{F(r)}{3}} \propto \; y = kV^h. \end{aligned} \tag{8.10}$$

Of particular interest is then the comparison between the exponents h and *F(r)/3*. Over the 0.05-2000 µm size range considered here the value of *F(r)* was found to vary between 3 and ~2, producing values of *F(r)/3* between 1 and 2/3. These not only cover the range of values of exponent h in literature-derived and UKCW optimised models for both POC and ChlA, but also echo general results found in literature for the volume scaling coefficients of chlorophyll and organic carbon and their relationship with cellular metabolism. Metabolic rate under optimal growth conditions is seen to scale with volume following a 3/4 exponent for a large number of organisms in what is known as the 3/4 rule or Kleiber's rule (Finkel *et al.*, 2004), and in phytoplankton this relationship is directly tied to the photosynthetic rate, and ultimately to the intracellular chlorophyll-A concentration; in general, phytoplankton cells regulate their pigment concentration in response to environmental irradiance changes (Finkel *et al.*, 2004). Values for the volume scaling coefficient are then variably predicted to range between 3/4 and 2/3 for optimal growth and light limited conditions respectively (Finkel *et al.*, 2004), or to reach ~1 when nutrients are abundant (Marañón *et al.*, 2007). Álvarez *et al.* (2017) report values between 3/4 and 1, and Mei *et al.* (2009) modelled cellular growth rate scaling exponents using chlorophyll-A content scaling exponents ranging between 2/3 and 1. Similarly, empirical estimates of the scaling between cell volume and particulate carbon also vary. Some authors find particulate carbon content to decrease proportionally with increasing cell size i.e. $h < 1$ (Verity *et al.*, 1992; Menden-Deuer & Lessard, 2000), while other find it to be isometric to cell size i.e. $h \sim 1$ (Montagnes *et al.*, 1994). Ultimately, the results presented in this chapter seem to suggest that, while cellular metabolic rate can be logically expected to play an important part in defining the chlorophyll and organic carbon content of organic particles, structural characteristics of the organic particles as described by fractal models can also offer a complementary interpretation for the proportionalities observed in nature and described in existing literature.

8.6 Conclusions

In this final chapter, the results obtained by the FC method for the UKCW dataset were combined with models of particle density and of organic carbon and chlorophyll-A content to investigate the biogeochemical properties of the particle populations. The success of the resulting TSM, ISM, OSM and (after empirical optimisation) POC and ChlA estimations lends further credibility to the PSD and PRID determination capabilities of the FC method, and further supports the usefulness of flow cytometry and of the FC method as a tool to complement other established techniques. The fact that a fractal model of particle structure was key to ensure the quality of the TSM, ISM and OSM match-ups suggests that the FC method observes particles as equivalent spheres, and is therefore to some extent resilient both to particles which violate Mie-compatible aspect ratios (cf. Chapter 7) and to the likely break-up of flocs and aggregates which derives from the flow cytometric measurement technique. Furthermore, this characteristic makes FC method results more readily comparable with those of other more common marine optics instruments, which for the most part observe bulk seawater IOPs and PSDs. The models used to calculate POC and ChlA, despite being at a first glance applied to simple spherical particles, were in fact revealed to hide proportionalities analogous to those caused by fractal structures in the TSM, ISM and OSM models, shining interesting new light on the volume scaling coefficients described in marine microbiology literature. Finally, the diameter discrimination offered by FC further allows for a better understanding of the contribution of different size classes to the bulk biogeochemical properties, as it did for the IOPs of natural particle populations in Chapter 7.

9. Summary and conclusions

Pioneering work carried out almost three decades ago demonstrated the possibility of associating particle scattering within a flow cytometer to the size and refractive index of each analysed particle (Ackleson & Spinrad, 1988), effectively providing a practical bridge between scattering and the physical properties of a particle. The effort remained isolated. The technique saw little use beyond a brief revival in 2003 (Green *et al*., 2003a; Green *et al*., 2003b), and to this day even within the context of marine sciences flow cytometry is chiefly used in its main capacity as an instrument for fluorescence detection and characterisation.

The aim of this work has been to recuperate this technique and to apply it anew, in an effort to determine whether the data quality provided by modern flow cytometry was sufficient to produce valid and useful PSDs and PRIDs from a particle-by-particle analysis of natural seawater particle population samples. While bulk measurements, with their relative ease of execution and high temporal resolution, remain a fundamental part of the marine optics toolkit, being able to associate single-particle measurements of particle size and refractive index to these techniques would unlock a wealth of knowledge which is still largely untapped.

The first step was to justify the need to complement bulk measurements with a flow cytometry method geared towards individual particles. The practice of approximating natural particle populations to a power law size distribution of spherical, homogeneous particles with a single, spectrally independent complex refractive index is commonplace in optical oceanography, and offered the opportunity for a sensitivity analysis to reveal the potential effect on IOPs of variations from assumed conditions. These are the kind of PSD and PRID variations which single-particle measurements of particle size and refractive index would be ideal to detect. The effect of the addition of spectral dependence to the complex refractive index was found to be small; however, marked differences were found in the results between organic- and inorganic-like particles, suggesting that while this case was not directly examined within the sensitivity study, a mixed particle population as opposed to a single complex RI would have a non-negligible effect on VSF and other IOPs. Local additions of log-normal populations to the power law were further found to induce sometimes large λ-dependence in the VSF and therefore in derived quantities like the backscattering ratio. Overall, the results demonstrated the potential usefulness of a reliable method for routine determination of the PSDs and PRIDs of marine particle populations.

The sensitivity study also offered the chance to assess the resilience of the Fournier-Forand phase function (Fournier & Forand, 1994; Mobley *et al*., 2002) and of the χ factor method for backscattering coefficient determination (Oishi, 1990; Maffione & Dana, 1997; Boss & Pegau, 2001). Both ultimately rely on the common approximations which were tested with the sensitivity study. The Fournier-Forand phase function was found unable to reproduce all the features generated on the scattering phase function by the PSD perturbations, especially in the forward and narrow backward direction; the angular neighbourhood around 120 degrees was however found as one of particular stability for the scattering angular structure. Conforming to this, the χ factor method for backscattering coefficient determination, which is valid for angles between 117 ± 3 degrees (Boss & Pegau, 2001), was found to be a robust one. Results indicated that the refractive index composition of the particle population has important consequences on the choice of the individual angle where the χ factor is measured, again underlying the usefulness of a method capable of directly assessing this composition. Results also confirmed that picking an angle at ~120 degrees may improve the resilience of the χ factor method.

The second step was to develop the FC method proper, employing Mie theory as the optical model of choice for scattering calculations. First of all, the flow cytometer was characterised in its capability of precisely determining particle concentrations within a sample. The characterisation procedure was carried out using polymer bead samples with known particle concentrations, and produced empirical corrections which ensured counting precision. The geometry of the detectors was then defined using weighting functions to account for forward and side scattering sensor aperture shape. Correct angular collection ranges were determined firstly in a rough approximation and then iteratively until agreement was reached with measured scattering from reference polymer beads of known diameter and refractive index.

These reference samples were further used to scale and centre the scattering model, establishing a relationship between particle physical properties and scattering as measured by the flow cytometer. Measurements taken at multiple PMT settings were combined to cover the entire FC detection range, spanning from ~0.5 microns to a few hundred micrometres, and resulting particle diameter and refractive index data was then double-binned within a matrix to produce final PSDs and PRIDs.

Once developed, the FC method was initially applied to samples containing particles of known size (polymer reference beads) and known refractive index (oil droplet suspensions), and found to be capable of identifying diameter and relative rRI within a range of values, i.e. ~0.5 to ~10 µm for diameter and up to 1.15 for rRIs. Particles larger than 5 µm presented increasingly anomalous side scattering profiles, possibly due to the breakdown of Mie theory assumptions i.e. particles becoming larger than the laser beam width (5 µm). This hampered correct diameter recognition in very large particles. Relative rRI values higher than 1.15 couldn’t be identified reliably, but still provided a generic high-rRI signifier as opposed to more specific low rRI values.

The method was then applied to the phytoplankton culture samples of the AC dataset. Size determination in phytoplankton culture samples using the FC method was found to be strongly influenced by shape and structure of the phytoplankton cells, as particularly evident in the case of species with extreme cell aspect ratio such as *P. seriata*. Rounded, smaller cells offered the best results as it is to be expected in the context of Mie theory. Real refractive index detection was found to be less affected than size detection by the aspect ratio of the particles, even though a general tendency for the FC method to underestimate rRI in very large particles may have biased the rRI results for phytoplankton towards organic-like values.

Finally, the FC method was applied to the natural seawater samples of the UKCW dataset, producing PSDs which compared well with the features and slope of independent LISST-100x PSD measurements. However, all PSDs were observed to rapidly lose statistical significance above ~20 µm due to few particle counts and particle concentrations reaching a floor corresponding to bins containing a single particle. This is likely to have been an effect of the measurement protocol, which selected for smaller particles. Finding statistically significant numbers of large particles will necessitate a revision of the measurement protocol or at the very least longer acquisition times. Furthermore, a lack of consensus on particle concentration was observed between the LISSTs and the FC, but also between one LISST instrument and the other. The calibration of the FC against particle concentration standards added weight to the FC data, but a definitive judgement on the reliability of the concentration measurements couldn't be made. FC PRIDs of the UKCW dataset were found to be largely homogeneous with a broad maximum between 1.05-1.15 and peak values at 1.07-1.09, consistent with a mix of organic and inorganic components. The range of values accounts for most of the common components of marine suspended particulate matter. A tail of particles with rRI > 1.15 was observed in all UKCW samples, with a particular concentration spike visible at ~1.3 found to be made up of particles outside the scattering model grid. These were found to account for less than 1% of the total particle population in all samples of the UKCW dataset, and interpreted as small, faceted, high-rRI particles incompatible with Mie theory assumptions. Due to their relatively small number, the influence of particles outside the scattering model grid was deemed negligible.

The results of the PSD and PRID retrieval demonstrated the potential of a FC method capable of simultaneous particle size and refractive index identification, both in controlled samples with monotypic spherical particles (polymer beads, oil suspensions) and in natural water samples containing mixed populations. Furthermore, the inherent ability of the FC method to partition the PSD by particle type offered the chance to investigate the links of PSDs and PRIDs to optical properties and the individual contributions of particle subpopulations to the bulk IOPs, as the size range covered by the method (0.5-10 µm) is found to be particularly relevant for optical studies (Davies *et al.*, 2014). The PSDs determined by the FC method for the AC and UKCW datasets were thus combined with their respective PRIDs and fed into Mie-based forward optical modelling to reconstruct total IOPs, which were then compared for optical closure against independent measurements of bulk absorption, scattering and backscattering. To successfully carry out IOP calculations, the FC PSDs were extrapolated beyond the FC detection range using a power law distribution model, in an effort to account for the whole optically relevant range from tens of nanometres to a few millimetres (Davies *et al*., 2014).

The results of closure analysis between modelled and measured IOPs in the AC dataset were found to vary between algal species. Large, rounded species *Heterocapsa sp.* and *A. minutum* were found to largely overestimate corresponding measured IOPs, while small-rounded and chain-forming species achieved good agreement across all IOPs considered. Disagreement for *Heterocapsa sp.* and *A. minutum* IOPs was diminished greatly when the PSD extrapolations were excluded from IOP calculations, suggesting that the power law extensions introduced more particles than were actually present in the samples; at the same time, other species achieved much better agreement using similar PSD extrapolations, and removing the contribution of undetected particles from the IOP calculations was found to predictably cause general underestimation of the measured IOPs. Overall, the results of Mie forward optical modelling for the AC dataset highlighted the risks of using power law extrapolations on samples which do not represent natural particle populations. The positive results of the closure analysis for chain-forming species were surprising, particularly in the case of *P. seriata*, and were interpreted to suggest that by nature the FC method produces diameter and rRI values corresponding to spheres which are optically equivalent to the measured particles.

The best results for the closure analysis between measured and modelled IOPs for the UKCW dataset were achieved for scattering and backscattering coefficients, although broad agreement was reached for all IOPs considered. Given the relatively narrow size range covered by the FC method, the need to extrapolate PSDs and then approximate rRIs in the PSD extensions, and the lack of directly determined iRI values, the degree to which forward modelled IOPs matched measured values is surprisingly good. This provides useful validation of the PSDs and PRIDs generated by the FC method, and in doing so also suggests that Mie theory is sufficient to provide useful estimates of bulk IOPs for randomly orientated, mixed populations of naturally occurring marine particles. Finally, PSDs and PRIDs as determined by the FC method did provide useful insight into the individual contributions of different particle subpopulations to the bulk IOPs of seawater samples; as a last test of the validity of FC method results and of the limits of the modelling enabled by them, a similar approach was taken to investigate to what degree the same PSDs and PRIDs could be used to successfully inform models of particle density and of organic carbon and chlorophyll content, and to explore the biogeochemistry of the particle populations of the UKCW dataset.

Interestingly, combining PSDs and PRIDs obtained by the FC method for the UKCW dataset with literature-derived models of particle density and simple spherical particle volumes produced TSM, ISM and OSM values which failed to compare successfully with corresponding measured quantities. However, the same PSDs and PRIDs and particle density models produced suspended matter values comparable with measured TSM, ISM and OSM if applied to a fractal model of particle volume instead. This result once more lends support to the validity of the PSDs and PRIDs determined by the FC method, and further demonstrates the usefulness of flow cytometry and of the FC method as a tool to complement other established techniques. As with the results obtained by IOP forward modelling in chain-forming species, it also offers additional proof that the particle diameters found by the FC method are indeed equivalent diameters rather than a measure of the physical dimension of the particles (cf. Tab. 6.3, Fig. 6.7).

POC and ChlA modelling had some success in reproducing measured quantities from literature-derived parameters, although the variable rate of success suggests that a single set of model parameters is likely to be insufficient to represent the metabolic complexity of all algal populations. Nevertheless, the further application of an empirical inversion procedure to select optimised POC and ChlA model parameters produced good match-ups between modelled and measured POC and ChlA values across the UKCW dataset. While the POC and ChlA models were applied to simple spherical volumes, their power law nature ultimately causes proportionalities which are in fact analogous to those seen in the TSM, ISM and OSM models for fractal particles. In other words, while they rigorously are power law equations applied to spheres, they are also ultimately revealed to be mathematically equivalent to polynomial equations applied to fractal structures, as was the case for TSM, ISM and OSM. These results shine new light on the volume scaling coefficients described in marine microbiology literature: while cellular metabolic rate can be logically expected to play an important part in defining the chlorophyll and organic carbon content of organic particles, structural characteristics of the organic particles as described by fractal models can also offer a complementary interpretation for the proportionalities observed in nature and described in literature.

In summary, this work has represented an effort to recover and re-evaluate a methodology that, at least to the best knowledge of the author, is unique in its capability of concurrently retrieving both size and refractive index of suspended marine particles on an individual basis. The information that the FC method provides in terms of PSD and PRID determination, coupled with the inherent fluorescence detection capabilities of flow cytometry, represents a large and yet to be used wealth of data which, once coupled with adequate models, straddles not only physical, but also optical and biogeochemical particulate properties. The FC method is not proposed as a technique capable of replacing any of the already established techniques for PSD, IOP or biogeochemistry measurements. Rather, it is proposed as a tool to assist and complement those other techniques, filling many blanks and unknowns which are usually glossed over or accounted for using wide assumptions. The capability of the FC method to identify individual fractions within the particle populations and to then estimate each fractional contribution to total particulate IOPs and biogeochemical properties in particular is seen as perhaps the most important new contribution that this technique can bring to marine sciences. In closing, while there is much scope for improvement on many aspects and especially on the front of the statistical significance of the results at large particle sizes, the reliability and usefulness of the FC method even in its present form have been vindicated multiple times throughout this work against both known standards and independent measurements obtained using other instruments. Although further development will be certainly needed, the methodology has the definite potential to become a relevant component of the toolkit of marine optics and of marine sciences in general.

10. Future work

The work carried out on the FC method has revealed a few outstanding points where improvements are necessary. Furthermore, a number of avenues for future research are apparent.

- The most immediate issue in the present iteration of the FC method is the lack of statistical significance for large particles beyond ~20 µm. This is in part an inherent problem of flow cytometry, due to the small sample volumes involved; however, there is scope for improvement in the measurement protocol used. The simplest option available is to extend the sampling time, although often the amount of time available for a measurement is limited by external circumstances, especially when working *in situ*. A second option is to reduce the number of PMT settings used to cover the particle size range, possibly moving from the current 4 settings to just 3 settings. This could naturally extend the sampling time for each PMT setting without impacting the current duration of FC measurements. Overall, improving the statistical significance of large particle counts remains a matter of compromise, and more data acquisition will be needed to evaluate the effectiveness of new measurement protocols.

- A second important issue revealed in this study was the underestimation of particle sizes determined by the FC method for large particles above ~5-10 µm; this was especially glaring in the case of large spherical polymer beads for which the method should be particularly suited. At least part of the problem seems to reside in a violation of the incident plane wave assumption of Mie theory, in the sense that the particles become increasingly larger than the width of the laser beam (5 µm). This particular issue looks difficult to surmount using the current laser beam shaping. However, there is anecdotal evidence that the issue may also be partly due to PMT effects caused by the large influx of scattered photons in the detectors, even before saturation actually occurs. Installing more resilient sensors may be a good starting point in investigating the extension of the functional size range of the FC method.

- There remains some uncertainty on the reliability of the particle concentrations retrieved by the FC method, as it was highlighted by the comparison between FC PSDs and LISST PSDs. Although the calibration of the FC method against standards of known concentration and the very disagreement between PSDs produced by different LISST instruments adds weight to the idea that the FC method is indeed capable of precision concentration measurements, a new set of measurements specifically geared towards resolving this outstanding issue is certainly desirable. A thorough measurement campaign carried out using both the FC method and other independent PSD meters on a set of various standards of known concentration seems the logical course of action towards this goal.

- Commercial flow cytometers in general are not geared towards the kind of scattering-based application which has been described in this work. Conforming to their common usage as biomedical instruments, most flow cytometers are specifically designed around fluorescence applications, and in marine sciences as well they are mostly used in phytoplankton studies. Furthermore, the beam shaping within the CytoSense is specifically designed to improve the resolution of particle profiling and of signal length determination, but is detrimental to scattering-based assessments of size and refractive index in large particles. Specifically designing a new flow cytometric instrument with scattering applications in mind may be an interesting endeavour which could result in better quality of results for the FC method. Additionally, the integration of holographic imaging techniques and flow cytometry is proposed as an extremely promising new frontier of sensor technology which could be advanced in parallel with the scattering-based FC method. Ideally, the two techniques may be integrated within a single instrument.

- A second aspect of desirable technological improvement would be to automate the FC method, at least to some degree. This would be done with the goal of direct *in situ* deployment, either for water column profiling from a vessel or for continued water monitoring from buoys. A precedent exists in

Imaging FlowCytobot (Olson & Sosik, 2007; Sosik & Olson, 2007), an imaging flow cytometer developed for long-term phytoplankton automated monitoring and recognition. Investigating the potential for the FC method to be employed in similar long-term fashion would require concerted advances in hardware, measurement protocol and data processing codes, and a rigorous testing of the consistency of results over long time periods. Nevertheless, automating the FC method procedure would greatly widen the scope of its application.

- The connection with fractal models of particle volume which was identified while investigating the biogeochemistry of the particle population appears to be a promising new avenue of research which could potentially shine new light on the optical behaviour of flocs and aggregates. The single-particle analysis approach inherent to flow cytometry, coupled with the availability of associated imaging, could offer information that bulk methods or pure imaging methods cannot deliver on their own. In this context, the integration of holographic imaging techniques and flow cytometry is seen as even more desirable, possibly in tandem with multiscale models of particle size e.g. circle-packing (Graham et al., 2012) and generalised Mie theory calculations suited for multiple packed spheres (Gouesbet & Grehan, 1999).

- Finally, the determination of particle physical parameters as equivalent spheres may potentially be exploited to determine the aspect ratio of the particles, providing more information than that available from simple sizing along one dimension and possibly informing more complex optical modelling or even just algal taxonomy. By evaluating the degree of underestimation between scattering as measured by the flow cytometer and theoretical scattering from a sphere with diameter equivalent to the physical length of the particle (as determined by signal length), a relationship may be established between underestimation and particle aspect ratio, the idea being that scattering underestimation will progressively tend towards zero as the aspect ratio itself becomes spherical.

References

Aas, E. (1996), “Refractive index of phytoplankton derived from its metabolite composition”, J. Plankton Res., 18, pp. 2223-2249.

Ackleson, S. G. and Spinrad, R. W. (1988), “Size and refractive index of individual marine particulates: a flow cytometric approach”, Applied Optics, Vol. 27, No. 7.

Adan, A., G. Alizada, Y. Kiraz, Y. Baran, and A. Nalbant (2016), “Flow cytometry: basic principles and applications”, Crit. Rev. Biotechnol., 37(2), pp. 163-176, DOI: 10.3109/07388551.2015.1128876.

Aden, A. L. and M. Kerker (1951), “Scattering of electromagnetic waves from two concentric spheres”, J. Appl. Phys., 22(10), pp. 1242-1246, DOI: 10.1063/1.1699834

Agrawal, Y. C. and H. C. Pottsmith (1994), “Laser Diffraction Particle Sizing in STRESS”, Continental Shelf Research, 14, pp. 1101–1121, doi:10.1016/0278-4343(94)90030-2

Agrawal, Y. C. and H. C. Pottsmith (2000), “Instruments for Particle Size and Settling Velocity Observations in Sediment Transport”, Marine Geology, 168, pp. 89–114, doi:10.1016/S0025-3227(00)00044-X

Ahn, Y.-H., A. Bricaud, and A. Morel (1992), “Light backscattering efficiency and related properties of some phytoplankters”, Deep-Sea Res., 39, pp. 1835–1855.

Álvarez, E., E. Nogueira, and Á. López-Urrutia (2017), “In vivo single-cell fluorescence and size scaling of phytoplankton chlorophyll content”, Appl. Environ. Microbiol., 83, e03317-16. https://doi.org/10.1128/AEM.03317-16.

Asano, S. and G. Yamamoto (1975), “Light scattering by a spheroidal particle”, Applied optics, 14(1), pp. 29-49, DOI: 10.1364/AO.14.000029

Ashley, L. E. and C.M. Cobb (1958), “Single particle scattering functions for latex spheres in water”, J. Opt. Soc. Am., 48, pp. 261-268.

Babin, M., A. Morel, V. Fournier-Sicre, F. Fell, and D. Stramski (2003), "Light scattering properties of marine particles in coastal and open ocean waters as related to the particle mass concentration", Limnol. Oceanogr., 48(2), pp. 843-859.

Bader, H. (1970), "The hyperbolic distribution of particles sizes", J. Geophys. Res. 75, pp. 2822-2830.

Barton, J. P. (2001), "Internal, near-surface, and scattered electromagnetic fields for a layered spheroid with arbitrary illumination", Applied Optics, 40(21), 3598.

Basiji, D. A., W. E. Ortyn, L. Liang, V. Venkatachalam, and P. Morrissey (2007), "Cellular Image Analysis and Imaging by Flow Cytometry", Clinics in Laboratory Medicine, 27(3), 653–viii. http://doi.org/10.1016/j.cll.2007.05.008.

Bauer, D. and A. Morel (1967), "Etude aux petits angles de l'indicatrice de diffusion de la lumière par les eaux de mer", Annales de Géophysique, 23, pp. 109-123.

Bohren, C. F., D. R. Huffman (1998), "Absorption and scattering of light by small particles", Wiley Professional Paperback ed., John Wiley & Sons, Inc., New York.

Borghese, F., O. I. Sindoni, P. Denti, and R. Saija (1992), "Optical properties of spheres containing a spherical eccentric inclusion", J. Opt. Soc. Am. A 9, pp. 1327-1335.

Boss, E. and W. S. Pegau (2001), "Relationship of light scattering at an angle in the backward direction to the backscattering coefficient", Appl. Opt., 40, pp. 5503-5507.

Bricaud, A. and A. Morel (1986), "Light attenuation and scattering by phytoplanktonic cells: a theoretical modelling", Appl. Opt. 25, pp. 571-580.

Bricaud, A., A.-L. Bedhomme, and A. Morel (1988), "Optical properties of diverse phytoplanktonic species: experimental results and theoretical interpretation", J. Plankton Res., 10, pp. 851-873.

Bricaud, A., M. Babin, A. Morel, and H. Claustre (1995), “Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization”, J. Geophys. Res., 100, pp. 13321–13332.

Brosnahan, M. L., S. Farzan, B. A. Keafer, H. M. Sosik, R. J. Olson, D. M. Anderson (2013), “Complexities of bloom dynamics in the toxic dinoflagellate Alexandrium fundyense revealed through DNA measurements by imaging flow cytometry coupled with species-specific rRNA probes”, Deep-SeaRes II, 103, pp. 185-198, doi: 10.1016/j.dsr2.2013.05.034

Bryant, F. D., B.A. Seiber, and P. Latimer (1969), “Absolute optical cross sections of cells and chloroplasts”, Arch. Biochem. Biophys., 135, pp. 97-108.

Chin, J. H., C. M. Sliepcevich, and M. Tribus (1955), “Determination of particle size distributions in polydispersed systems by means of measurements of angular variation of intensity of forward-scattered light at very small angles”, J. of Phys. Chem, 59, pp. 841-845.

Connor, D. (2017), “Investigating the Use of Fast Repetition Rate Fluorometry in Understanding Algal Physiology in Optically Complex Oceans”, Doctoral thesis, University of Strathclyde, Glasgow.

Cunningham, A. and G. A. Buonaccorsi. (1992), “Narrow-angle forward light scattering from individual algal cells: implications for size and shape discrimination in flow cytometry”, J. Plankton Res., 14, pp. 223-234.

Daugbjerg, N., G. Hansen, J. Larsen, and Ø. Moestrup (2000), “Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates”, Phycologia, 39, pp. 302-317.

Dave, J. V. (1968), “Subroutines for Computing the Parameters of the Electromagnetic Radiation Scattered by a Sphere”, IBM Palo Alto Scientific Center, Publication 320-3237.

Davies, E. J., D. McKee, D. Bowers, G. W. Graham, and W. A. M. Nimmo-Smith (2014), “Optically significant particle sizes in seawater”, Applied Optics, Vol. 53, No. 6.

Draine, B. T. and P. J. Flatau (1994), “Discrete-Dipole Approximation For Scattering Calculations”, J. Opt. Soc. Am. A, 11, pp. 1491-1499.

Dubelaar, G. B. J. and R. R. Jonker (2000), “Flow cytometry as a tool for the study of phytoplankton”, Sci. Mar. 64, pp. 135–156.

Duforêt-Gaurier, L., W. Moutier, N. Guiselin, M. Thyssen, G. Dubelaar, X. Meriaux, L. Courcot, D. Dessailly, and H. Loisel (2015), “Determination of backscattering cross-section of individual particles from cytometric measurements: a new methodology”, Opt. Express 23, pp. 31510-31533.

Elliott, G. S. (2009), “Moving pictures: imaging flow cytometry for drug development”, Comb. Chem. High Throughput Screening, 12, pp. 849–859.

Fenn, R. W. and H. Oser (1965), “Scattering properties of concentric soot-water spheres for visible and infrared light”, Applied Optics, 4(11), 1504, DOI: 10.1364/AO.4.001504

Fikioris, J. G. and N. K. Uzunoglu (1979), “Scattering from an eccentrically stratified dielectric sphere”, J. Opt. Soc. Am. 69, pp. 1359-1366.

Finkel, Z. V., A. J. Irwin, and O. Schofield (2004), “Resource limitation alters the 3/4 size scaling of metabolic rates in phytoplankton”, Mar. Ecol. Prog. Ser., 273, pp. 269–279. https://doi.org/10.3354/ meps273269

Fournier, G. R. and J. L. Forand (1994), “Analytic phase function for ocean water”, in Ocean Optics XII, J. S. Jaffe, ed., Proc. SPIE 2258, pp. 194–201.

Fuller, K. A. (1993), “Scattering and absorption by inhomogeneous spheres and sphere aggregates”, Proc. SPIE 1862, Laser Applications in Combustion and Combustion Diagnostics, 249, DOI:10.1117/12.145702

Fuller, K. A. (1993), “Scattering of light by coated spheres”, Optics letters, 18(4), 257.

Goda, K., A. Ayazi, D. R. Gossett, J. Sadasivam, C. K. Lonappan, E. Sollier, A. M. Fard, S. C. Hur, J. Adam, C. Murray, C. Wang, N. Brackbill, D. Di Carlo, and B. Jalali (2012), “High-throughput single-microparticle imaging flow analyser”, Proc. Natl. Acad. Sci. USA, 2012, 109, pp. 11630–11635, doi: 10.1073/pnas.1204718109.

Gordon, H. R. and.O. B. Brown (1972), “A theoretical model of light scattering by Sargasso Sea participates”, Limnol. Oceanogr., 17, pp. 826-832.

Gorthi, S. S., D. Schaak and E. Schonbrun (2013), “Fluorescence imaging of flowing cells using a temporally coded excitation”, Opt. Express, 21, pp. 5164–5170.

Gouesbet, G. and G. Grehan (1999). “Generalized Lorenz-Mie theory for assemblies of spheres and aggregates”. Journal of Optics A: Pure and Applied Optics, 1(6), pp. 706-712.

Graham, G. W., E. J. Davies, W. A. M. Nimmo-Smith, D. G. Bowers, and K. M. Braithwaite (2012), “Interpreting LISST-100X measurements of particles with complex shape using digital in-line holography”, J. Geophys. Res., 117, C05034, doi:10.1029/2011JC007613

Green, R. E., H. M. Sosik and R. J. Olson (2003), “Contributions of phytoplankton and other particles to inherent optical properties in New England continental shelf waters”, Limnol. Oceanogr., 48(6), pp. 2377–2391.

Green, R. E., H. M. Sosik, R. J. Olson and M. D. DuRand (2003), “Flow cytometric determination of size and complex refractive index for marine particles: comparison with independent and bulk estimates”, Applied Optics, Vol. 42, No. 3.

Jerlov, N. G. (1953), “Particle distribution in the ocean”, Rep. Swedish Deep-sea Exped., 3, pp. 73-97.

Jerlov, N. G. (1961), “Optical measurements in the eastern North Atlantic”, Medd. Oceanogr. Inst. Göteborg, 30, pp. 1–40.

Jerlov, N. G. (1963), “Optical Oceanography”, Oceanogr. Mar. Biol. Annu. Rev., 1, pp. 89-114.

John, M. D., B. A. Whitton, and A. J. Brook (2002), “The freshwater algal flora of the British Isles: an identification guide to freshwater and terrestrial algae”, Cambridge University Press.

Jonasz, M. (1983), “Particle size distribution in the Baltic”, Tellus Ser. B, 35, pp. 346–358.

Jonasz, M. and G. R. Fournier (1996), “Approximation of the size distribution of marine particles by a sum of log-normal functions”, Limnol. Oceanogr., 41(4), 1996, pp. 744-754.

Jonasz, M., and G. R. Fournier (2007), “Light Scattering by Particles in Water, Theoretical and Experimental Foundations”, 704 pp., Elsevier, Amsterdam.

Junge, C. E. (1963), “Air Chemistry and Radioactivity”, Academic Press, New York.

Khelifa, A. and P. S. Hill (2006), “Models for effective density and settling velocity of flocs”, J. Hydraul. Res., 44, pp. 390-401.

Kitchen, J. C. and J. R. V. Zaneveld (1992), “A three-layered sphere model of the optical properties of phytoplankton”, Limnol. Oceanogr. 37(8), pp. 1680–1690.

Kullenberg, G. (1968), “Scattering of light by Sargasso Sea water”, Deep-Sea Res. 15, pp. 423–432.

Lee, M.E. and M. R. Lewis (2003), “A new method for the measurement of the optical volume scattering function in the upper ocean”, J. Atmos. Ocean. Technol., 20, pp. 563–571.

Lefering, I., F. Bengil, C. Trees, R. Röttgers, D. Bowers, A. Nimmo-Smith, J. Schwarz, and D. McKee (2016), "Optical closure in marine waters from in situ inherent optical property measurements", Opt. Express, 24, pp. 14036-14052.

Lide, D. R. (Ed.) (1997), "Physical and optical properties of minerals". In "CRC Handbook of Chemistry and Physics", 77th ed., pp. 4130-4136, CRC Press, Boca Raton, Fla.

Maffione, R. A. and D. R. Dana (1997), "Instruments and methods for measuring the backward-scattering coefficient of ocean waters", Appl. Opt., 36, pp. 6057–6067.

Marañón, E., P. Cermeño, J. Rodríguez, M. V. Zubkov, and R. P. Harris (2007), "Scaling of phytoplankton photosynthesis and cell size in the ocean", Limnol. Oceanogr., 52, pp. 2190–2198. https://doi.org/10.4319/lo.2007.52.5.2190.

McCrone, W. C., R.G. Draftz, and J. G. Delly (1967), "The Particles Atlas", Ann Arbor Science Publ., Ann Arbor.

McKee, D., J. Piskozub, R. Röttgers, and R. A. Reynolds (2013), "Evaluation and Improvement of an Iterative Scattering Correction Scheme for in situ Absorption and Attenuation Measurements", J. Atmos. Oceanic Technol., 30, pp. 1527–1541.

Mei, Z. P., Z. V. Finkel, and A. J. Irwin (2009), "Light and nutrient availability affect the size-scaling of growth in phytoplankton", J. Theor. Biol., 259, pp. 582–588. https://doi.org/10.1016/j.jtbi.2009.04.018.

Menden-Deuer, S., and E. J. Lessard (2000), "Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton", Limnol. Oceanogr., 45(3), 2000, pp. 569–579

Mie, G. (1908), "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen", Annalen der Physik, 330 (3), pp. 377–445.

Mikulski, C. M., S. L. Morton, and G. J. Doucette (2005), “Development and application of LSU rRNA probes for Karenia brevis in the Gulf of Mexico”, USA Harmful Algae, 4, pp. 49-60.

Mishchenko, M. I., L. D. Travis, D. W. Mackowski (1996), “T-matrix computations of light scattering by nonspherical particles: A review”, J.Quant.Spectrosc.Radiat.Transf., 55(5), pp. 535-575, DOI: 10.1016/0022-4073(96)00002-7

Mobley, C. D. (1994), “Light and water: Radiative transfer in natural waters”, Academic Press, San Diego.

Mobley C. D., L. K. Sundman and E. Boss (2002), “Phase function effects on oceanic light fields”, Appl Opt., 41(6), pp. 1035-1050.

Montagnes, D. J. S., J. A. Berges, P. J. Harrison, and F. J. R. Taylor (1994), “Estimating carbon, nitrogen, protein and chlorophyll a from volume in marine phytoplankton”. Limnol. Oceanogr., 39, pp. 1044–1060.

Morel, A. (1965), “Interprétation des variations de la forme de l'indicatrice de diffusion de la lumière par les eaux de mer”, Annales de Géophysique, 21, pp. 281-284.

Morel, A. and Y.-H. Ahn (1990), “Optical efficiency factors of free-living marine bacteria: influence of bacterioplankton upon the optical properties and particulate organic carbon in oceanic waters”, J. Mar. Res., 48, pp. 145-175.

Morel, A. (1991). “Optics of marine particles and marine optics”, in “Particle analysis in Oceanography”, S. Demers (ed.), NATO ASI Series, Vol G27, pp. 141-188.

Moutier, W., L. Duforêt-Gaurier, M. Thyssen, H. Loisel, X. Mériaux, L. Courcot, D. Dessailly, and S. Alvain (2016), "Scattering of individual particles from cytometry: tests on phytoplankton cultures," Opt. Express 24, pp. 24188-24212.

Ngo, D., S. Christesen, and G. Videen (1996), “Light Scattering from Nonconcentric Spheres”, (No. ERDEC-TR-329). Edgewood Research Development And Engineering Center, Aberdeen Proving Ground MD.

Oishi, T. (1990), “Significant relationship between the backward scattering coefficient of sea water and the scatterance at 120°”, Appl. Opt., 29, pp. 4658–4665.

Olson, R. J. and Sosik, H. M. (2007), “A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot”, Limnol. Oceanogr.: Methods 5, pp. 195–203.

Pegau, W. S., D. Gray, and J. R. V. Zaneveld (1997), “Absorption and attenuation of visible and near-infrared light in water: Dependence on temperature and salinity”, Appl. Opt., 36, pp. 6035–6046.

Peng, F. and S. W. Effler (2007), “Suspended minerogenic particles in a reservoir: Light-scattering features from individual particle analysis”, Limnol. Oceanogr., 52, pp. 204–216.

Petzold, T. J. (1972), “Volume scattering functions for selected ocean waters”, Visibility Laboratory Tech. Report, pp. 72-78, Scripps Institution of Oceanography, San Diego.

Purcell, E. M. and C. R. Pennypacker (1973), “Scattering and Absorption of Light by Nonspherical Dielectric Grains”, Astrophysical Journal, 186, pp. 705-714.

Quinten, M. (2011), “Optical Properties of Nanoparticle Systems”, Wiley-VCH, Berlin.

Reynolds, R. A., D. Stramski, V. M. Wright, and S. B. Woźniak (2010), “Measurements and characterization of particle size distributions in coastal waters”, J. Geophys. Res., 115, C08024, doi:10.1029/2009JC005930.

Risović, D. (1993), “Two-component model of sea particle size distribution”, Deep. Sea Res. Part I 40, pp. 1459–1473.

Röttgers, R., D. McKee, and S. B. Wozniak (2013), “Evaluation of scatter corrections for ac-9 absorption measurements in coastal waters”, Methods Oceanogr., 7, pp. 21-39.

Röttgers, R., K. Heymann, and H. Krasemann (2014), “Suspended matter concentrations in coastal waters: Methodological improvements to quantify individual measurement uncertainty”, Estuar Coast Shelf Sci., 151, pp. 148–155.

Sarno, D., W. H. C. F. Kooistra, L. K. Medlin, I. Percopo, and A. Zingone (2005), “Diversity in the genus Skeletonema (Bacillariophyceae). II. An assessment of the taxonomy of S. costatum-like species with the description of four new species”, Journal of Phycology, 41, pp. 151–176, doi:10.1111/j.1529-8817.2005.04067.x

Sequoia Scientific Inc. (2013), “LISST-100x Particle Size Analyzer”, User’s Manual, Version 5.0.

Shapiro, H. M. (1988), “Practical flow cytometry”, 4th ed., Wiley-Liss, Inc., New York.

Shapiro, H. M. and W. G. Telford (2009), “Lasers for Flow Cytometry”, Curr. Protoc. Cytom. 49:1.9.1-1.9.17, DOI: 10.1002/0471142956.cy0109s49.

Sharma, S. K., G. Ghosh, and D. J. Somerford (1997), “S approximation for light scattering by an infinitely long cylinder”, Appl. Opt. 36, pp. 6109-6114.

Sheldon, R. W., A. Prakash, and W. H. Sutcliffe, Jr. (1972), “The size distribution of particles in the ocean”, Limnol. Oceanogr. 17, pp. 327-340.

Slade, W. H. (2006), FASTMie, http://www.scattport.org/www.scattport.org/index.php/light-scattering-software/mie-type-codes/list/264-fastmie.html

Sosik, H. M. and Olson, R. J. (2007), “Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry”, Limnol. Oceanogr.: Methods 5, pp. 204–216.

Stramski, D., A. Morel, and A. Bricaud (1988), “Modelling the light attenuation and scattering by spherical phytoplankton cells: A retrieval of the bulk refractive index”, Appl. Opt., 27 (19), pp. 3953-3956.

Stramski, D. and D. A. Kiefer (1990), “Optical properties of marine bacteria”. In Spinrad.R.W. (ed.), “Ocean Optics X. Proceedings 16-18 April”, Orlando, Florida. SPIE, 1302, pp. 250-268.

Stramski. D. and A. Morel (1990), “Optical properties of photosynthetic picoplankton in different physiological states as affected by growth irradiance”. Deep-Sea Res., 37, pp. 245-266.

Stramski, D. and D. A. Kiefer (1991), “Light scattering by microorganisms in the open ocean”, Prog. Oceanog., Vol. 28, pp. 343-383.

Stramski, D. and C. D. Mobley (1997), “Effects of microbial particles on oceanic optics: A database of single-particle optical properties”, Limnol. Oceanogr. 42, pp. 538-549.

Strickland, J. D. H. and T. R. Parsons (1972), “A practical handbook of seawater analysis”, Fisheries Research Board of Canada, Ottawa, Ontario.

Sullivan, J. M. and M. S. Twardowski (2009), “Angular shape of the oceanic particulate volume scattering function in the backward direction”, Appl. Opt., 48, pp. 6811-6819.

Sullivan, J. M., M. S. Twardowski, J. Ronald, V. Zaneveld, and C. C. Moore (2013), “Measuring optical backscattering in water”, in: Light Scattering Reviews 7, edited by: Kokhanovsky, A. A., Springer Berlin Heidelberg, pp. 189–224.

Tan, H., R. Doerffer, T. Oishi, and A. Tanaka (2013), “A new approach to measure the volume scattering function”, Opt. Express 21, pp. 18697-18711.

Thyssen, M., G. A. Tarran, M. Z. Zubkov, R. J. Holland, G. Gregori, P. H. Burkill, and M. Denis (2008), “The emergence of automated high-frequency flow cytometry:

revealing temporal and spatial phytoplankton variability", J Plankton Res, 30, pp. 333–343.

Thyssen, M., B. Beker, D. Ediger, D. Yilmaz, N. Garcia, and M. Denis (2011), "Phytoplankton distribution during two contrasted summers in a Mediterranean harbour: combining automated submersible flow cytometry with conventional techniques" Environ. Monit. Assess., 173 (1–4), pp. 1-16.

Tomas, C. R., G. R. Hasle, E. E. Syvertsen, K. A. Steidinger, K. Tangen, J. Throndsen, and B. R. Heimdal (1997), "Identifying marine phytoplankton", Academic Press, San Diego.

Traykovski, P., R. Latter, and J.D. Irish (1999), "A laboratory evaluation of the laser in situ scattering and transmissometry instrument using natural sediments", Marine Geology, 159, pp. 355– 367, doi:10.1016/S0025-3227(98)00196-0.

Twardowski, M. S., J. M. Sullivan, P. L. Donaghay, and J. R. V. Zaneveld (1999), "Microscale quantification of the absorption by dissolved and particulate material in coastal waters with an ac-9", J. Atmos. Oceanic Technol., 16, pp. 691–707.

Twardowski, M. S., E. Boss, J. B. Macdonald, W. S. Pegau, A. H. Barnard, and J. R. V. Zaneveld (2001), "A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters", J. Geophys. Res., 106(C7), pp. 14129–14142, doi:10.1029/2000JC000404.

Ulloa, O., S. Sathyendranath and T. Platt (1994), "Effect of the particle-size distribution on the backscattering ratio in seawater", Applied Optics, 33(30).

van de Hulst, H. C. (1957), "Light scattering by small particles", John Wiley & Sons, Inc., New York.

Verity, P. G., C. Y. Robertson, C. R. Tronzo, M. G. Andrews, J. R. Nelson, and M. E. Sieracki (1992), "Relationship between cell volume and carbon and nitrogen content of marine photosynthetic nanoplankton", Limnol. Oceanogr., 37, pp. 1434–1446.

Volten, H., J. F. Haan, J. W. Hovenier, R. Schreurs, W. Vassen, A. G. Dekker, H. J. Hoogenboom, F. Charlton, and R. Wouts (1998), "Laboratory measurements of angular distributions of light scattered by phytoplankton and silt," Limnol. Oceanogr. 43(6), pp. 1180–1197.

Waterbury, J. B., S. W. Watson,R. R. L. Guillard, and L. E. Brand (1979), "Wide-spread occurrence of a unicellular, marine planktonic, cyanobacterium", Nature, 277 (5694), pp. 293–294, doi:10.1038/277293a0

Waterman, P.C. (1965), "Matrix formulation of electromagnetic scattering", Proceedings of the IEEE, 53(8), pp. 805-812, DOI: 10.1109/PROC.1965.4058

Waterman, P.C. (1971), "Symmetry, unitarity, and geometry in electromagnetic scattering", Physical Review D, 3(4), pp. 825-839, DOI: 10.1103/PhysRevD.3.825

WET Labs Inc. (2008), "Absorption and attenuation meter, ac-9", User's Guide, Revision Q.

WET Labs Inc. (2010), "Scattering meter, ECO BB9", User's Guide, Revision K.

WET Labs Inc. (2011), "ac Meter", Protocol Document, Revision Q.

Xu, F., K. Ren, G. Gouesbet, G. Gréhan, and X. Cai (2007), "Generalized Lorenz-Mie theory for an arbitrarily oriented, located, and shaped beam scattered by a homogeneous spheroid", J. Opt. Soc. Am. A 24, pp. 119-131.

Yang, P. and K. N. Liou (1996), "Finite-difference time domain method for light scattering by small ice crystals in three-dimensional space", J. Opt. Soc. Am. A, 13, pp. 2072-2085.

Yang, P. and K. N. Liou (1996), "Geometric-optics–integral-equation method for light scattering by nonspherical ice crystals", Appl. Opt., 35, pp. 6568-6584.

Yurkin, M. A. and Hoekstra, A. G. (2007), “The discrete dipole approximation: an overview and recent developments”, J.Quant.Spectrosc.Radiat.Transf. 106, pp. 558-589, DOI: 10.1016/j.jqsrt.2007.01.034.

Zaneveld, J. R. V., D. M. Roach, and H. Pak (1974), “The determination of the index of refraction distribution of oceanic particles”, J. Geophys. Res., 79, pp. 4091-4095.

Zaneveld, J. R. V., J. C. Kitchen, and C. Moore, “The scattering error correction of reflecting-tube absorption meters”, Ocean Optics XII, June 13-15 1994, Bergen, Norway, SPIE, pp. 44-55.

Zapata, M., F. Rodrígues, and J. L. Garrido (2000), “Separation of chlorophylls and carotenoids from marine phytoplankton: A new HPLC method using a reversed phase C8 column and pyridine-containing mobile phases”, Mar. Ecol. Prog. Ser., 195, pp. 29–45.

Zhang, X., M. Twardowski, and M. Lewis (2011), “Retrieving composition and sizes of oceanic particle subpopulations from the volume scattering function”, Appl. Opt. 50, pp. 1240-1259.

Zhang, X., R.H. Stavn, A.U. Falster, D. Gray, and R.W. Gould Jr. (2014), “New insight into particulate mineral and organic matter in coastal ocean waters through optical inversion”, Estuarine, Coastal and Shelf Science, 149, pp. 1–12

A. Instrument operation

In this appendix section a primer on the operation and maintenance of the CytoBuoy CytoSense is given as a reference manual for the reader. While it is no substitute for actual practice and experience in flow cytometer operation, it should nevertheless equip the reader with the knowledge necessary to carry out measurements with the CytoSense and keep the instrument in working order.

A.1 Measurement setup

The CytoSense is controlled by the operator through its dedicated software, CytoUSB. Here the instrument can be operated either in free mode or by setting up one or more executable measurements.

In *free mode* the instrument does not log measurements, but its pumps and peripherals can be actively and dynamically controlled by the operator. This mode is especially useful for maintenance. The most important instrument features to be controllable in free mode are the sheath fluid vacuum pump (the global flow rate of which is also set at this level), the sample peristaltic pump, and the laser.

An *executable measurement* is a measurement setup which can be edited and saved by the operator. It can be executed alone or within a series of other measurements which are then carried out sequentially by the software. Unlike free mode operation, the results of an executable mesurement are logged and saved by the software in a dedicated file format (.cyz) which is specific to the CytoClus data processing software supplied by CytoBuoy. The main variables which can be set in the measurement setup are: sample peristaltic pump flow rate for the measurement; triggering channel(s); PMT sensitivity for the side scattering and red fluorescence channels; and end-of-measurement conditions, which can be defined on a set time duration, a set number of detected particles, a set processed sample volume or a combination thereof (with the measurement ending as soon as at least one of the conditions is met). The CytoSense also offers the possibility of taking up to 150 pictures of the particles, triggered either in smartgrid mode or in gated mode. The so-called *smartgrid mode* uses an algorithm to take pictures of particles in a logarithmically even fashion across the value range of a variable of choice. *Gated mode* takes pictures of particles within a range of x/y values for two variables of choice instead. The range can be selected and saved by the operator within the dedicated the CytoClus data processing software, to be later invoked in CytoUSB within the executable measurement setup.

A.2 Maintenance

A.2.1 Sample carryover and system cleaning

Flow cytometers are delicate devices and great care has to be taken to maintain them in working order. One of the most common issues in flow cytometry is sample carryover from previous runs. Particles will often stick to the walls of tubing, flow cell and other components within the fluidics system, only to be dislodged by the continued flow at later times and mixed with subsequent samples. The problem can be negligible when measuring a series of samples which are similar and have similar concentrations, even more so if the samples are measured in order of concentration, from lowest to highest, but highly concentrated samples can seriously affect subsequent measurements. Filtering systems can be employed, as is the case with the CytoSense, but can also act as reservoirs of particles and exacerbate the problem if no other care is taken; keeping the instrument reasonably clean therefore calls for a combination of approaches, which were followed in this study in the fashion detailed below.

The fluidics system must be run often even during downtime and continuously during a measurement session. During downtime periods in which the instrument was not actively used for measurement, the CytoSense was kept running using copious amounts of ultrapure filtered and deionized water (Milli-Q) through the sample line, 3-4 days a week, ~6 hours a day. No performance analysis was carried out on the length of rinsing periods during downtime, which was chosen to be as long as practically possible (the compromise here being with the amount of waste fluid generated). The sheath pump of the CytoSense was also kept on at all times during downtime rinsing and especially between and during sample runs. This continuously circulates the sheath fluid through the filters and keeps the core flow confined, impeding particle-rich, turbulent plumes from reaching the recesses of the flow cell and injection stages.

The filters must be cleaned periodically. The filters within the CytoSense were cleaned at irregular intervals following episodes of fluidics fouling or evidence of large particulate buildup in the measurements. The task can be accomplished through flushing and backflushing of the filters with a weak ~0.5% sodium hypochlorite solution using 50 ml Luer-lock plastic syringes (as suggested by CytoBuoy, private communication), following this procedure:

- Clamp the tubing sections around both ends of each filter.
- Remove the filters.
- Backflush each of the filters with plenty of Milli-Q water.
- Fill each of the filters with the sodium hypochlorite solution for about 30 seconds.
- Flush the filters with plenty of Milli-Q water before installing them back on the instrument.
- Restore and unclamp all the tubing.

The flow cell itself may also be cleaned. A good indicator of the necessity to clean the flow cell is diminishing scattering values in subsequent measurements, although the effect is usually only visible on long time scales. The cleaning procedure is simple and can be accomplished with a solution of Milli-Q water, Tween20 non-ionic viscous detergent (Sigma-Aldrich Co. LLC.) at 0.1% concentration and isopropyl alcohol at 20% concentration (as suggested by CytoBuoy, private communication):

- Run the detergent solution through the sample line for a few minutes at high sample pump flow rate (~5-6 μl/s).
- Lower the sample pump flow rate (~1-2 μl/s) and turn off the sheath pump.

Leave the system running in this fashion for a few minutes to let the solution diffuse into the flow cell and injectors.

- Turn off the sample pump and let the system settle for 2-3 minutes.
- Turn on both sheath and sample pumps again and rinse with plenty of Milli-Q water.

Alternatively, the flow cell can be manually cleaned using the same solution and 50 ml Luer-lock plastic syringes.

- Make sure that both sample and sheath pumps are turned off.
- Isolate the flow cell by clamping before and after the two Luer locks located where the sheath line separates into the two sheath fluid injection lines and after the flow cell. Open the Luer locks.
- Fill a 50 ml Luer-lock plastic syringe with the detergent solution and connect it to one of the open ends; connect a second empty syringe to the other open end.
- Manually flush and backflush the solution between the two syringes, multiple times.
- Restore and unclamp all the tubing.
- Turn on both sheath and sample pumps again and rinse with plenty of Milli-Q water.

Finally, the CytoSense flow cell may also be accessed vertically through an o-ring sealed screw opening on the top.

- Make sure that both sample and sheath pumps are turned off.
- Remove the screw at the top of the flow cell.

- Dip a small brush in the detergent solution (bespoke brushes were supplied with the instrument).
- Insert the brush in the flow cell and scrub gently.
- Place the sealing screw back – don't tighten excessively.
- Turn on both sheath and sample pumps again and rinse with plenty of Milli-Q water.

A.2.2 Core stability and laser alignment

A second major concern in flow cytometry is the stability of the core flow and its alignment with respect to the flow cell and the laser beam. Bad laser alignment generally produces bad quality data as the particles intercept the least flat and least bright parts of the radiant power distribution, while bad stability of the core flow leads to variance in the scattering values produced by particles, especially evident in the case of particles of the same type and size.

Bad laser alignment is caused by the laser beam not being centred between the walls of the flow cell and on the centre of the flow core. For the CytoSense, the laser alignment procedure is simple, as the mirror controlling the horizontal position of the laser beam can be moved by rotating a single screw. However, the laser alignment procedure requires the laser light from the instrument to be shone on a screen. Take all the precautions that apply.

- Remove the laser beam stop.
- Turn on the laser and shine it upon a screen, preferably rigid and writable upon.
- Turn the alignment screw in one direction until the flow cell wall position becomes evident on the surface. Mark it with a pen.
- Turn the screw in the opposite direction and repeat the procedure.

- Turn the screw until the laser light rests in the middle between the two wall marks.

- Put the beam stop back.

Bad stability of the core is usually caused by either or both of two processes. The first is the presence of air bubbles in the injectors; actual clogging is less problematic for the CytoSense than for most flow cytometers as the instrument has a large operating size range and can accept particles a few hundred microns wide. Air bubbles may be simply removed by clamping the tubing leading to the sheath injectors for a couple of seconds while the sheath pump is running. The sudden release of pressure buildup is normally sufficient to dislodge the air bubbles. The second cause of bad core stability is a badly arranged geometry of the sample and sheath fluid injectors. In the CytoSense the relative geometry of sample injector and sheath fluid injectors can be acted upon using two sets of screws: once again, the laser alignment procedure requires the laser light from the instrument to be shone on a screen. Take all the precautions that apply.

- Remove the imaging camera module.

- Install the viewer supplied with the instrument. The viewer magnifies the flow cell and lets the operator verify the position of the flow core relative to the centre of focus of the beam, parallel to the laser beam direction.

- Remove the laser beam stop.

- Turn on the laser and shine it upon a screen, preferably rigid and writable upon. If needed, repeat the marking procedure as described for laser alignment. This lets the operator verify the position of the flow core relative to the centre of the laser beam, perpendicular to the laser beam direction.

- Turn on both sheath and sample pumps. Run the detergent solution through the sample line for contrast.

- Align the field of view of the viewer until the laser going through the flow cell is clearly visible.

- Use the *outer* set of screws (those without springs) to align the flow with the laser. The core must rest in the middle of the laser beam on the screen and where the beam is thinnest in the viewer. It is advisable to move the core as close as possible to the center of the flow cell and to finalize alignment using the alignment screw of the laser mirror.

- Use the *inner* set of screws to adjust the alignment of the three second-stage sheath injectors. Unfortunately, this is a trial-and-error process: slowly adjust the three screws until the core can be seen reaching a stable configuration on both screen and viewer.

Ideally, the core must be both stable (no bumps or vibrations apart from the periodic ones induced by the peristaltic pump) and coherent (no splitting in double or triple cores). Care must also be taken not to loosen the inner screws excessively, as the flow cell assembly is likely to start leaking fluid.

B. Flow cytometric method of size and RI detection: MATLAB code

In this first appendix, the core sections of the MATLAB code which was used to determine PSDs and PRIDs of the particle populations are presented along with a flow chart of the code structure (Fig. B.1). The function of each line is described in the comments within the code. Code sections which require data inputs from the user are highlighted in light blue, final outputs are highlighted in orange.

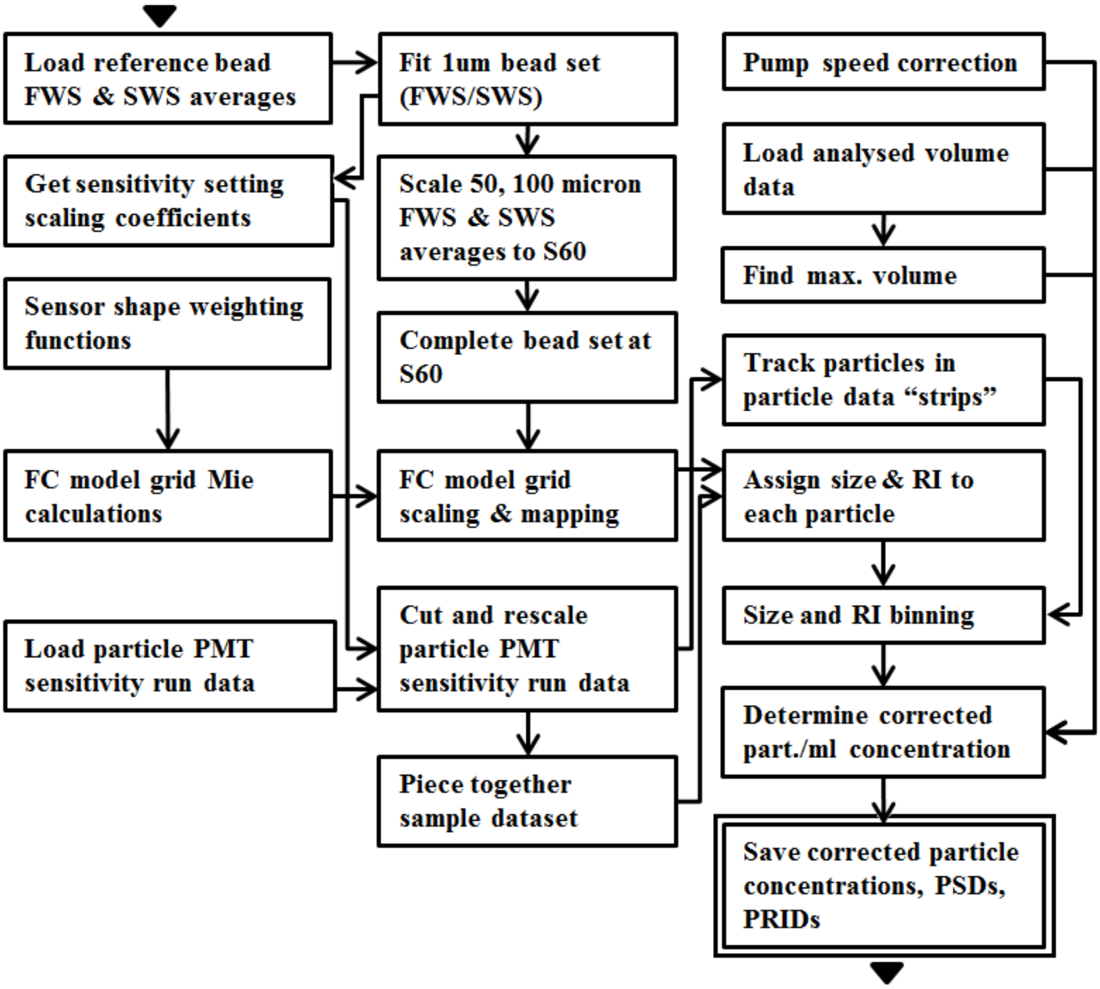


Figure B.1 – Flow chart for the FC method of size and RI detection procedure described in Chapter 6. The pump speed correction is the one presented in Paragraph 3.4.3.3.

B.1 Angular ranges, sensor shape and PMT sensitivity scaling coefficients

```
%%% 1.1 Angular ranges

%%% The next few lines define the scattering angles array. The
%%% resolution is higher at small angles.

theta1 = (0:0.01:1);      % Angle range 1.
theta2 = (1.1:0.1:10);    % Angle range 2.
theta3 = (11:1:180);      % Angle range 3.

theta = [theta1 theta2 theta3];     % Scattering angles array.
thetalength = numel(theta);         % Gets the number of scattering
angles.
theta_rad = theta./180*pi;          % Converts the scattering angles
array into radians.

%%% The collection angle ranges of the forward and side scattering
%%% sensors are defined here. These fine-tune the shape of the model %%%
curves.

ang1 = 2;                 % Forward scattering sensor MIN angle.
ang2 = 9.7;               % Forward scattering sensor MAX angle.

ang3 = 45;                % Side scattering sensor MIN angle.
ang4 = 135;               % Side scattering sensor MAX angle.

[~,Iindex1] = min(abs(theta-ang1));    % Singles out the element in the
scattering angles array corresponding to ang1 for integration purposes.
[~,Iindex2] = min(abs(theta-ang2));    % Singles out the element in the
scattering angles array corresponding to ang2 for integration purposes.
[~,Iindex3] = min(abs(theta-ang3));    % Singles out the element in the
scattering angles array corresponding to ang3 for integration purposes.
[~,Iindex4] = min(abs(theta-ang4));    % Singles out the element in the
scattering angles array corresponding to ang4 for integration purposes.

dtheta = ones(1,thetalength);          % Preallocates the array which
contains the increments between angles in the scattering angles array.
For integration purposes.

for S1 = 1:1:(thetalength-2)           % This small loop fills the angle
increments array.
      dtheta(S1+1) = (theta_rad(S1+2)-theta_rad(S1))/2;
end;

dtheta(1) = dtheta(2);    % Defines the first element of dtheta.
dtheta(thetalength) = dtheta(thetalength-1);      % Defines the last
element of dtheta.

sintheta = sin(theta_rad);             % The sines of the scattering
angles.

%%% 1.2 Sensor shape weighting functions

%%% This section defines the weighting functions which account for the
%%% collecting area of the sensors.
```

```
% Forward scattering weighting function
shapecorr1=((.5*pi)-asin(sin(theta_rad(Iindex1))./ ...
sin(theta_rad(Iindex1:Iindex2)))))/(.5*pi);

% Side scattering weighting function element 1
shapecorr2space=sin(linspace(0,pi,numel(theta_rad(Iindex3:Iindex4))));

% Side scattering weighting function element 2
Kprime = sin(theta_rad(Iindex4)-(.5*pi));

% Side scattering weighting function element 3
shapecorr2core = Kprime*shapecorr2space;

% Side scattering weighting function
shapecorr2=asin(shapecorr2core)/(.5*pi);

%%% 1.3 Reference polymer beads and PMT sensitivity settings

%%% This section will define reference polymer bead scattering averages
%%% and PMT sensitivity settings which will be used both to scale and
%%% centre the scattering values of the FC model grid/look-up table and
%%% to combine data from multiple sensitivity runs into a single sample
%%% dataset. PMT setting 60 was the one used as reference.

sensSetting = [50;60;70;80];          % Sensitivity settings array.
sensS_num = numel(sensSetting);       % Gets the number of elements in
the sensitivity settings array.

%%% The following line will require input from the user. It is meant to
%%% collect average scattering values for 1 micron reference beads into
%%% a 4x2 matrix, with each row corresponding to a PMT setting and the
%%% two columns corresponding to forward and side scattering
%%% respectively.

set1um = [avgs_1um{1};avgs_1um{2};avgs_1um{3};avgs_1um{4}];   % The set
of scattering averages corresponding to 1 micron beads

%%% The following line will also require input from the user. It is
%%% meant to collect average scattering values for 5 and 10 micron
%%% reference beads into a 2x2 matrix, with each row corresponding to a
%%% bead diameter and the two columns corresponding to forward and side
%%% scattering respectively.

setMed60 = [avgs_5um;avgs_10um];      % The set of scattering averages
corresponding to 5 and 10 micron beads

%%% The following line will require a final input from the user. It is
%%% meant to collect average scattering values for 50 and 100 micron
%%% reference beads into a 2x2 matrix, with each row corresponding to a
%%% bead diameter and the two columns corresponding to forward and side
%%% scattering respectively. PMT setting 50 was the only one for which
%%% side scattering of 50 and 100 micron reference beads didn’t
%%% saturate.

setLarge50 = [avgs_50um;avgs_100um];  % The set of scattering averages
corresponding to 50 and 100 micron beads
```

```
%%% This next section uses the 1 micron set of bead scattering averages
%%% to work out scaling coefficients necessary to compose data
%%% corresponding to each sensitivity setting into one single dataset.
%%% The type of the following fit models was determined after direct
%%% examination of the data.

f1 = coeffvalues(fit(sensSetting,set1um(:,1),'poly1'));    % Calculates
the linear fit of the FWS averages of the 1 um set.
f2 = coeffvalues(fit(sensSetting,set1um(:,2),'power1'));   % Calculates
the power model fit of the SWS averages of the 1 um set.

fws50coeff = set1um(2,1)/(f1(1)*sensSetting(1)+f1(2));      % FWS scaling
coefficient for sensitivity setting 60.
sws50coeff = set1um(2,2)/(f2(1)*(sensSetting(1)^f2(2)));   % SWS scaling
coefficient for sensitivity setting 60.
fws70coeff = set1um(2,1)/(f1(1)*sensSetting(3)+f1(2));      % FWS scaling
coefficient for sensitivity setting 70.
sws70coeff = set1um(2,2)/(f2(1)*(sensSetting(3)^f2(2)));   % SWS scaling
coefficient for sensitivity setting 70.
fws80coeff = set1um(2,1)/(f1(1)*sensSetting(4)+f1(2));      % FWS scaling
coefficient for sensitivity setting 80.
sws80coeff = set1um(2,2)/(f2(1)*(sensSetting(4)^f2(2)));   % SWS scaling
coefficient for sensitivity setting 80.

sensCorrCoeff = [[fws50coeff;1;fws70coeff;fws80coeff], ...
[sws50coeff;1;sws70coeff;sws80coeff]];       % Scaling coefficient
matrix

%%% The next two lines scale the scattering averages corresponding to
%%% 50 and 100 micron beads to PMT setting 60

fwsLarge60 = setLarge50(:,1)*sensCorrCoeff(1,1);
swsLarge60 = setLarge50(:,2)*sensCorrCoeff(1,2);

%%% The next line compiles the set of bead scattering averages
%%% corresponding to reference PMT setting 60. It also requires input
%%% of the 0.5 micron bead scattering averages at PMT setting 60, which
%%% was the lowest setting for which 0.5 micron beads were detectable.

modelBeadSet = [[avgs_05um(1);avgs_1um{2}(1);setMed60(:,1); ...
fwsLarge60],[avgs_05um(2);avgs_1um{2}(2);setMed60(:,2);swsLarge60]];
```

B.2 FC model grid calculations, scaling and mapping

```
%%% 2.1 Refractive indices & wavelength

%%% The refractive indices used in the model are defined here, as well
%%% as the wavelength of the laser light incident on the particles.

wnr = 1.333;         % Water refractive index

lambda_air = 488e-9;          % Laser light wavelength in air.
lambda = lambda_air/wnr;      % Laser light wavelength in water.

%%% The refractive index of the standard is put first to simplify
%%% scaling later on
```

```
nr = [1.595 (1.335:0.01:1.585) (1.605:0.01:1.725)]./wnr;     % Relative
real refractive indices array.
NV = numel(nr);     % Gets the number of elements in the real refractive
index array.
ni = 0;             % Imaginary refractive index - set at negligible
absorption.

%%% 2.2 Size & refractive index isolines i.e. Look-up table

%%% This section of code calculates scattering value nodes, forming
%%% curves of constant size and refractive index. These nodes
%%% constitute a look-up table of scattering values which will relate
%%% particle scattering to particle physical properties. Calculations
%%% are based on Mie theory and handled by an implementation of the
%%% FASTMie code by W. H. Slade, http://www.scattport.org/index.php/
%%% light-scattering-software/mie-type-codes/list/264-fastmie.html
%%% The curves are then mapped to the dataset by using reference bead
%%% scattering averages.

%%% Two lines of code reassigning the two columns in the modelBeadSet
%%% array to two different variables.

beadsFSC = modelBeadSet(:,1);
beadsSSC = modelBeadSet(:,2);

%%% An array of particle sizes is now generated.

D = (logspace(-8,-4,300));      % Array of log-spaced virtual particle
diameters.
r = D/2;                        % Array of log-spaced virtual particle
radii.
RV = numel(D);                  % Gets the number of virtual particles.

sizes = [0.498;0.994;4.993;10.12;50.2;100]*1e-6;
% The actual NIST diameters for the 6 bead groups used as standard, in
metres.

[~,Bindex1] = min(abs(D-sizes(1)));      % Singles out the element in
the diameters array corresponding to sizes(1).
[~,Bindex2] = min(abs(D-sizes(2)));      % Singles out the element in
the diameters array corresponding to sizes(2).
[~,Bindex3] = min(abs(D-sizes(3)));      % Singles out the element in
the diameters array corresponding to sizes(3).
[~,Bindex4] = min(abs(D-sizes(4)));      % Singles out the element in
the diameters array corresponding to sizes(4).
[~,Bindex5] = min(abs(D-sizes(5)));      % Singles out the element in
the diameters array corresponding to sizes(5).
[~,Bindex6] = min(abs(D-sizes(6)));      % Singles out the element in
the diameters array corresponding to sizes(6).

If = ones(1,RV);        % Preallocates the array which will contain the
computed total forward scattered intensity of the virtual particles.
Is = ones(1,RV);        % Preallocates the array which will contain the
computed total side scattered intensity of the virtual particles.
```

```
FSC = ones(NV,RV);      % Preallocates the array which will store
modeled forward scatter for each refractive index/diameter pair.
SSC = ones(NV,RV);      % Preallocates the array which will store
modeled side scatter for each refractive index/diameter pair.
SIZEs = repmat(D,NV,1);       % Preallocates the array which will store
the virtual particle diameter for each refractive index/diameter pair.
RIs = repmat(nr',1,RV);       % Preallocates the array which will store
the virtual particle real RI for each refractive index/diameter pair.

%%% The following loop calculates the total forward and side scatter
%%% for each virtual particle.

for N1=1:1:NV     % Iterates through refractive indices.
      for R1=1:1:RV        % Iterates through radii.
            m = nr(N1)+(1i*ni);      % Complex refractive index.
            k = 2*pi/lambda;         % Wavenumber.
            x = k*r(R1);             % Size parameter.
            [S1,S2,Qb,Qc,Qbb] = fastmie(x,m,[],theta_rad);     % The
core of the calculation is handled by the FASTMie script (by W. H.
Slade, 2006).
            i1=abs(S1).^2;           % Scattered intensity functions.
            i2=abs(S2).^2;             % Scattered intensity functions.

%%% The total forward and side scatter are now calculated using
%%% an extremely stripped down version of the VSF integral
%%% normally used to calculate scattering - all missing
%%% factors and coefficients are taken care of by the
%%% subsequent mapping.

            Iratio=i1+i2;            % Simplified total scattered
intensity (angular).

            tsd = 2*pi*dtheta'.*sintheta'.*Iratio;     % The stripped
down version of the VSF integral.
            tsdf = tsd(Iindex1:Iindex2);              % Preparing the
forward scatter sensor angle integration.
            tsdfMOD = shapecorr1.*tsdf';              % Applying the
sensor shape correction.
            If(:,R1) = sum(tsdfMOD);                  % Integrates and
computes the total forward scattering by the virtual particle (N1,R1).

tsds = tsd(Iindex3:Iindex4);             % Preparing the side scatter
sensor angle integration.
            tsdsMOD = shapecorr2.*tsds';              % Applying the
sensor shape correction.
            Is(:,R1) = sum(tsdsMOD);                  % Integrates and
computes the total forward scattering by the virtual particle (N1,R1).
      end

%%% 2.2.1 FC model grid/look-up table mapping

%%% In the following section the script takes the refractive index
%%% isolines it has computed and maps them to the standard beadset.
%%% To do so, it first calculates the X-axis and Y-axis
%%% displacement between the averages for the 0.5 um bead group and
%%% its virtual counterpart. It then translates the whole isoline
%%% so that the positions of the two match before rescaling the
%%% whole isoline by matching the X-axis and Y-axis distance
```

```
%%% between the 0.5 and 1 um bead groups and its virtual
%%% counterpart.

      if N1 == 1            % The mapping is done with the RI nodes
corresponding to the latex beads of the standard.

            DisplX = If(:,Bindex1)-beadsFSC(1);       % 0.5 um bead
      group X-axis displacement.
            DisplY = Is(:,Bindex1)-beadsSSC(1);       % 0.5 um bead
      group Y-axis displacement.

            FSCtemp = If-DisplX;             % Applies the X-axis
      displacement to the whole isoline.
            SSCtemp = Is-DisplY;             % Applies the Y-axis
      displacement to the whole isoline.

            ScaleX = (beadsFSC(2)-beadsFSC(1))/(FSCtemp(:,Bindex2)- ...
                beadsFSC(1));         % X-Axis scaling coefficient.
            ScaleY = (beadsSSC(2)-beadsSSC(1))/(SSCtemp(:,Bindex2)- ...
                beadsSSC(1));         % Y-Axis scaling coefficient.

            FSC(N1,:)=((FSCtemp-beadsFSC(1))*ScaleX)+beadsFSC(1);      %
      FWS values of the mapped isoline.
            SSC(N1,:)=((SSCtemp-beadsSSC(1))*ScaleY)+beadsSSC(1);      %
      SWS values of the mapped isoline.
      end

%%% The process is then applied to all other refractive indices.

      FSCtemp = If-DisplX;              % Applies the X-axis
displacement to the whole isoline.
      SSCtemp = Is-DisplY;              % Applies the Y-axis
displacement to the whole isoline.

      FSC(N1,:)=((FSCtemp-beadsFSC(1))*ScaleX)+beadsFSC(1);         %
FWS values of the mapped isoline.
      SSC(N1,:)=((SSCtemp-beadsSSC(1))*ScaleY)+beadsSSC(1);         %
SWS values of the mapped isoline.
end
```

B.3 PMT sensitivity scaling of particle data and sample dataset composition

```
%%% 3.1 Rescaling and composition of measured particle scattering data

%%% Measured particle scattering data (main_data) is input here. In the
%%% present iteration of the code it is meant to be input as a cell
%%% array of four Nx3 matrices, each corresponding to a separate PMT
%%% sensitivity run for the same sample. N is the number of particles
%%% within each sensitivity run, and the three columns represent
%%% forward scattering, side scattering and red fluorescence signal
%%% respectively.
```

```
main_data_corr = main_data;           % Preallocates cell array which
will contain rescaled particle scattering values

for q = 1:sensS_num
      if not(isempty(main_data{q}))
            main_data_corr{q}(:,1) = main_data{q}(:,1).* ...
                  sensCorrCoeff(q,1);       % Sensitivity setting
      correction coefficient applied to FWS.
            main_data_corr{q}(:,2) = main_data{q}(:,2).* ...
            sensCorrCoeff(q,2);       % Sensitivity setting correction
coefficient applied to SWS.
      end
end

%%% The following part of the code will select "strips" of particle
%%% data to eliminate overlap between rescaled PMT sensitivity run data
%%% and avoid inflating particle counts.

main_data_slice = cell(sensS_num,1);       % Preallocates cell array to
contain the data "strips"

cutoff = [1e7,3500,700,200,1];             % SWS cutoff values for each
sensitivity setting. User determined by direct examination of the data
cutoff_ind = cell(sensS_num,1);            % Prepares an array of cells
to contain the arrays of indices of particles contained within the cut-
off boundaries.

for q = 1:sensS_num                          % This loop operates all
the data cuts
      if not(isempty(main_data_corr{q}))        % Checks if the cell
isn't empty (to prevent the code from stopping).
            cutoff_ind{q} = find(main_data_corr{q}(:,2) ...
            < cutoff(q) & main_data_corr{q}(:,3) >= cutoff(q+1));    %
      Gets the indices of data points in main_data_corr{q} between
      cutoff(q) and cutoff(q+1).
            main_data_slice{q} = main_data_corr{q}(cutoff_ind{q},:);  %
      Assigns those data points to main_data_slice{q}.
      end
end

%%% The data "strips" are stitched back together into one single large
%%% matrix. Fluorescence data is separated into a dedicated matrix.

main_data_total = [main_data_slice{1};main_data_slice{2}; ...
main_data_slice{3};main_data_slice{4}];            % All cells from
main_data_slice are compiled together into a single large array.

if not(isempty(main_data_total))
      main_data_total_FLR = main_data_total(:,3);      % Extracts the
Max FLR column from the main_data_total matrix.
      main_data_total = main_data_total(:,1:2);         % Reduces the
main_data_total matrix to the FWS and SWS columns only.
end

%%% The number of particles left into each of the cells inside
%%% main_data_slice are counted and kept track of.
```

```
main_data_nums = [size(main_data_slice{1},1); ...
size(main_data_slice{2},1);size(main_data_slice{3},1);...
size(main_data_slice{4},1)];
```

B.4 Size and RI binning: concentration corrections, PSDs and PRIDs

```
%%% 4.1 Size and refractive index binning

%%% Size and refractive index are assigned to each particle based on
%%% their proximity to the nodes in the FC model grid. Before binning
%%% proper, a few correction terms are needed.

%%% 4.1.1 Sample pump flow rate correction

%%% The first correction term accounts for particle concentration
%%% underestimation at low sample pump flow rates. This correction will
%%% need the parameter pumpSpeed to be set by the user to be equal to
%%% that used during the measurement. The function coeffCorrFind simply
%%% finds the correction curve value corresponding to pumpSpeed - see
%%% pump flow rate correction results, Chapter 3.

platCorr = 1.087628356;                     % Correction factor derived
from flow rate test data - see pump flow rate correction results,
Chapter 3
pumpSpeed = 0.5;                            % Pump speed in uL/s.
cCorr = coeffCorrFind(pumpSpeed);           % Sets the pump speed
correction coefficient - more on this in the function definition and
further on in this script.
coeffCorr = (1/cCorr)*platCorr;       % Defines the final flow rate
correction coefficient

%%% 4.1.2 Sensitivity run volume correction

%%% A further correction is required to account for data transfer
%%% overhead time. While the instrument software transfers particle
%%% data, it cannot analyse the sample. This means not only that the
%%% analysed volume is always less than the processed volume, but that
%%% the analysed volume depends on the amount of detected particles
%%% too. The more particles there are, the larger the difference
%%% between analysed and pumped volume. In turn, this means that the
%%% analysed volume will be different for different sensitivity
%%% settings (high sensitivity generally means that more particles are
%%% detected by the sensors). Individual analysed volumes are imported
%%% and compared to produce a correction that is applied to each
%%% sensitivity setting.

%%% The following line will require input from the user (sens_vols). It
%%% is meant to identify the largest analysed volume among those
%%% corresponding to each sensitivity run. These are provided directly
%%% by the flow cytometer data analysis software (CytoClus).

sens_vols_max = max(sens_vols);         % Gets the largest analysed
volume.

%%% 4.1.3 Binning
```

```
FSClin = reshape(FSC,[1,NV*RV]);             % Reshaping FWS array.
SSClin = reshape(SSC,[1,NV*RV]);             % Reshaping SWS array .
SIZElin = reshape(SIZEs,[1,NV*RV]);          % Reshaping diameters
array.
RIslin = reshape(RIs,[1,NV*RV]);             % Reshaping RI array.

dataN = size(main_data_total,1);           % Gets the number of data
points in main_data_total.

particleSIZEs = ones(dataN,1);               % Preallocates the array
that will contain the sizes of each particle.
particleRIs = ones(dataN,1);                 % Preallocates the array
that will contain the RIs of each particle.
NUMb = ones(dataN,1);                        % A simple array of ones.
Summing along this array effectively counts particles, and will be the
basis of the concentration calculations.

for q = 1:1:dataN                             % This loop assigns size & RI
of the closest grid point to each particle.
       [~,ind] = min(abs(FSClin-main_data_total(q,1)) + ...
       abs(SSClin-main_data_total(q,2)));      % Finds closest grid
       point.

       particleSIZEs(q) = SIZElin(ind);      % Assigns the size to the
particle.
      particleRIs(q) = RIslin(ind);          % Assigns the RI to the
particle.
end

nrsort = sort(nr);             % Sorts the real refractive index values
in ascending order
B1 = ones(1,numel(nrsort)-1);        % Preallocates the array of values
used as RI bin boundaries

for h = 1:1:numel(nrsort)-1
       B1(h) = (nrsort(h)+nrsort(h+1))/2;          % Defines the values
used as RI bin boundaries
end
B1 = [nrsort(1)-((nrsort(2)-nrsort(1))/2),B1,nrsort(end)+ ...
                ((nrsort(end)-nrsort(end-1))/2)];  % Further redefines
the first and last RI bin boundary values

B1N = numel(B1);                     % Gets the number of values used as
RI bin boundaries

%%% The following line will require input from the user (size_bins).
%%% These are log-spaced size bin boundaries.

B2 = size_bins;           % The array of log-spaced values used as size
bin boundaries
B2N = numel(B2);          % Gets the number of log-spaced values used as
size bin boundaries

binMAT = zeros(B1N-1,B2N-1);         % Preallocates the matrix that will
contain the particles after being binned by RI and by size
binMAT_FLRaye = zeros(B1N-1,B2N-1);            % Preallocates the matrix
that will contain the FL particles after being binned by RI and by size
```

```
binMAT_FLRnay = zeros(B1N-1,B2N-1);           % Preallocates the matrix
that will contain the non-FL particles after being binned by RI and by
size

loB = 1;                        % Primes the low boundary of the
particles (relative to sensitivity setting).
hiB = 0;                        % Primes the high boundary of the
particles (relative to sensitivity setting).

for u = 1:1:sensS_num
      if main_data_nums(u) ~= 0      % Checks if there are particles
corresponding to the sensitivity setting.
             if u ~= 1               % Updates the low boundary with
      each step save the first.
                   loB = loB + main_data_nums(u-1);
            end
             hiB = hiB + main_data_nums(u);    % Updates the high
      boundary.
            for tR = 1:1:B1N-1
                  indRI = find(particleRIs(loB:hiB) >= B1(tR) & ...
                        particleRIs(loB:hiB) < B1(tR+1)); % Indices of
the particles belonging to the RI bin

                  pinPoint = particleSIZEs(loB:hiB);     % Singles out
particles within the sensitivity boundaries
                  SEL_binMAT = pinPoint(indRI);    % Singles out an
array containing the particles within the boundaries and belonging to
the RI bin, ready to be binned by size.
                  NUMbSEL = NUMb(indRI);              % A reduced array
of ones. Summing along this array effectively counts particles.

                  pinPoint_FLR = main_data_total_FLR(loB:hiB);  %
Singles out FL values of particles within the sensitivity boundaries

                  SEL_binMAT_FLR_pre = pinPoint_FLR(indRI);  % Singles
out an array containing FL values of particles within the boundaries and
belonging to the RI bin, ready to be binned by size.

                  indRIFLRaye = find(SEL_binMAT_FLR_pre > 10); % Indices
of the FL particles within the boundaries and belonging to the RI bin.
10 is the value chosen to separate FL particles from the FL noise
background

                  indRIFLRnay = find(SEL_binMAT_FLR_pre <= 10); %
Indices of the non-FL particles within the boundaries and belonging to
the RI bin.

                  SEL_binMAT_FLRaye = SEL_binMAT(indRIFLRaye);  %
Singles out an array containing FL particles within the boundaries and
belonging to the RI bin, ready to be binned by size.
                  SEL_binMAT_FLRnay = SEL_binMAT(indRIFLRnay);  %
Singles out an array containing non-FL particles within the boundaries
and belonging to the RI bin, ready to be binned by size.

                  NUMbSEL_FLRaye = NUMbSEL(indRIFLRaye);       % A
reduced array of ones. Summing along this array effectively counts FL
particles.
```

```
                  NUMbSEL_FLRnay = NUMbSEL(indRIFLRnay); % A reduced
array of ones. Summing along this array effectively counts non-FL
particles.

                  for tS = 1:1:B2N-1
                        indSIZE= find(SEL_binMAT >= B2(tS) & ...
                         SEL_binMAT < B2(tS+1)); % Indices of the
particles belonging to the size bin.

                         indSIZE_FLRaye= find(SEL_binMAT_FLRaye >= ...
                   B2(tS) & SEL_binMAT_FLRaye < B2(tS+1));
% Indices of FL particles belonging to the size bin.

                        indSIZE_FLRnay= find(SEL_binMAT_FLRnay >= ...
                   B2(tS) & SEL_binMAT_FLRnay < B2(tS+1));
% Indices of non-FL particles belonging to the size bin.

                        binMAT_mult = sens_vols_max/...
                            sens_vols(u);          % This is the
analysed volume multiplier.

                        binMAT(tR,tS) = binMAT(tR,tS) + ...
                         (sum(NUMbSEL(indSIZE))*binMAT_mult);  %
Modified number of elements inside the bin. Adds number of elements from
previous sensitivity settings to itself.

                        binMAT_FLRaye(tR,tS) = ...
                               binMAT_FLRaye(tR,tS) + ...
                               (sum(NUMbSEL_FLRaye(indSIZE_FLRaye))*...
 binMAT_mult);
% Modified number of FL elements inside the bin. Adds number of elements
from previous sensitivity settings to itself.

                        binMAT_FLRnay(tR,tS) = ...
                               binMAT_FLRnay(tR,tS) + ...
                               (sum(NUMbSEL_FLRnay(indSIZE_FLRnay))*...
 binMAT_mult);
% Modified number of non-FL elements inside the bin. Adds number of
elements from previous sensitivity settings to itself.
                   end
             end
      end
end

binMAT = (binMAT.*coeffCorr)./(sens_vols_max*(1e-3));    % Final
particle concentrations for the RI and size bins, after application of
the flow rate and volume corrections. Converted to number concentration
per millilitre
binMAT_FLRaye = (binMAT_FLRaye.*coeffCorr)./(sens_vols_max*(1e-3)); %
Final FL particle concentrations for the RI and size bins, after
application of the flow rate and volume corrections. Converted to number
concentration per millilitre
binMAT_FLRnay = (binMAT_FLRnay.*coeffCorr)./(sens_vols_max*(1e-3)); %
Final non-FL particle concentrations for the RI and size bins, after
application of the flow rate and volume corrections. Converted to number
concentration per millilitre
```

```
PSD = sum(binMAT,1);           % Final total PSD
PRID = sum(binMAT,2);          % Final total PRID

FL_PSD = sum(binMAT_FLRaye,1);           % Final total FL PSD
FL_PRID = sum(binMAT_FLRaye,2);          % Final total FL PRID

nFL_PSD = sum(binMAT_FLRnay,1);           % Final total non-FL PSD
nFL_PRID = sum(binMAT_FLRnay,2);          % Final total non-FL PRID
```

C. Mie forward optical modelling: MATLAB code

The core sections of the Mie forward modelling MATLAB code which was used to calculate total particulate IOPs from FC PSDs and PRIDs are presented along with a flow chart of the code structure (Fig. C.1) in this second appendix. The function of each line is described in the comments within the code. Code sections which require data inputs from the user are highlighted in light blue, final outputs are highlighted in orange.

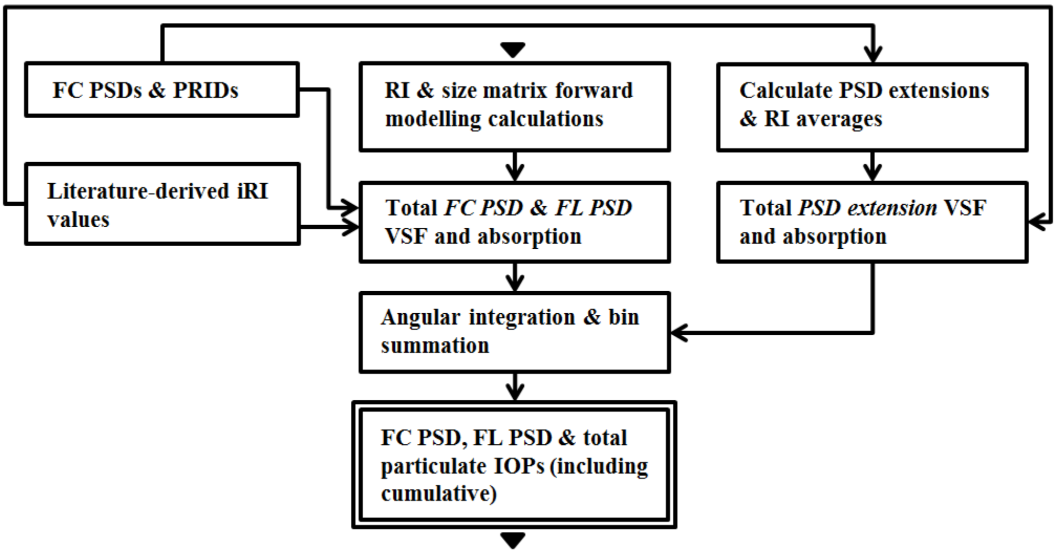


Figure C.1 – Flow chart for the IOP forward modelling procedure. PSDs and PRIDs are those determined by the flow cytometric method which described in Chapter 6; the iRIs of the particles cannot be determined by the method, and are instead adapted from literature as outlined in Chapters 4 & 7.

C.1 PSD power law best fit and PSD extrapolation

```
%%% 1.1 PSD extrapolation

%%% This section of the code defines the PSD extrapolations and
%%% calculates the slope of the size distribution. It requires the
%%% user to input the total PSD and the median particle size of each
%%% size bin (plotSIZEs). Note that plotSIZEs must already contain
%%% the bins that will correspond to the extrapolations. PSD must
%%% accordingly contain empty (0) values in correspondence with
%%% these bins.

[~,indminA] = max(PSD);     % This index defines the particle size at
which the PSD peaks before falling off due to size detection limit
[~,indmaxA] = min(abs(plotSIZEs-2e-5));     % This index corresponds
to 20 micron, the limit of statistical significance in a majority of
UKCW PSDs

%%% The next two lines will define derived indices. These further
%%% neglect two size bins on each end of the PSD to avoid boundary
%%% effects
indminAf = indminA+2;
indmaxAf = indmaxA-2;

%%% The next two lines will define a reduced PSD using the two
%%% indices just defined.
reducedPSD = PSD*0;
reducedPSD(indminAf:indmaxAf) = PSD(indminAf:indmaxAf);

%%% The next lines define the arrays over which fitting takes place.
%%% A further input from the user is required (plotBINs). This
%%% contains the width of the size bins, and it will be required to
%%% calculate the density function of the PSD. This is the one for
%%% which distribution slope is calculated.
fitmeXa = plotSIZEs(indminAf:indmaxAf);
fitmeBa = plotBINs(indminAf:indmaxAf);
fitmeYa = PSD(indminAf:indmaxAf);

%%% The next line defines the number of bins considered
binNum = numel(plotBINs);

expungeInd = find(fitmeYa == 0);     % Finds the zeroes corresponding
to the yet-to-be-calculated extrapolations in preparation for the
best fit calculations

%%% The zeroes are subsequently eliminated from the fitting arrays
fitmeXa(expungeInd) = [];
fitmeBa(expungeInd) = [];
fitmeYa(expungeInd) = [];

pFITA = coeffvalues(fit(log(fitmeXa)',log(fitmeYa)','poly1')); %
Power law best fit of the PSD. This one will be used for the PSD
extrapolations proper. Note that the fit is actually carried out as
a linear fit of the logarithm of the PSD and bin values as to limit
the excessive influence of the largest values on the fit

PDF = fitmeYa./fitmeBa;        % Density function of the PSD
```

```
pFITnA = coeffvalues(fit(log(fitmeXa)',log(PDF)','poly1'));   %
Power law best fit of the density function, similarly carried out as
a linear fit of log values

slope = -pFITnA(1);     % The slope of the PSD, given as the slope
of the underlying density function

%%% The following lines calculate the PSD extrapolations proper.

N = ((pFITA(1).*log(plotSIZEs))+pFITA(2));      % Linear
extrapolation
jungePSD = exp(N);      % Resulting power law extrapolation

extPSD = reducedPSD;    % Preallocates extended PSD array

%%% The next line preallocates the complementary PSD array - this
%%% contains ONLY the extensions. This is needed because IOP
%%% calculation for the PSD will be made using the size & RI
%%% concentration matrix rather than the single PSD. This avoids
%%% averaging and maintains full details of the organic/inorganic
%%% content of the particulate. On the other hand, the extensions
%%% are NOT a matrix, but an array (which uses rRI approximations).
%%% Calculations must be done separately and then combined.
complPSD = PSD*0;

for v = 1:1:binNum      % This loop fills in complPSD and extPSD
      if reducedPSD(v) == 0
            complPSD(v) = jungePSD(v);
            extPSD(v) = jungePSD(v);
      end
end
```

C.2 Total particulate IOPs calculations

```
%%% 2.1 FC PSD

%%% This section calculates the IOPs limitedly to the FC PSD, that
%%% is the unextended size distribution. In practice, calculations
%%% are made for the size & RI concentration matrix (PSDmat), which
%%% maintains all details about composition of the particulate and
%%% allows for organic and inorganic IOP calculations

N = PSDmat;             % The size & RI concentration matrix
N = N*1e6;              % This is a cm3 to m3 conversion

D = plotSIZEs;          % Median size of the bins
RIs = plotRIs;          % Median RI of the bins. This needs to be
input by the user, consistently with PSDmat

%%% The following contains the RI approximations in the PSD
%%% extrapolations. The lines given here are for Mode B (see Chapter
%%% 7). The code checks for the first and last non-zero elements of
%%% the PSD and uses those as starting and ending point for a 4-bin
%%% average of RI (i.e. average over first non-zero element and next
%%% three bins, and average over last non-zero element and previous
%%% three bins)
```

```
multRIs1 = repmat(RIs',1,65); % Constructs a RI matrix to
superimpose the size & RI matrix
RI_avgMe = multRIs1.*N;       % Multiplies matrices in preparation
for weighted average
avgRI = sum(RI_avgMe(:))/sum(N(:));       % Weighted average over
entire PSD

check = 0;        % Primes non-zero element locator
for nz = 1:1:size(N,2)  % Cycles through PSD matrix
      check = check+sum(N(:,nz));   % Sums over PSD matrix column
      if check ~= 0           % If “check” becomes non-zero...
            nz1 = nz;
            break             % ... break from loop
      end
end
check = 0;        % Primes non-zero element locator
for nz = 1:1:size(N,2)  % Cycles through PSD matrix
      check = check+sum(N(:,end-nz+1));   % Sums over PSD matrix
column
      if check ~= 0           % If “check” becomes non-zero...
            nz2 = size(N,2)-nz+1;
            break             % ... break from loop
      end
end

multRIs2 = repmat(RIs',1,4);        % Constructs a RI matrix to
superimpose the size & RI matrix
N1 = N(:,nz1:nz1+3);          % Isolates the section of the size &
RI matrix needed for the first averaging
N2 = N(:,nz2-3:nz2);          % Isolates the section of the size &
RI matrix needed for the second averaging
RI_avgMe1 = multRIs2.*N1;           % Multiplies matrices in
preparation for weighted average
RI_avgMe2 = multRIs2.*N2;           % Multiplies matrices in
preparation for weighted average
avgRI1 = sum(RI_avgMe1(:))/sum(N1(:));    % Weighted average
avgRI2 = sum(RI_avgMe2(:))/sum(N2(:));    % Weighted average

RV = numel(D);          % Number of elements in the diameter array
RIV = numel(RIs);       % Number of elements in the RI array

%%% 2.1.1 FC PSD Mie calculations

%%% This section will manage the Mie forward modelling calculations.
%%% The first step is to calculate the single-particle VSF and
%%% absorption corresponding to each size/RI bin of the size &
%%% refractive index matrix. Although in this appendix the
%%% corresponding code lines are given within the rest of the code,
%%% it is advisable to conduct these calculations in a separate
%%% script and to store the resulting data in a file, since these
%%% don’t change if the RI and size arrays remain the same. The
%%% resulting VSFs and absorptions can then be simply multiplied by
%%% the number of elements within each bin to determine the total
%%% particulate VSF and absorption. The angular resolution of the
%%% VSF is defined here similarly to Appendix A.

%%% Angular ranges
```

```
%%% The next few lines define the scattering angles array. The
%%% resolution is higher at small angles.

theta1 = (0:0.01:1);      % Angle range 1.
theta2 = (1.1:0.1:10);    % Angle range 2.
theta3 = (11:1:180);      % Angle range 3.

theta = [theta1 theta2 theta3];     % Scattering angles array.
thetalength = numel(theta);         % Gets the number of scattering
angles.
theta_rad = theta./180*pi;          % Converts the scattering angles
array into radians.

[~,index1] = min(abs(theta-90));    % Singles out the element in the
scattering angles array corresponding to pi/2 for integration
purposes.
[~,index2] = min(abs(theta-180));    % Singles out the element in
the scattering angles array corresponding to pi for integration
purposes.

dtheta = ones(1,thetalength);          % Preallocates the array
which contains the increments between angles in the scattering
angles array. For integration purposes.

for S1 = 1:1:(thetalength-2)           % This small loop fills the
angle increments array.
      dtheta(S1+1) = (theta_rad(S1+2)-theta_rad(S1))/2;
end;

dtheta(1) = dtheta(2);    % Defines the first element of dtheta.
dtheta(thetalength) = dtheta(thetalength-1);      % Defines the last
element of dtheta.

sintheta = sin(theta_rad);             % The sines of the scattering
angles.

r = D/2;    % Array of log-spaced virtual particle radii.

Beta_ind_wvl = cell(RIV,RV);        % Preallocates the cell matrix
which will contain all single-particle VSFs corresponding to each
size/RI bin of the size & refractive index matrix for a certain
wavelength. In this study this was chosen as 532 nm (see Section
7.3)
abs_ind_wvl = ones(RIV,RV);         % Preallocates the matrix which
will contain all single-particle absorption corresponding to each
size/RI bin of the size & refractive index matrix for the selected
wavelength.

%%% Wavelength is defined here
lambda_air = 532*1e-9; % Chosen wavelength in air (nanometres).
wnr = 1.333; % Absolute refractive index of water
lambda = lambda_air/wnr; % Chosen wavelength in water.

for RI1 = 1:1:RIV             % Iterates through RIs
      for R1 = 1:1:RV         % Iterates through radii

%%% This part will require the user to input the imaginary
```

```
%%% refractive indices for the selected wavelength. In this study
%%% these were adapted from Babin et al. (2003).
            if RIs(RI1) < 1.1
                  ni = niOrg_wvl;
            elseif RIs(RI1) >= 1.1
                ni = niMin_wvl;
            end
            m=RIs(RI1)+(1i*ni);     % Complex refractive index.
            w=2*pi/lambda;          % Wavenumber.
            x=w*r(R1);              % Size parameter.
            [S1,S2,Qb,Qc,Qbb] = fastmie(x,m,[],theta_rad);  % The
core of the calculation is handled by the FASTMie script (by W. H.
Slade, 2006).
            i1=abs(S1).^2;    % Scattered intensity functions.
            i2=abs(S2).^2;    % Scattered intensity functions.
            Beta_ind_wvl{RI1,R1}= ((1/w)^2)*0.5*((i1+i2));        %
      Single-particle VSF for diameter r(R1) and refractive index
      RIs(RI1)
            abs_ind_wvl(RI1,R1) = (pi/4)*(D(R1).^2).*(Qc-Qb);     %
      Single-particle absorption for diameter r(R1) and refractive
      index RIs(RI1)
      end
end

Beta_ind = cell(RIV,RV);      % Preallocates the cell matrix which
will contain all total particulate VSFs corresponding to each
size/RI bin of the size & refractive index matrix for the selected
wavelength.
abs_ind = ones(RIV,RV);       % Preallocates the cell matrix which
will contain all total particulate absorptions corresponding to each
size/RI bin of the size & refractive index matrix for the selected
wavelength.

for RI1 = 1:1:RIV             % Iterates through RIs
      for R1 = 1:1:RV         % Iterates through radii
            Beta_ind{RI1,R1}= Beta_ind_wvl{RI1,R1}.*N(RI1,R1);    %
      total particulate VSF for diameter r(R1) and refractive index
      RIs(RI1)
            abs_ind(RI1,R1) = abs_ind_wvl(RI1,R1).*N(RI1,R1);     %
      total particulate absorption for diameter r(R1) and refractive
      index RIs(RI1)
      end
end

Beta_tot = 0;                 % Primes the total particulate VSF
(all sizes and RIs)
for RI1 = 1:1:RIV
      for R1 = 1:1:RV
            Beta_tot = Beta_tot + Beta_ind{RI1,R1};         % Total
      VSF for this wavelength
      end
end
tsd = 2*pi*dtheta.*sintheta.*Beta_tot';   % Total particulate VSF
integral

%%% 2.1.2 FC PSD IOP calculations
%%% This section handles the calculations of the IOPs proper. The
%%% first part will calculate total FC PSDs, with cumulative,
```

```
%%% organic and inorganic IOPs following.

FCPSD_b = sum(tsd);           % FC PSD total scattering coefficient
tsda = tsd(index1:index2);    % Backscattering VSF integral
FCPSD_bb = sum(tsda);   % FC PSD total backscattering coefficient
FCPSD_bbr = (FCPSD_bb/FCPSD_b)*100; % FC PSD total backscattering
ratio
FCPSD_a_Qa = sum(sum(abs_ind));     % FC PSD total absorption
coefficient

%%% This small section calculates the FC PSD component of the
%%% cumulative IOPs.

cmlPSD1_b = zeros(RV);        % Preallocates cumulative scattering
array
cmlPSD1_bb = zeros(RV);       % Preallocates cumulative
backscattering array
cmlPSD1_a_Qa = zeros(RV);     % Preallocates cumulative absorption
array

for c = 1:1:binNum      % Iterates through size bins
      Beta_tot_cml = 0;       % Primes the total particulate VSF
      for RI1 = 1:1:RIV       % Iterates through RIs
            for R1 = 1:1:c    % Iterates through radii up to bin c
                  Beta_tot_cml = Beta_tot_cml + Beta_ind{RI1,R1};
            % Total VSF for this wavelength
            end
      end
      tsd = 2*pi*dtheta.*sintheta.*Beta_tot_cml';     % Total
particulate VSF integral
      cmlPSD1_b(c) = sum(tsd);      % FC PSD component of cumulative
scattering
      tsda = tsd(index1:index2);    % Backscattering VSF integral
      cmlPSD1_bb(c) = sum(tsda);    % FC PSD component of cumulative
backscattering
      cmlPSD1_a_Qa(c) = sum(sum(abs_ind(:,1:c),2),1); % FC PSD
component of cumulative absorption
end

%%% The following section determines organic and inorganic IOPs.

%%% This next line finds the RI element closest to 1.1.
[~,RIind] = min(abs(RIs-1.1));

Beta_tot = 0;                 % Primes the total particulate VSF
orgPSD_a_Qa = 0;              % Primes the total particulate organic
absorption
for RI1 = 1:1:RIind-1         % Iterates through organic RIs
      for R1 = 1:1:RV         % Iterates through radii
            Beta_tot = Beta_tot + Beta_ind{RI1,R1};       % Total VSF
for this wavelength
            orgPSD_a_Qa = orgPSD_a_Qa + abs_ind(RI1,R1);          %
Total organic absorption for this wavelength
      end
end
tsd = 2*pi*dtheta.*sintheta.*Beta_tot';   % Total particulate VSF
integral
orgPSD_b = sum(tsd);          % FC PSD total organic scattering
coefficient
```

```
tsda = tsd(index1:index2);     % Backscattering VSF integral
orgPSD_bb = sum(tsda);   % FC PSD total organic backscattering
coefficient
orgPSD_bbr = (orgPSD_bb/orgPSD_b)*100;          % FC PSD total
organic backscattering ratio

Beta_tot = 0;                  % Primes the total particulate VSF
minPSD_a_Qa = 0;               % Primes the total particulate
inorganic absorption
for RI1 = RIind:1:RIV          % Iterates through inorganic RIs
      for R1 = 1:1:RV          % Iterates through radii
            Beta_tot = Beta_tot + Beta_ind{RI1,R1};      % Total VSF
for this wavelength
            minPSD_a_Qa = minPSD_a_Qa + abs_ind(RI1,R1);          %
Total inorganic absorption for this wavelength
      end
end
tsd = 2*pi*dtheta.*sintheta.*Beta_tot';   % Total particulate VSF
integral
minPSD_b = sum(tsd);           % FC PSD total inorganic scattering
coefficient
tsda = tsd(index1:index2);     % Backscattering VSF integral
minPSD_bb = sum(tsda);   % FC PSD total inorganic backscattering
coefficient
minPSD_bbr = (minPSD_bb/ minPSD_b)*100;         % FC PSD total
inorganic backscattering ratio

%%% 2.2 Extended PSD

%%% This section calculates the IOPs of the extended PSD by
%%% integrating the results of the FC PSD sections with IOP
%%% calculations for the sole extensions to the PSD. The first step
%%% calculates total particulate IOPs, followed by total cumulative
%%% IOPs and IOPs corresponding to the upper and lower extensions of
%%% the PSD.

N = complPSD;      % The complementary PSD array
N = N*1e6;         % This is a cm3 to m3 conversion

extRIs = D*0;            % Primes RI array in the extensions

%%% The next loop assigns the average RIs to the extension as they
%%% have been calculated earlier
for p = 1:1:RV     % Iterates through radii
      if plotSIZEs(p) < 1e-6        % First average assigned to
particles smaller than 1 micron
            extRIs(p) = avgRI1;
      elseif plotSIZEs(p) > 1e-5    % Second average assigned to
particles larger than 10 microns
            extRIs(p) = avgRI2;
      else                          % Overall average assigned to
particles between 1 and 10 microns
            extRIs(p) = avgRI;
      end
%%% Imaginary RIs are assigned similarly as before.
      if extRIs(R1) < 1.1
            ni = niOrg_wvl;
      elseif extRIs(R1) >= 1.1
```

```
            ni = niMin_wvl;
      end
end

Beta_ind = ones(RV,thetalength);     % Primes total VSF
abs_ind = ones(RV,1);                % Primes total absorption

for R1 = 1:1:RV          % Iterates through radii
      m=extRIs(R1)+(1i*ni);    % Complex refractive index
      w=2*pi/lambda;           % Wavenumber
      x=w*r(R1);               % Size Parameter
      [S1,S2,Qb,Qc,Qbb] = fastmie(x,m,[],theta_rad);           % The
core of the calculation is handled by the FASTMie script (by W. H.
Slade, 2006).
      i1=abs(S1).^2;           % Scattered intensity functions.
      i2=abs(S2).^2;           % Scattered intensity functions.

      Beta_ind_core = ((1/w)^2)*0.5*((i1+i2));  % Core VSF
calculation for r(R1)
      abs_ind_core = (pi/4)*(D(R1).^2).*(Qc-Qb);      % Core
absorption calculation for r(R1)

      Beta_ind(R1,:)= Beta_ind_core.*N(R1);     % Total particulate
VSF for r(R1)
      abs_ind(R1) = abs_ind_core.*N(R1);        % Total particulate
absorption for r(R1)
end

Beta_tot = sum(Beta_ind,1);        % Total VSF for the selected
wavelength
tsd = 2*pi*dtheta.*sintheta.*Beta_tot;     % Total particulate VSF
integral
complPSD_b = sum(tsd);   % Total scattering coefficient in the
extensions
tsda = tsd(index1:index2);     % Backscattering VSF integral
complPSD_bb = sum(tsda);       % Total backscattering coefficient in
the extensions
complPSD_a_Qa = sum(abs_ind); % Total absorption coefficient in the
extensions

extPSD_b = FCPSD_b+complPSD_b;       % Total particulate scattering
coefficient
extPSD_bb = FCPSD_bb+complPSD_bb;    % Total particulate
backscattering coefficient
extPSD_bbr = (extPSD_bb/extPSD_b)*100;           % Total particulate
scattering ratio
extPSD_a_Qa = FCPSD_a_Qa+complPSD_a_Qa;   % Total particulate
absorption coefficient

%%% This small section calculates the extension component of the
%%% cumulative IOPs.

cmlPSD2_b = zeros(RV);         % Preallocates cumulative scattering
array
cmlPSD2_bb = zeros(RV);        % Preallocates cumulative
backscattering array
cmlPSD2_a_Qa = zeros(RV);      % Preallocates cumulative absorption
array
```

```
for c = 1:1:binNum      % Iterates through size bins
      Beta_tot_cml = sum(Beta_ind(1:c,:),1);      % Total VSF for
this wavelength
      tsd = 2*pi*dtheta.*sintheta.*Beta_tot_cml;      % Total
particulate VSF integral
      cmlPSD2_b(c) = sum(tsd);      % Extension component of
cumulative scattering
      tsda = tsd(index1:index2);    % Backscattering VSF integral
      cmlPSD2_bb(c) = sum(tsda);    % Extension component of
cumulative backscattering
      cmlPSD2_a_Qa(c) = sum(abs_ind(c));  % Extension component of
cumulative absorption
end

cmlPSD_b = cmlPSD1_b+cmlPSD2_b;      % Total cumulative particulate
scattering coefficient
cmlPSD_bb = cmlPSD1_bb+cmlPSD2_bb;  % Total cumulative particulate
backscattering coefficient
cmlPSD_a_Qa = cmlPSD1_a_Qa+cmlPSD2_a_Qa;  % Total cumulative
particulate absorption coefficient

%%% The following section determines IOPs in the upper and lower
%%% extensions

Beta_tot_upx = sum(Beta_ind(1:nz1-1,:),1);    % Total upper
extension VSF for the selected wavelength
tsd = 2*pi*dtheta.*sintheta.*Beta_tot_upx;      % Total particulate
VSF integral
upxPSD_b = sum(tsd);          % Total upper extension particulate
scattering coefficient
tsda = tsd(index1:index2);    % Backscattering VSF integral
upxPSD_bb = sum(tsda);  % Total upper extension particulate
backscattering coefficient
upxPSD_bbr = (upxPSD_bb/upxPSD_b)*100;    % Total upper extension
particulate backscattering ratio
upxPSD_a_Qa = sum(abs_ind(1:nz1-1));      % Total upper extension
particulate absorption coefficient

Beta_tot_lox = sum(Beta_ind(nz2+1:end,:),1);    % Total lower
extension VSF for the selected wavelength
tsd = 2*pi*dtheta.*sintheta.*Beta_tot_lox;      % Total particulate
VSF integral
loxPSD_b = sum(tsd);          % Total lower extension particulate
scattering coefficient
tsda = tsd(index1:index2);    % Backscattering VSF integral
loxPSD_bb = sum(tsda);  % Total lower extension particulate
backscattering coefficient
loxPSD_bbr = (loxPSD_bb/loxPSD_b)*100;    % Total lower extension
particulate backscattering ratio
loxPSD_a_Qa = sum(abs_ind(nz2+1:end)); % Total lower extension
particulate absorption coefficient

%%% 2.3 Fluorescing PSD

%%% This section repeats the calculations made in section 2.1 but
%%% this time only on the fluorescent fraction of the FC PSD.
```

```
N = FLPSDmat;                   % The size & RI concentration matrix
N = N*1e6;                % This is a cm3 to m3 conversion

Beta_ind = cell(RIV,RV);        % Re-preallocates the cell matrix
which will contain total fluorescent particulate VSFs
abs_ind = ones(RIV,RV);         % Re-preallocates the matrix which
will contain total fluorescent particulate absorptions

for RI1 = 1:1:RIV               % Iterates through RIs
      for R1 = 1:1:RV           % Iterates through radii
            Beta_ind{RI1,R1}= Beta_ind_wvl{RI1,R1}.*N(RI1,R1);    %
      total particulate VSF for diameter r(R1) and refractive index
      RIs(RI1)
            abs_ind(RI1,R1) = abs_ind_wvl(RI1,R1).*N(RI1,R1);     %
      total particulate absorption for diameter r(R1) and refractive
      index RIs(RI1)
      end
end

Beta_tot = 0;                   % Primes the total particulate VSF
FLPSD_a_Qa = 0;                 % Primes the total particulate
fluorescent absorption
for RI1 = 1:1:RIind-1           % Iterates through organic RIs
      for R1 = 1:1:RV           % Iterates through radii
            Beta_tot = Beta_tot + Beta_ind{RI1,R1};       % Total VSF
for this wavelength
            FLPSD_a_Qa = FLPSD_a_Qa + abs_ind(RI1,R1);            %
Total fluorescent absorption for this wavelength
      end
end
tsd = 2*pi*dtheta.*sintheta.*Beta_tot';    % Total particulate VSF
integral
FLPSD_b = sum(tsd);             % FC PSD total fluorescent scattering
coefficient
tsda = tsd(index1:index2);      % Backscattering VSF integral
FLPSD_bb = sum(tsda);     % FC PSD total fluorescent backscattering
coefficient
FLPSD_bbr = (FLPSD_bb/FLPSD_b)*100;        % FC PSD total fluorescent
backscattering ratio
```