Characterization and correction of errors in measured inherent optical properties: Multiple scattering errors in LISST-VSF measurements

Håvard Stavn Ugulen

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2023

UNIVERSITY OF BERGEN

Characterization and correction of errors in measured inherent optical properties: Multiple scattering errors in LISST-VSF measurements

Håvard Stavn Ugulen

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 28.11.2023

Year: 2023

Title: Characterization and correction of errors in measured inherent optical properties: Multiple scattering errors in LISST-VSF measurements

Name: Håvard Stavn Ugulen

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

This study has been carried out at the Department of Physics and Technology, University of Bergen.

Acknowledgements

First of all, I would like to thank my supervisors Camilla Sætre, Børge Hamre, and Arne Kristoffersen for the guidance and support throughout this thesis work. I would also like direct a big thanks to the optics group at the University of Bergen, not only for contributing to my work through discussions and cooperation, but also for creating a very pleasant working and social environment. I must also acknowledge the work of Håkon Sandven and Daniel Koestner, who have contributed with measurements from field work. Field work is hard work (or so I've been told), so I am very grateful for their efforts. Finally, I would like to thank my family and friends for the support during this time as a PhD candidate and throughout my time at the University of Bergen.

Håvard Stavn Ugulen
Bergen, June 19th 2023

Abstract

Accurate measurements of light-particle interactions in marine environments have become increasingly important over the years due to the growing interest in environmental monitoring, remote sensing, and underwater optical wireless communication. Interactions in terms of absorption and scattering are described by the medium's inherent optical properties (IOPs), which are properties that are solely dependent on the medium itself. Thus, obtaining knowledge of the IOPs for a body of water provides valuable information of what type of particles constitutes the water, e.g. phytoplankton and minerals, and the size distribution of these particles.

The volume scattering function (VSF) is one of the fundamental IOPs, describing the angular distribution and relative intensity of scattered light. Measurements of the VSF have for many years been difficult to obtain, requiring complex instrumentation, while being practical to deploy for *in situ* measurements. The LISST-VSF is a commercial instrument developed by Sequoia Scientific, and is able to measure the VSF from $0.1 - 150^{\circ}$. The availability of a commercial instrument enables *in situ* measurements of VSFs on a larger scale, and offers a more standardized method for doing so. In addition to being a characterizing property of the medium, the VSF is also commonly used in light transport simulations. Analytical models of the VSF have traditionally been employed for this purpose, but these do generally not accurately model the VSF in natural waters.

This thesis work revolves around characterization and correction of VSF measurements. The methods presented are general in nature, but are here applied to the LISST-VSF specifically. The LISST-VSF operates under the assumption that detected light is only scattered once (single scattering). However, studies have shown that this is not the case. In turbid waters, a significant fraction of the detected light is scattered two or more times (multiple scattered light). This results in elevated and distorted measured VSFs compared to the desired single scattering VSF.

In Paper I, a Monte Carlo simulation was developed to simulate the LISST-VSF instrument. The simulation was validated by comparing simulated VSFs to controlled LISST-VSF measurements in the lab. The results confirmed that discrepancies seen in LISST-VSF measurements in turbid waters were indeed due to multiple scattering. The results also demonstrated that errors from multiple scattering can be 10% at a scattering coefficient of only $b = 1\,\mathrm{m}^{-1}$.

In Paper II, the Monte Carlo simulation was improved by implementing a number of variance reduction methods. The most important of these is known as detector directional importance sampling, which significantly increases the number of photons detected in the simulation. The results demonstrated a 10x increase in efficiency at scattering coefficient $b < 6\,\mathrm{m}^{-1}$. In the second part of the study, we further investigated the

effects of multiple scattering on LISST-VSF measurements. By simulating VSFs with many different Fournier-Forand and Henyey-Greenstein phase functions, we showed that the multiple scattering error significantly depends on the phase function.

In Paper III, a feedforward neural network with three hidden layers was developed to correct multiple scattering errors in LISST-VSF measurements. The neural network was trained on a training set generated with the Monte Carlo simulation developed in Paper I and II. Testing of the neural network was carried out on two different sample sets, where dilution corrected VSFs had been obtained, serving as benchmarks for the neural network. The results showed that the neural network estimated VSFs was very similar to the dilution corrected VSFs, both in terms of shape and magnitude. The neural network was also found to provide good corrections to a large sample set consisting of *in situ* samples collected from many different places.

Sammendrag

Viktigheten av nøyaktige målinger av interaksjoner mellom lys og partikler i marine miljøer har økt de siste årene grunnet den økende interessen for miljøovervåkning, satellittmålinger, og trådløs optisk undervannskommunikasjon. Interaksjoner i form av absorbsjon og spredning er beskrevet av mediets iboende optiske egenskaper, som er egenskaper som er kun avhengige av selve mediet. Kunnskap om de iboende optiske egenskapene til vannmasse kan derfor gi verdifull informasjon om hvilke typer partikler vannet inneholder, f.eks. planteplankton og mineraler, og størrelsesfordelingen av disse partiklene.

Volumspredningsfunksjonen (VSF) er en av de fundamentale iboende optiske egenskapene, og beskriver vinkelfordelingen og den relative intensiteten av spredt lys. Det har i mange år vært vanskelig å måle VSF fordi det krever komplisert instrumentering, men som også skal være praktisk å bruke i feltarbeid. LISST-VSF er et kommersielt instrument utviklet av Sequoia Scientific, som kan gjøre målinger av VSF fra $0.1 - 150^\circ$. Et kommersielt tilgjengelig instrument gjør det mulig å gjennomføre *in situ* målinger av VSF i stor skala, og tilbyr en mer standardisert metode for å gjøre slike målinger. I tillegg til å være karakteriserende egenskap av mediet, så brukes VSF ofte i lystransportsimuleringer. Analytiske modeller for VSF har tradisjonelt blitt brukt for slike simuleringer, men disse er generelt ikke en presis modell for VSF i naturlige vannmasser.

Arbeidet i denne avhandlingen dreier seg om karakterisering og korreksjon av VSF-målinger. Metodene som presenteres er generelle, men anvendes spesifikt til LISST-VSF målinger i dette arbeidet. Instrumentet LISST-VSF opererer under antagelsen av at detektert lys kun er spredt en gang (singel spredning). Tidligere studier har vist at dette ikke stemmer, hvor en stor andel av det detekterte lyset kan være spredt to eller flere ganger (multippel spredning) i vann med høy turbiditet. Dette resulterer i for høye verdier og endret form på VSF-målingen sammenlignet med den ønskede målingen, hvor kun lys som er spredt en gang er detektert.

I den første artikkelen (Paper I) ble en Monte Carlo simulering utviklet for å simulere LISST-VSF instrumentet. Simuleringen ble validert ved å sammenligne simulert VSF med kontrollerte målinger av VSF gjennomført på laboratoriet. Resultatene bekreftet at avvikene observert i målinger med LISST-VSF i turbide vannmasser var et resultat av multippel spredning av lys. Resultatene demonstrerte også at feil i VSF målinger kan være 10% for en spredningskoeffisient på bare $b = 1\,\mathrm{m}^{-1}$.

I den andre artikkelen (Paper II) ble Monte Carlo simuleringen forbedret ved å implementere en rekke variansreduksjonsmetoder. Den viktigste av disse er kalt detector directional importance sampling, som øker andelen av fotoner som blir detektert i simuleringen. Resultatene viste 10x økt effektivitet for spredningskoeffisienter $b < 6\,\mathrm{m}^{-1}$. I

den andre delen av denne studien undersøkte vi nærmere effekten av multippel spredning på VSF målt med LISST-VSF. Ved å simulere VSF med mange forskjellige Fournier-Forand og Henyey-Greenstein fasefunksjoner fant vi at feil i målinger pga. multippel spredning avhenger av fasefunksjonen.

I den tredje artikkelen (Paper III) ble et nevralt nettverk med tre skjulte lag utviklet for å korrigere feil grunnet multippel spredning i målinger gjort med LISST-VSF. The nevrale nettverket ble trent med data generert med Monte Carlo simuleringen utviklet i den første og andre artikkelen. Testing av det nevrale nettverket ble gjennomført med to forskjellige datasett, hvor "sann" VSF ble estimert ved hjelp av utvannede prøver. Disse ble brukt som en referanse for VSF korrigert med det nevrale nettverket. Resultatene viste at korrigert VSF var veldig lik referanse VSF i både størrelse og form for begge datasett. Det nevrale nettverket ga også gode korreksjoner for et stort datasett med *in situ* VSF målinger fra mange forskjellige steder.

List of papers

This thesis consists of an introductory part and three scientific papers.

1. Ugulen, H. S., Sandven, H., Hamre, B., Kristoffersen, A. S., and Sætre, C., (2021) *Analysis of multiple scattering errors in LISST-VSF volume scattering function measurements using Monte Carlo simulations and experimental data*, Optics Express **29(8)**

2. Ugulen, H. S., Sandven, H., Hamre, B., Kristoffersen, A. S., and Sætre, C., (2022) *Efficient Monte Carlo simulation reveals significant multiple scattering errors in underwater angular scattering measurements*, Optics Express **30(7)**

3. Ugulen, H. S., Koestner, D., Sandven, H., Hamre, B., Kristoffersen, A. S., and Sætre, C., (2023) *A neural network approach for correction of multiple scattering errors in the LISST-VSF instrument*, Optics Express (Submitted, May 2023)

Contents

Chapter 1

Introduction

The ocean covers 71% of the earths surface and constitutes a significant part of global natural resources (*Costanza*, 1999). It also serves as a crucial part of the earths climate, as part of the water cycle, and acting as massive reservoir for heat and carbon. Since the industrial revolution, the concentration of carbon dioxide in the atmosphere has increased continually, resulting in a stronger greenhouse effect and warmer climate. Similar trends are observed for the ocean, with increasing average surface temperature (*Hansen et al.*, 2010) and carbon concentration (*Caldeira and Wickett*, 2003). The rapid changes observed in the state of the worlds oceans may have devastating consequences, both for the organisms that inhabit it, and for those that exploit its resources.

Despite its economical and ecological importance, the worlds oceans have been underexplored (*Levin et al.*, 2019). Historically, measurements of ocean properties are made by lowering measuring equipment into water from boats and ships (*in situ* measurements). An early example is the Secchi disk, popularized by the Italian physicist Angelo Secchi, who made several measurements in 1866 (*Preisendorfer*, 1986). The Secchi disk is a white disk that can be lowered into the water until it is no longer visible, known as the Secchi depth, providing a visual measure of the clarity of the water. In modern times, new technologies have made gathering of ocean data both more efficient and accurate. Especially, remote sensing through satellite imaging has become very important, being able to cover large areas. Another example is the very successful Argo program, which utilizes autonomous profiling floats to measure temperature, salinity, currents, and certain bio-optical properties (*Jayne et al.*, 2017). The Argo program provides freely available data collected from approximately 3800 (per 2017) floats deployed all around the worlds oceans.

While large scale methods, such as satellite imaging, can provide useful data for large areas, there is still a need for *in situ* measurements. Remote sensing through satellite imaging relies on sunlight reflected and scattered from the surface layer of the ocean. Hence, this method measures apparent optical properties (AOPs), which are properties that do not only depend on the constituents of the water, but also the incident light field. On the other hand, the inherent optical properties (IOPs) of a medium describe the scattering and absorption of light that is only dependent on the medium itself. Hence, these properties can be used to describe the concentration and composition of particles suspended in the water mass. A significant effort is directed to deriving IOPs from AOPs measured by remote sensing, for which *in situ* measurements of IOPs are an integral part (*Werdell et al.*, 2018).

The volume scattering function (VSF) is one of the fundamental IOPs, describing the angular dependent light intensity scattered by a medium. Measurements of the full VSF demand detection of scattered light over the angular range $0 - 180°$, requiring complex instrumentation. Hence, *in situ* measurements of the VSF of natural waters have historically been few. The LISST-VSF (Sequoia Scientific) is a commercially developed instrument, designed to perform *in situ* measurements of the VSF over the angular range $0.1 - 150°$, and offers a standardized method for measuring the VSF of natural waters. However, all instruments suffer from errors and uncertainties. Thus, work is being done to validate the instrument, exploring its limits, and investigate sources of error, producing more reliable and accurate measurements.

In this thesis, I have developed methods for characterization and correction of IOPs measured with optical instruments. This work focuses specifically on multiple scattering errors observed in LISST-VSF measurements in turbid waters, but the methods are general in nature. First, a Monte Carlo simulation was developed to simulate LISST-VSF measurements, and was validated through comparison with bench-top measurements of polymer beads. The simulation was then used to establish that multiple scattering is a significant source of error in LISST-VSF measurements in turbid waters. Second, the Monte Carlo simulation was improved by implementing several variance reduction methods. The improved simulation was then used to quantify the multiple scattering error dependency on scattering coefficient and phase function. Finally, a neural network was developed to correct LISST-VSF measurements for multiple scattering errors.

Outline

This thesis begins with an introduction to the scientific background in Chapter 2, presenting key concepts for the methods used in the research presented. Chapter 3 provides a detailed description of the methods used. Chapter 4 is dedicated to multiple scattering, which is the focus of this thesis work, giving a thorough description of multiple scattering and how it affects LISST-VSF measurements. Then, an introduction to each publication is given in Chapter 5. Finally, this thesis scientific results, consisting of three publications, are presented in Chapter 6.

Chapter 2

Scientific background

2.1 Inherent optical properties

The inherent optical properties (IOPs) describe scattering and absorption of light by particulate and dissolved matter in water. IOPs are properties of the medium itself, meaning that they do not depend on the ambient light. Thus, measurements of IOPs of natural waters yield information about the constituents of the water. While the absorption coefficient is an important IOP in natural waters, it is not considered relevant for this thesis work and is thus not described here.

2.1.1 Scattering

Scattering of unpolarized incident light is fully described by the volume scattering function (VSF). The VSF, denoted $\beta(\theta)$, is defined as the radiant intensity, dI, scattered per elemental volume, dV, in the direction θ per unit incident irradiance E. Assuming that particles are randomly oriented in the water, one may also assume azimuthal symmetry so that the VSF for wavelength λ can be expressed independent of the azimuth angle ϕ:

$$\beta(\theta,\lambda) = \frac{dI}{EdV} \quad [\mathrm{m}^{-1}\mathrm{sr}^{-1}] \tag{2.1}$$

This function describes the angular distribution and magnitude of the scattered light. The VSF can be divided into two parts, where the total amount of scattering (i.e. magnitude of the VSF) is described by the scattering coefficient:

$$b(\lambda) = 2\pi \int_0^\pi \beta(\theta,\lambda)\sin\theta d\theta \quad [\mathrm{m}^{-1}]. \tag{2.2}$$

The angular distribution is described by the phase function, which can be calculated as the ratio

$$\tilde{\beta}(\theta,\lambda) = \frac{\beta(\theta,\lambda)}{b(\lambda)} \quad [\mathrm{sr}^{-1}]. \tag{2.3}$$

In some cases one is also interested in the backscattering coefficient,

$$b_b(\lambda) = 2\pi \int_{\pi/2}^{\pi} \beta(\theta,\lambda)\sin\theta d\theta \quad [\mathrm{m}^{-1}], \tag{2.4}$$

which describes the amount of light that is scattered backwards. This quantity is especially important when looking at the reflectance from a body of water, such as in remote sensing applications. The backscattering ratio $\tilde{b}_b = b_b/b$ is also frequently used to characterize the shape of the phase function.

Phase function

The phase function is an important property for several reasons. It is a fundamental characteristic that is dependent on the particles suspended in the water. Particles that are large compared to the wavelength of the incident light generally scatter more in the forward direction, and as the particles become smaller, scattering becomes more isotropic (*Slade and Boss*, 2015). This relation between particle size and phase function is frequently used in inversion algorithms to determine the particle size distribution in a water sample.

The phase function (and VSF) has historically been difficult to measure, as it requires scattering measurements over the angular range $0-180^\circ$. Hence, several analytical phase functions have been proposed over the years, in an attempt to approximate the phase function of natural particles, both in water and in the atmosphere. The Henyey-Greenstein (HG) phase function, proposed by *Henyey and Greenstein* (1941) for scattering by interstellar dust, is a well known phase function that has been adopted to various fields, such as biomedical optics (*Binzoni et al.*, 2006) and oceanography (*Haltrin*, 2002). The HG phase function can be expressed as

$$\tilde{\beta}_{\mathrm{HG}}(\theta) = \frac{1}{4\pi}\frac{1-g^2}{(1+g^2-2g\cos\theta)^{3/2}}, \tag{2.5}$$

where $-1 < g < 1$ is an asymmetry parameter defined as the average of the cosine of the scattering angle. Setting $g = 0$ corresponds to isotropic scattering, while $g > 0$ and $g < 0$ correspond to forwards and backwards peaked scattering, respectively. This is known as a single-parameter phase function, as the shape is only dependent on the parameter g. Generally, the HG phase function is not a good approximation for particles suspended in natural waters. A two parameter phase function, commonly known as the Fournier-Forand (FF) phase function, was proposed by *Fournier and Forand* (1994). It is the phase function of Mie scatterers integrated over an inverse power-law particle size distribution, where Mie scattering (or Mie theory) is an exact solution to scattering of electromagnetic radiation from a spherical dielectric particle of arbitrary size (*Jonasz and Fournier*, 2011; *Mie*, 1908). Hence, the FF phase function is grounded in physics, and is calculated as

$$\begin{aligned}\tilde{\beta}_{\mathrm{FF}}(\theta) = {} & \frac{\left(v(1-\delta)-(1-\delta^v)+\frac{4}{u^2}[\delta(1-\delta^v)-v(1-\delta)]\right)}{4\pi(1-\delta)^2\delta^v} \\ & -\frac{1}{16\pi}\frac{1-\delta_\pi^v}{(1-\delta_\pi)\delta_\pi^v}(3\cos^2\theta-1),\end{aligned} \tag{2.6}$$

where

$$v = \frac{3-m}{2}, \quad \delta = \frac{4}{3}\Big(\frac{\sin(\theta/2)}{n-1}\Big)^2, \quad \delta_\pi = \delta(\theta = \pi).$$

Here, $1 < n < 1.35$ is the refractive index of the particles relative to seawater, and $3.5 < m < 5$ is the slope of the inverse power-law particle size distribution. While the FF phase function is considerably more complicated than the HG phase function, it has the advantage of being defined by the physical quantities n and m. Furthermore, an additional parameter allows for more flexibility in fitting procedures. This has since been taken a step further with the development of two-term phase functions. These are made from a superposition of two phase function, adding additional parameters that can be tweaked. The two-term Reynolds-McCormick (TTRM) phase function is an example of this, based on the phase function proposed by *Reynolds and McCormick* (1980). This is a generalized version of the HG phase function, and is expressed as

$$\tilde{\beta}_{\mathrm{RM}}(\theta) = \frac{\alpha g(1-g^2)^{2\alpha}}{\pi\Big((1+g)^{2\alpha} - (1-g)^{2\alpha}\Big)(1+g^2-2g\cos\theta)^{\alpha+1}}. \tag{2.7}$$

Here, α is a fitting parameter, where $\alpha = 0.5$ results in the HG phase function. The TTRM phase function is then expressed as (*Harmel et al.*, 2021)

$$\tilde{\beta}_{\mathrm{TTRM}}(\theta) = \gamma\tilde{\beta}_{\mathrm{RM}}(\theta, g_1, \alpha_1) + (1-\gamma)\tilde{\beta}_{\mathrm{RM}}(\theta, g_2, \alpha_2), \tag{2.8}$$

where $0 < \gamma < 1$ is a weighting factor. This phase function takes five parameter inputs, allowing for greater flexibility and fine-tuning of the shape. Usually, the weighting parameter γ is large, so that there is a dominant term in the phase function. The second term is then used to tailor the backward scattering part of the angular spectrum, where setting $g_2 < 0$ allows the phase function to rise as θ approaches 180°. This approach for making two-term phase functions may be applied to other phase functions as well, such as the FF phase function. A comparison between the HG, FF, and TTRM phase function is shown in Fig. 2.1.

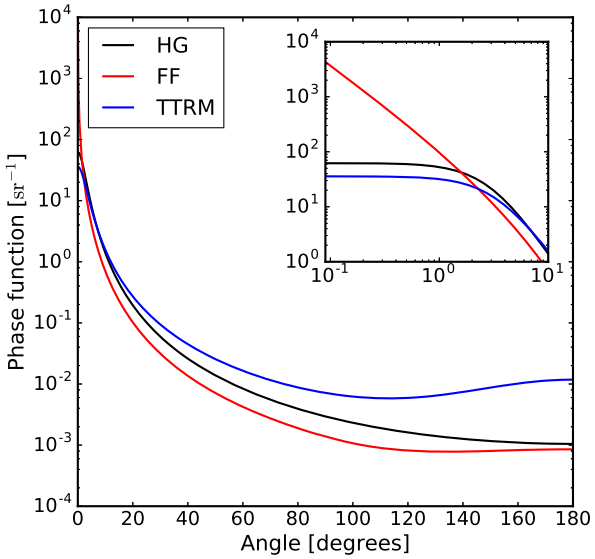

Figure 2.1: Semi-log plot of a Henyey-Greenstein (black), Fournier-Forand (red), and two-term Reynolds-McCormick (blue) phase function. A log-log plot for the angular range $0.09 - 10^{\circ}$ *is shown in the inset.*

2.2 Radiative transfer computation

Radiative transfer (or radiation transport) is the propagation of electromagnetic waves through a medium. It is mathematically expressed by the radiative transfer equation (RTE), which describes absorption, scattering, and emissions in the medium. Solutions to the RTE are usually found numerically, where approximations are in most cases necessary. Different approaches to numerical modeling of radiative transfer have been developed over the years. One of the most influential methods is the two-stream approximation, where radiative transfer is considered in only two discrete directions (usually up and down) (*Chandrasekhar*, 2013; *Stamnes et al.*, 2017). Some models make use of more streams, i.e., solve the RTE in more directions, to achieve higher accuracy. However, this comes at the cost of longer computation time.

2.2.1 The Monte Carlo method

Another approach for simulating radiative transfer is the Monte Carlo method. As opposed to the RTE, the Monte Carlo method can in principle be applied to any geometry and without approximations. The Monte Carlo method is a computer algorithm used to solve problems that can be interpreted probabilistic, and is based on repeated random sampling from probability density functions to obtain numerical results. The first real use case of the modern Monte Carlo method was applied to neutron transport during the development of nuclear weapons at Los Alamos National Laboratory (*Metropolis*

and Ulam, 1949), and has since been applied to several areas related to physics, such as medicine (*Rogers*, 2006), chemistry (*Hammond et al.*, 1994), nuclear physics, and quantum mechanics (*Carlson et al.*, 2015).

A simple example to demonstrate the Monte Carlo method is by using it to estimate π. First, we can draw a square with sides $l = 1$, and then a quarter of circle with radius $r = 1$ inside the square. The ratio of the two areas can then be calculated as

$$R = \frac{A_{\text{circle}}}{A_{\text{square}}} = \frac{\frac{\pi r^2}{4}}{l^2} = \frac{\pi}{4}. \tag{2.9}$$

Now, lets say we do not know the value of π, but we want to estimate it. From Eq. (2.9), we see that π can be expressed as $\pi = 4R$. Hence, by estimating the ratio of two areas, we simultaneously get an estimate of π. This is were we apply the Monte Carlo method. Lets pick a random point inside the square defined above, where all points inside the square have an equal probability of being picked. Now, note that the probability of this random point being inside of the quarter circle is the same as the ratio of the two areas. Hence, by picking many random points, one can estimate the ratio (and π) by simply counting how many points are within the quarter circle and divide it by the total number of points. The results from a simulation of this scenario is shown in Fig. 2.2.

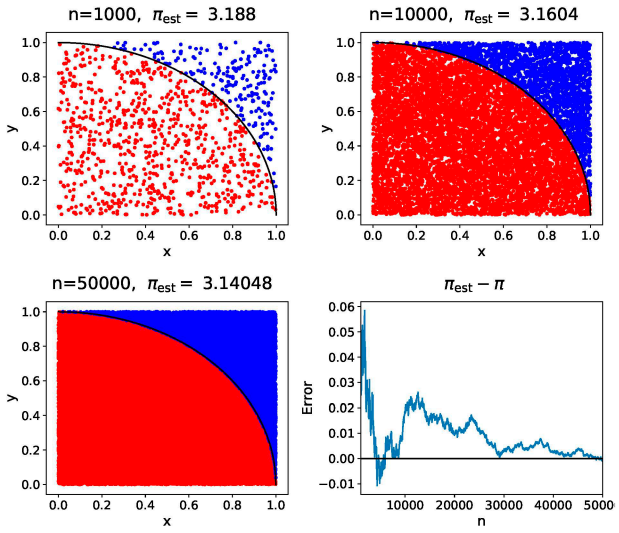

Figure 2.2: Estimation of π from sampling random points inside a square for 1000 points (top left), 10000 points (top right), and 50000 points (bottom left). Points that are within a quarter of a circle with radius $r = 1$ are colored red, and points outside are colored blue. The error $\pi_{est-\pi}$ between the estimate and true value is plotted in the bottom right panel from n=1000-50000.

The results presented in Fig. 2.2 illustrate a key aspect of Monte Carlo simulations. The estimate of π has a high variance when the number of sampled points are few,

meaning that the estimate tends to fluctuate. As more points are sampled and added to the total, the variance becomes smaller and the estimate converges towards the correct answer. This illustrates both the strength and weakness with the Monte Carlo method. At one hand, one can obtain arbitrary high accuracy by simply increasing the number of sampled points. At the other hand, accurate simulation may require a very large number of sample points, resulting in a long simulation time.

The continuing rise of computational power has made Monte Carlo modelling of radiative transfer increasingly attractive due to its accuracy. It has evolved over the years to become a competitive method with great success within fields such as astrophysics (*Noebauer and Sim*, 2019). It is also frequently used for radiative transfer simulations in water, such as for simulations of underwater optical communication (*Gabriel et al.*, 2013; *Jasman and Green*, 2013). The flexibility of the Monte Carlo method makes it an effective tool for simulation of instruments, were geometries are varied and in some cases complex. In a study by *Doxaran et al.* (2016), Monte Carlo simulations was used to improve the correction of the HydroScat backscattering sensor, while *McKee et al.* (2008) used Monte Carlo simulations to derive an improved scattering correction for the AC-9 instrument (WETLabs Inc.).

2.3 Machine learning

Over the past couple of decades, machine learning (ML) has gone from being a subject for the particularly interested, to be widely used within the industry and scientific community. While traditional computer software and systems work by following a set of instructions coded by the developer, ML revolves around making computer algorithms that can learn from experience. The recent rise of ML algorithms is usually attributed to increasing availability of data (often referred to as "big data"), and the ever increasing computing power that enabled larger and more complex ML algorithms, resulting in the emergence of deep learning algorithms (*Aggarwal et al.*, 2018; *Fradkov*, 2020).

One of the earliest examples of modern ML was a machine called "perceptron", developed to recognize letters in the alphabet, and created by a group headed by the psychologist Frank Rosenblatt (*Fradkov*, 2020; *Rosenblatt*, 1957, 1958, 1960). The machine was based on models for learning in humans and animals, and is seen as the prototype for artificial neural networks (ANN). After the birth in 1958, further development in ML was slow before the 21 century. Some contributions have proven valuable, such as the backpropagation learning algorithm (*Rumelhart et al.*, 1985, 1986), and the improved support vector machine (SVM) algorithm (*Cortes and Vapnik*, 1995).

From the start of the 21 century and onward, the field of ML has experienced a tremendous growth. Currently, ML techniques are integrated in many of the cutting edge technologies, such as in voice and face recognition (*Ali et al.*, 2021; *Chandolikar et al.*, 2022; *Tripathi*, 2017). Some of the more commonly known applications are the chess computer AlphaZero, self driving vehicles developed by companies such as Tesla and Waymo, and the recently developed language model ChatGPT.

2.3.1 Concepts in machine learning

The basic idea of machine learning is to create an algorithm that improves some measure of performance on solving a task, provided some type of training experience (*Jordan and Mitchell*, 2015). Over the years, several machine learning techniques have been developed, such as different models and learning algorithms.

Artificial neural networks

Artificial neural networks (ANNs) is a class of models inspired by the biological brain. Usually, an ANN consists of an input layer, a number of hidden layers, and an output layer. Each layer contains a certain number of nodes, which are sometimes referred to as artificial neurons, or just neurons. These nodes are capable of taking an input, apply some simple processing, and send out an output signal (*Svozil et al.*, 1997). The nodes in the network are connected to each other by a series of weights, where a weight describes the strength of the connection between the nodes it connects. A network with one or zero hidden layers are commonly referred to as shallow neural networks, while networks with more than one hidden layer are called deep neural networks.

A feedforward neural network (FFNN) is an ANN where the weights connecting the nodes only work in the forward direction (*Svozil et al.*, 1997). The input to the ANN, usually a vector of values, is fed to the first hidden layer in the network. Here, the output of each node is calculated as a function of the input values, which in turn is fed forward as inputs to the next layer. This process repeats going forward in the ANN, until reaching the output layer. This is illustrated in Fig. 2.3, for a FFNN with 2 hidden layers.

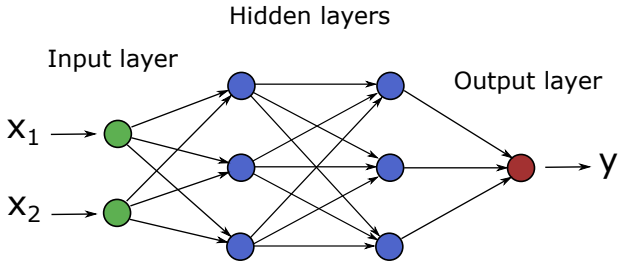

Figure 2.3: A feedforward neural network with two input values x_1 and x_2, one output value y, and two hidden layers. Connections are indicated by the arrows.

Recurrent neural networks (RNNs) is a class of ANN used to process a sequence of inputs, often involving time and movement in space. They are very similar to FFNN in terms of structure, but unlike FFNNs, the connections in RNNs are not restricted to the forward direction. Here, output from a node can be fed as input to nodes in previous layers, or even to itself. This enables the network to develop internal representation and memory properties (*Medsker and Jain*, 2001). For input ordered as a time sequence, the output from nodes at time step $t-1$ can be fed backwards as input to the network at time step t. Thus, information from previous time steps are retained in the network. This type of ANN is illustrated in Fig. 2.4, for a RNN with 2 hidden layer.

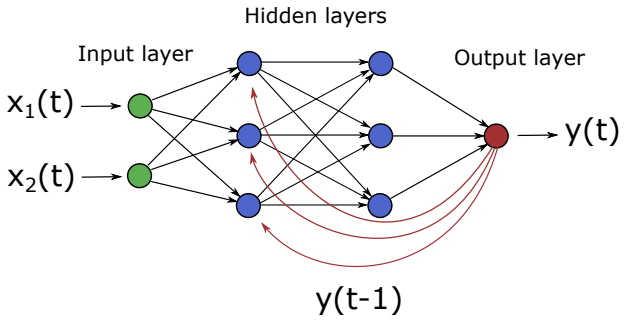

Figure 2.4: A recurrent neural network with two input values $x_1(t)$ and $x_2(t)$, one output value $y(t)$, and two hidden layers. Forward connections are indicated by straight black arrows, and backward connections are indicated by curved red arrows.

Learning paradigms

Typically, the method of learning is divided into three main categories: Supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is used for data with existing labels, meaning that the desired output is known. In this case, the aim of the training process is to find a general function that maps the input data to the desired output data (labels) (*Schmidhuber*, 2015). The deviation between the network output and the desired output can be expressed by an error function E, e.g the mean squared error. This error can then be minimized during the training process, typically by utilizing gradient descent with backpropagation, changing a weight ω_i in proportion to the derivative $\frac{\partial E}{\partial \omega_i}$.

Unsupervised learning is usually applied to data sets without labels, i.e. there is no specific desired output. One of the most common subfields in unsupervised learning is known as clustering (*Celebi and Aydin*, 2016). Here, the goal is to find patterns in the data, and group objects in clusters based on similarities. This method can also be used to discover anomalies and outliers in a data set, i.e. objects that do not fit well into any cluster.

Finally, reinforcement learning typically revolves around having an agent, e.g. a computer controlled character in a video game, learning to interact with and/or navigate an environment (*Wiering and Van Otterlo*, 2012). Usually, the feedback is simply a number, termed reward, telling the agent how well it is doing. Do to the limited feedback, this method is often based on trial and error. A typical application is in computer games, where rewards (positive actions) and punishments (negative actions) are often easily defined.

2.3.2 Machine learning in Oceanography

The recent development in machine learning has increasingly attracted the attention of researchers in the natural sciences. Machine learning techniques are now being applied within fields such as material science, ocean engineering, cosmology, and many others (*Carleo et al.*, 2019; *Juan and Valdecantos*, 2022; *Wei et al.*, 2019). The field of oceanography is no exception, where the increasing amount of available data calls for more efficient analysing and processing techniques (*Lou et al.*, 2021).

Satellites are a popular tool for gathering of ocean data due to its accessibility, temporal and spatial resolution, surface coverage, and easy access to remote locations. However, these advantages also result in a large quantity of data. Coupling satellite data with machine learning techniques has become a popular way to extract useful data from satellite images, including water quality monitoring (*Chebud et al.*, 2012; *Sharaf El Din et al.*, 2017), estimation of the surface seawater partial pressure of CO_2 (*Chen et al.*, 2019), shallow water bathymetry (*Sagawa et al.*, 2019), and more.

Chapter 3

Methods

3.1 LISST-VSF

The LISST-VSF is a complex optical instrument that is able to measure the VSF, in addition to attenuation and polarization properties. The VSF and attenuation coefficient are measured by detecting the scattered and transmitted light from a monochromatic and collimated incident laser beam. The incident laser interacts with the water sample, resulting in either scattering, absorption or transmittance. The attenuation coefficient c can be directly calculated from the transmitted intensity following Beer-Lambert law $I_t = I_0 e^{-cL}$, which in terms of power becomes

$$P_t = P_0 e^{-cL}, \tag{3.1}$$

where P_t is the transmitted power, P_0 is the incident laser power, and $L = 15\,\text{cm}$ is the pathlength of the sample chamber. Measuring the VSF is more complicated, as it requires detection of light from a large angular range. The detection system for scattered light consists of a combination of ring detectors and a rotating eyeball detector.

3.1.1 Ring detector

The ring detector system consists of 32 detectors, each covering a different part of the angular spectrum in the approximate range $0.1 - 15^\circ$. Each detector covers 1/6 of a full ring, with in an inner and outer radius r_i and r_{i+1}, respectively. Light scattered at an angle θ passes through an optical window followed by a convex lens, directing the light to the ring detectors positioned at the focal length f behind the lens, as illustrated in Fig. 3.1 (optical window is not depicted). All light incident on the lens are refracted to the radial distance $r = f \tan\theta_a$ from the center of the ring detectors. The angle in air θ_a is derived from Snell's law $\sin\theta_a = n_w \sin\theta$, where n_w is the refractive index of water. Thus, each individual ring detector detects light at angles $\theta_i < \theta < \theta_{i+1}$, corresponding to the ring's inner and outer radius r_i and r_{i+1}. The angles θ_i and θ_{i+1} are provided by the manufacturer (Sequoia Scientific).

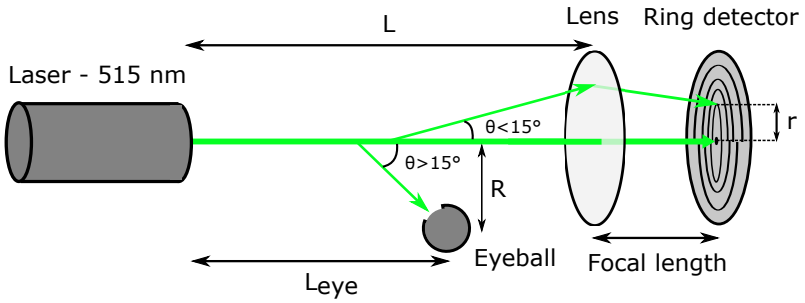

Figure 3.1: Schematics of the LISST-VSF instrument. The lens refracts all light with a certain incidence angle to a distance r from the center of the ring detector. The eyeball is positioned at a distance R from the laser beam and rotates to view the beam at different angles for $\theta > 15°$. The optical window situated in front of the convex lens is not shown.

The VSF can be calculated from the measured optical power. Starting from the definition of the VSF (Eq. 2.1), we have

$$dP_s(\theta) = \beta(\theta) E dV d\Omega, \tag{3.2}$$

where $dP_s = dI d\Omega$ is the optical power from an infinitesimal volume dV with a VSF $\beta(\theta)$, scattered into an infinitesimal solid angle $d\Omega$, due to an incident irradiance E. As the laser beam is collimated, the incident irradiance can be expressed as $E = \exp(-cx)P_0/A$, where P_0 is the laser power, and A is the cross-section of the beam. The factor $\exp(-cx)$ represents attenuation by the water sample, where c is the attenuation coefficient, and x is the pathlength from the beginning of the sample to the scattering volume dV. The beam is further attenuated traveling from the scattering volume dV to the detector, so that the measured power can be expressed as

$$dP(\theta) = dP_s(\theta) \exp\Big[-\frac{c(L-x)}{\cos\theta}\Big], \tag{3.3}$$

where $L = 15\,\text{cm}$ is total length of the sample chamber. Inserting Eq. 3.3 into Eq. 3.2, the measured power can be related to the VSF by the following expression

$$\begin{aligned} dP(\theta) &= \exp\Big[-\frac{c(L-x)}{\cos\theta}\Big]\beta(\theta)\exp(-cx)\frac{P_0}{A}A dx 2\pi\phi \sin\theta d\theta \\ &= \exp\Big[-c\Big(\frac{L-x}{\cos\theta} - x\Big)\Big]\beta(\theta)P_0 dx 2\pi\phi \sin\theta d\theta. \end{aligned} \tag{3.4}$$

Here, the scattering volume is expressed as $dV = Adx$, where dx is an infinitesimal length of the laser beam. Under the assumption of azimuthal symmetry, the solid angle is written as $d\Omega = 2\pi\phi \sin\theta d\theta$, where $\phi = 1/6$ is the fraction of a full ring covered by the detector. The total power measured by each ring is found by integrating x over the total pathlength L, and the polar angle θ from the inner scattering angle θ_i to the outer scattering angle θ_{i+1} covered by the respective detectors. The resulting integral is

$$P_i = 2\pi\phi P_0 \int_0^L \int_{\theta_i}^{\theta_{i+1}} \exp\Big[-c\Big(\frac{L-x}{\cos\theta} - x\Big)\Big]\beta(\theta) dx \sin\theta d\theta. \tag{3.5}$$

In order to calculate this integral, we assume that $\beta(\theta)$ varies slowly between the angles θ_i and θ_{i+1}, and treat it as constant representing an approximation of the average value $\beta_i =< \beta(\theta) >$ in this angular range. Additionally, we can apply the approximation $\cos\theta = 1$, since the ring detectors only measure light scattered at angles $\theta < 15°$. Thus the integral in Eq. 3.5 becomes

$$\begin{aligned} P_i &= 2\pi\phi P_0 \beta_i \exp(-cL) \int_0^L dx \int_{\theta_i}^{\theta_{i+1}} \sin\theta d\theta. \\ &= 2\pi\phi P_0 \beta_i \exp(-cL) L(\cos\theta_i - \cos\theta_{i+1}) \end{aligned} \tag{3.6}$$

To obtain the particulate VSF $\beta_{i,p}$, one must also account for scattering and absorption of pure water, as well as optical losses in the instrument. This can be done by performing background measurements on ultra-pure water, e.g. milli-Q water. As the total VSF $\beta_i = \beta_{i,w} + \beta_{i,p}$, where $\beta_{i,w}$ is the VSF of pure water, the following expression can be derived

$$P_{i,p} = 2\pi\phi P_0 \beta_{i,p} e^{-c_w L} L(\cos\theta_i - \cos\theta_{i+1}) = \frac{P_i}{\tau} - P_{i,w}. \tag{3.7}$$

Here, τ is the attenuation only due to particles, and is related to the transmission measurements by the equation

$$\tau = e^{-c_p L} = \frac{e^{-(c_p + c_w)L}}{e^{-c_w L}} = \frac{P_t}{P_{t,w}}, \tag{3.8}$$

where $P_{t,w}$ is the transmission measurement of pure water. Finally, the particulate VSF is calculated as

$$\beta_{i,p} = \frac{P_{i,p}}{P_0} \frac{e^{c_w L}}{2\pi\phi L(\cos\theta_i - \cos\theta_{i+1})} \tag{3.9}$$

3.1.2 Eyeball detector

The angular range $15 - 150°$ is covered by an eyeball detector, rotating around its own axis to detect light scattered from different parts of the laser beam (different scattering angles). Unlike the ring detector, the eyeball detector is designed to measure polarization properties in addition to the VSF. The incident light passes through a half-wave plate, resulting in either perpendicular (r) or parallel (p) polarized light. At the other end, scattered light passes through a polarizing beam splitter, separating the detected light into its perpendicular and parallel polarization components. Thus, there are four combinations of incident and detected light polarization, resulting in the four scattering measurements I_{rp} (perpendicular incident light and parallel detected light), I_{rr}, I_{pp}, and I_{pr}. While this enables measurements of the polarization properties of the sample, it also complicates the measurement processing. The processing steps are not described in detail here, but generally follow the steps described in (*Hu et al.*, 2019; *Sandven et al.*, 2020).

The uncalibrated VSF measured by the eyeball detector can be calculated from applying Mueller calculus to the four scattering measurements I_{rp}, I_{rr}, I_{pp}, and I_{pr}. For each measurement, several factors must be accounted for. The eyeball detector operates over a large dynamic range, depending on angle of detection, turbidity, and phase function. Hence, the detector uses photomultiplier tubes with automatically adjusted gain levels, and the laser is dimmed between 15° and ∼45°. Geometrical effects must also be considered, where a $\sin\theta$ factor accounts for the angular dependent volume viewed by the detector, and the factor $e^{c_p l(\theta)}$ accounts for attenuation. Here, the particulate attenuation c_p is calculated according to Eq. 3.8, and the angular dependent pathlength is calculated as

$$l(\theta) = L_{\mathrm{eye}} - R\cot\theta + \frac{R}{\sin\theta}. \tag{3.10}$$

Here, $L_{\mathrm{eye}} = 10\,\mathrm{cm}$ is the distance along the laser axis from the beginning of the chamber to the eyeball detector, and $R = 2\,\mathrm{cm}$ is the distance from the laser axis to the eyeball (see Fig. 3.1).

The uncalibrated VSF measured by the eyeball detector can be calibrated in several ways. The default method applies a relative calibration, where the two outermost ring detectors are linearly extrapolated to match the the first eyeball detection angle $\theta = 15°$. The ratio between the extrapolated ring VSF and the uncalibrated eyeball VSF at $\theta = 15°$ is then used as scaling factor for the entire uncalibrated eyeball VSF. This method is very sensitive to measurement errors in the data used to calculate the scaling factor, since these errors will propagate to the entire eyeball VSF. The method applied to the experimental data in Paper I, II and III, is an absolute calibration. This method makes use of monodisperse polymer beads, for which the VSF is accurately known from Mie theory. A correction function $\kappa(\theta)$ can then be found by comparing measurement to theoretical VSF. Hence, the ring and eyeball detectors are treated separately.

3.2 Monte Carlo simulations

The Monte Carlo (MC) method can be a powerful tool for simulating light transport due to its flexibility. As oppose to solving the radiative transfer equation, the MC method can be applied to any geometry, making it very useful for simulating optical instruments. The MC method is a statistical approach, making use of probability density functions (PDFs) to represent the scattering and absorption properties of a sample, and must satisfy the condition

$$\int_{x_1}^{x_2} p(x)dx = 1, \tag{3.11}$$

where x_1 and x_2 is the lower and upper boundary of the variable x, respectively. The cumulative distribution function (CDF) gives the probability $P(x)$ that the variable x' is below or equal to the value of x, and is defined as

$$P(x) = \int_{x_1}^{x} p(x')dx'. \tag{3.12}$$

The trick used for MC simulation is to map the PDF onto a uniform distribution of numbers r limited to the interval [0,1], so that

$$\int_0^r p(r')dr' = \int_{x_1}^x p(x')dx' = P(x). \tag{3.13}$$

Hence, the probability that a number r' lies in the interval $[r, r+dr]$ is the same as the probability that the variable x' lies in the interval $[x, x+dx]$. Since r is uniformly distributed in the interval $[0,1]$, integrating the left-hand side of Eq. (3.13) gives

$$r = \int_{x_1}^x p(x')dx' = P(x) \tag{3.14}$$

Thus, by drawing a random number ξ from the uniform distribution of r and solving $\xi = P(x)$, one simultaneously draws a random variable x according to the PDF $p(x)$.

3.2.1 Monte Carlo simulation for the LISST-VSF

To accurately simulate the LISST-VSF instrument, the simulation should be as close to the real instrument as possible in terms of geometry and processing. Assuming that the geometry is very similar, the processing applied to the LISST-VSF ring detector signal can also be applied to the simulated ring detector signal, with only two exceptions. Firstly, the attenuation of pure water is not simulated, hence the processing does not correct for this. Secondly, the simulated signal is measured in terms of photon counts, rather than power. Hence, slight modification can be made to Eq. (3.9) to obtain the simulated VSF for the ring detector,

$$\beta_{i,p}^{\text{ring}} = \frac{N_{i,p}}{N_0} \frac{1}{2\pi\phi L(\cos\theta_i - \cos\theta_{i+1})}, \tag{3.15}$$

Here N_0 is the number of simulated photons, and the term $N_{i,p} = N_i/\tau$ is the number of detected photons N_i corrected for the attenuation $\tau = e^{-c_p L}$ (see Eq. (3.7)).

Replicating the LISST-VSF processing for the simulated eyeball detector signal is not as straight forward. The LISST-VSF eyeball signal is made up of four measurements with different combinations of incident light polarization and detected light polarization. Additionally, the eyeball detector has automatically adjusted gain levels and the laser is dimmed between 15° and ∼ 45°. Applying the absolute calibration described in Section 3.1.2, the eyeball VSF is obtained directly. Hence, the simulated eyeball detector signal can be processed similarly to the simulated ring detector signal, giving

$$\beta_{i,p}^{\text{eye}} = \frac{N_{i,p}}{N_0} \frac{\sin\theta/d}{2\pi\phi(\cos\theta_i - \cos\theta_{i+1})}. \tag{3.16}$$

As oppose to the ring detector, the eyeball only views a fraction of the laser beam. Thus, L in Eq. (3.15) is replaced by $d/\sin\theta$, where $d = 0.6\,\text{cm}$ is related to the simulation geometry (see Fig. 3.2(c)).

Geometry

While the MC method can in principle be applied to any geometry, this is not always straight forward. In some cases one may instead use a simulation geometry that

achieves a similar signal to the instrument geometry, but is much more simple to define in terms of mathematical expressions in the simulation code. In case of the ring detector, it is not necessary to simulate refraction from the lens and the photons position in the ring detector plane. Instead, one may only simulate the photon until it reaches the lens and register the angle of incidence, as this is the same as the angle of detection. The eyeball is more complicated, where the light passes through an opening in the eyeball, before it is guided by two mirrors to a spatial filter consisting of two lenses and a pinhole (see Fig. 3.2(b)). Rather than simulating this complex geometry, the detector is defined as a sphere of diameter $d = 0.6\,\text{cm}$ (see Fig. 3.2(c)). A photon passing through this sphere is approximately the same as a photon passing through the opening of the eyeball and being within the acceptance angle of the spatial filter.

While the design of the LISST-VSF instrument is restricted in terms of practicality, e.g. space for wires and electrical components, the simulation is not. Hence, one can make further adjustments to the detection geometry to increase the efficiency of the simulation. The ring detectors in the LISST-VSF only cover 1/6 of full circle, which can be extended to a full circle in the simulation. The same goes for the eyeball detector, where the detection sphere in the simulation can be extended around the full circumference of the chamber, as illustrated in Fig. 3.2. Furthermore, the simulated eyeball detector does not need to rotate in order to detect light at different angles, but can detect light at all angles simultaneously.

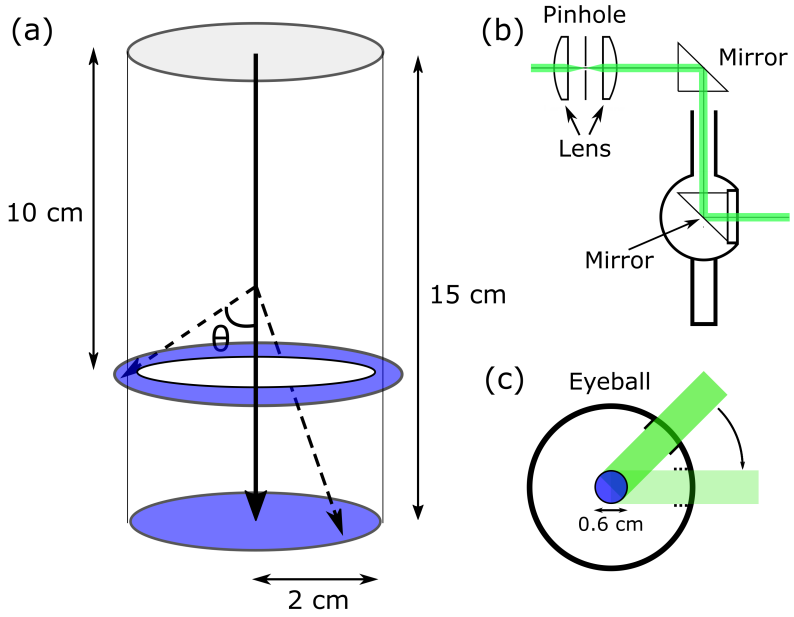

Figure 3.2: (a) Illustration of the Monte Carlo simulation geometry for the LISST-VSF instrument. Blue areas represent detection areas, while black lines represent the boundaries of the simulated volume. (b) Working principle of the eyeball detection scheme in the LISST-VSF instrument. (c) Detection scheme in the Monte Carlo simulation for the eyeball detector.

Probability density functions

In the LISST-VSF MC simulation, there are three variables that must be sampled from their respective PDFs: The pathlenght l between photon-particle interactions, scattering angle θ, and the azimuth angle φ. Assuming that the particles suspended in the water sample are randomly oriented, light scattering can be treated as azimuthal symmetric. Hence, the PDF for the azimuthal angle is

$$p(\varphi) = \frac{1}{2\pi}, \tag{3.17}$$

with the corresponding CDF

$$P(\phi) = \int_0^{\varphi} \frac{1}{2\pi} d\varphi = \frac{\varphi}{2\pi}. \tag{3.18}$$

Thus, the sampled azimuth angle is $\varphi = 2\pi\xi$, where ξ is a random number between 0 and 1.

The PDF for the photon pathlenght can be derived from Beer-Lamberts law

$$I(l) = I_0 e^{-cl}, \tag{3.19}$$

where I_0 is the intensity of the incident beam, c is the attenuation coefficient of the medium, and $I(l)$ is the intensity of the beam after traveling the pathlength l. As the fraction of photons that are not attenuated after traveling the distance l is $I(l)/I_0$, the CDF can be expressed as

$$P(l) = 1 - \frac{I(l)}{I_0} = 1 - e^{-cl}. \tag{3.20}$$

From Eq. (3.12), the PDF is then calculated as

$$p(l) = \frac{dP(l)}{dl} = ce^{-cl}. \tag{3.21}$$

Inserting Eq. (3.20) into Eq. (3.14), and substituting r with the random number ξ, one can solve for the pathlength l to obtain the equation

$$l = \frac{-\ln(1-\xi)}{c} = \frac{-\ln(\xi)}{c} \tag{3.22}$$

Here, $1-\xi$ and ξ are equivalent, as they are both uniformly distributed on the interval [0,1].

The PDF for the scattering angle θ is defined by the phase function $\tilde{\beta}(\theta)$, so that

$$p(\theta) = 2\pi \sin\theta \tilde{\beta}(\theta). \tag{3.23}$$

For natural samples, $\tilde{\beta}(\theta)$ is not an analytical expression, meaning that there is no analytical solution to Eq. (3.14) for this PDF. Instead, the integral in Eq. (3.14) must be solved numerically to find the CDF, which can be saved as a look-up table. Drawing a random number ξ, the sampled scattering angle corresponds to the point satisfying CDF $= \xi$. For the example shown in Fig. 3.3, drawing a random number $\xi = 0.6$ corresponds to a scattering angle $\theta = 54^\circ$, as indicated by the black lines in Fig. 3.3(c).

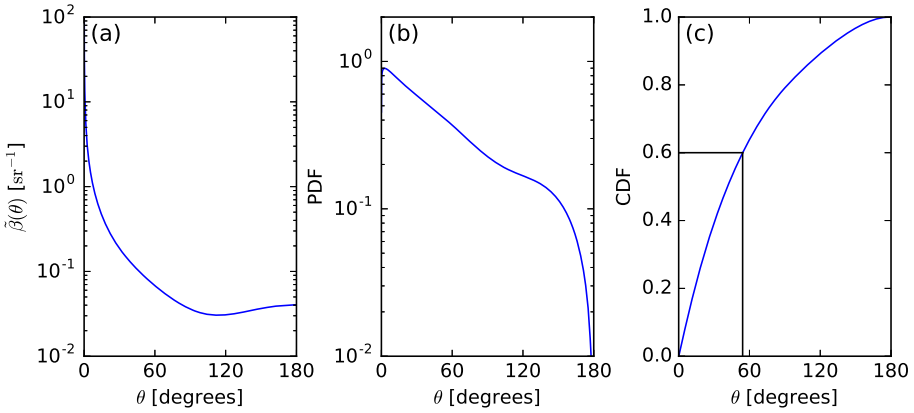

Figure 3.3: Phase function (a), probability density function (b), and cumulative density function (c) plotted for a Fournier-Forand phase function.

Historically, accurate measurements of VSFs have been few. Hence, analytical phase functions are frequently used to approximate the shape of natural phase functions. The Henyey-Greenstein phase function is a one parameter function that has been widely used in light transport simulations due to its simplicity, and is written as (see Eq. 2.5)

$$\tilde{\beta}_{\mathrm{HG}}(\theta) = \frac{1}{4\pi}\frac{1-g^2}{(1+g^2-2g\cos\theta)^{3/2}}. \tag{3.24}$$

Another advantage of the Henyey-Greenstein phase function is that Eq. 3.14 can be solved analytically, removing the need for look-up tables. The relation between the scattering angle θ and random number ξ can be expressed as

$$\theta = \cos^{-1}\left[\frac{1+g^2}{2g} - \frac{1}{2g}\left(\frac{1-g^2}{1+g-2g\xi}\right)^2\right] \tag{3.25}$$

There are other analytical phase functions, such as the Fournier-Forand and Two-term Reynolds-McCormick, that offer more flexibility in approximating natural phase functions. However, these are more complex and require that Eq. (3.14) is solved numerically.

3.2.2 Variance reduction methods

One of the major drawbacks of using MC methods for light transport simulations, is that it can be very time consuming. Hence, developing and implementing methods that increase the efficiency can save time in the long run. For optical instrumentation, one should first consider what does and what does not contribute to the signal. Depending on the sample properties, a major loss of signal comes from absorbed photons. Hence, simulating these photons is a waste of computation time. A very simple way of negating this problem, is to give the photons an initial weight of $w = 1$, and multiply the weight

with the single scattering albedo b/c for each scattering event. The detected signal is now the sum of weights rather than the sum of photons. For a specific number of simulated photons, the average detected signal would be the same as before, but is now distributed over more photons, reducing the variance. For the LISST-VSF specifically, the attenuation coefficient is defined in the simulation, eliminating the need to simulate the transmission measurement. Hence, all photons that are not scattered within the sample volume, but are directly transmitted, do not contribute to the signal. Normally, the pathlength is drawn from the PDF $p(l)$ given in Eq. 3.21. However, one can apply a method called mean free path transform (MFPTR), which utilizes a modified PDF $p_{\text{mod}}(l)$ to alter the pathlength distribution of photons. A change in the PDF must be reflected in the weight of the photon, which can be calculated as

$$w_{i+1} = w_i \frac{p(l)}{p_{\text{mod}}(l)}. \tag{3.26}$$

where w_i and w_{i+1} is the weight before and after the MFPTR, respectively. For the LISST-VSF simulation, we want all photons to be scattered before reaching the end of the sample volume. A very simple PDF that ensures this is $p_{\text{mod}}(l) = 1/L$ for $0 < l < L$, and $p_{\text{mod}}(l) = 0$ anywhere else. Here, L is the length of the sample volume. With this altered PDF, the randomly sampled pathlength is $l = \xi L$, and the weight is calculated as

$$w_{i+1} = w_i \frac{ce^{-cl}}{1/L}, \tag{3.27}$$

where c is the attenuation coefficient.

The final portion of photons that do not contribute to the signal, are photons that are scattered, but not detected. A technique, similar to the MFPTR, can be applied to the scattering angle, where a modified PDF is used to increase the probability of scattering towards the detector. This is called detector directional importance sampling (DDIS) (*Buras and Mayer*, 2011). To accurately simulate the LISST-VSF, photons must be allowed to interact with the entire sample volume. Hence, the modified PDF should not be applied to every scattering event, as this would send the majority of photons directly to the detector. Instead, the photon is given a probability ε to be scattered according to the modified PDF $p'_{\text{mod}}(\theta)$, meaning that the PDF used in the simulation can be expressed as

$$p_{\text{mod}}(\theta) = (1-\varepsilon)p(\theta) + \varepsilon p'_{\text{mod}}(\theta), \tag{3.28}$$

were $p(\theta)$ is the normal PDF calculated according to Eq. (3.23). The probability ε is user defined, and should be chosen to provide a good balance between photons that are scattered normally to explore the sample volume, and photons that are scattered towards the detector for increased efficiency. As with the MFPTR, the weight of the photons must be changed to reflect the change in the PDF, giving

$$w_{i+1} = w_i \frac{p(\theta)}{p_{\text{mod}}(\theta)}. \tag{3.29}$$

For the LISST-VSF specifically, there are several detectors. In general, the phase function of natural samples are forward peaked. Thus, the majority of photons are scattered

forward, as shown by the PDF in Fig. 3.3(b). As a result, the simulated signal is usually much stronger for the ring detector than for the eyeball detector. Thus, applying DDIS to increase scattering towards the eyeball detector may results in large improvements to the overall signal. Choosing a $\varepsilon \sim 0.2$, the weak eyeball detector signal can be significantly increased, while the strong ring detector signal is only slightly decreased.

3.3 Artificial neural network

The correction of multiple scattering errors is a very complicated problem, as the error depends on both the magnitude and shape of the VSF. In principle, the amount of the multiple scattered photons is determined by the scattering coefficient of the medium, while the angular distribution of these are determined by the phase function. Hence, information of the true VSF is embedded in the erroneous measurement, but is not easily retrievable by analytical means. Artificial neural networks (ANNs) have become an increasingly popular tool to solve such complex multivariate problems. ANNs are especially known for their generalization capabilities, making it a promising candidate for correction of multiple scattering errors in the LISST-VSF.

3.3.1 Feedforward neural network

Consider a feedforward neural network with input layer $\mathbf{x} = [x_1, x_2, \dots x_m]$, one hidden layer $\mathbf{h} = [h_1, h_2, \dots h_n]$ and output layer $\mathbf{y} = [y_1, y_2, \dots y_p]$, as illustrated in Fig. 3.4. The inputs $\mathbf{x}$ are fed forward to the hidden layer, as indicated by the connections (arrows). At the hidden layer, each node h_j is calculated as a function F of a linear combination z of the inputs,

$$h_j = F(z_j), \qquad z_j = \sum_{i=1}^{m} W_{ij}^{(1)} x_i + b_j^{(1)}. \tag{3.30}$$

Here, $W_{ij}^{(k)}$ are the weights (coefficients) that describe how much a node h_j depends on the input value x_i. The term $b_j^{(k)}$ is the bias for node h_j, and has a similar role to the constant b in a linear equation $y = ax + b$.

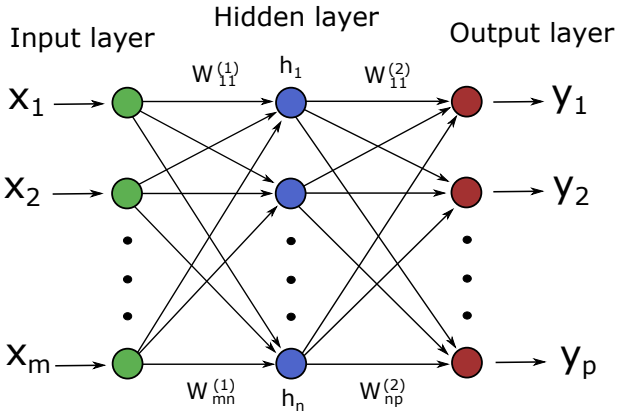

Figure 3.4: A neural network with m input nodes (green), p output nodes (red), and one hidden layer with n nodes (blue). The arrows indicate the direction of the connections between the nodes, where the weights $W_{ij}^{(k)}$ represent the strength of the connection between node i and node j.

The function F is called an activation function, and is the element that transforms the otherwise linear model to a nonlinear model. There are many different activation functions with their respective advantages and drawbacks. The Sigmoid function is a logistic function (s-shape), and is widely used as activation function. The function is limited between 0 and 1, and is thus suited for ANNs where the output is expressed in terms of probabilities. The hyperbolic tangent function is very similar to the Sigmoid function in terms of shape, but is limited between -1 and 1. An advantage of using the hyperbolic tangent function rather than the Sigmoid function, is that negative inputs are mapped to negative outputs. If the linear combination of inputs to node h_j is a large negative value, the output of the Sigmoid function is positive and close to 0, while the output of the hyperbolic tangent function is negative and close to -1, as shown in Fig. 3.5.

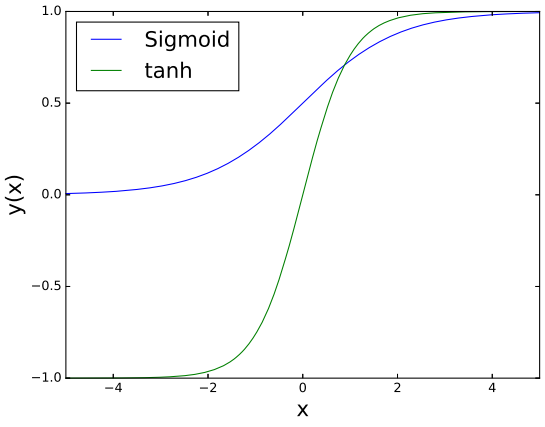

Figure 3.5: Plot of the Sigmoid (blue) and hyperbolic tangent (green) activation functions.

The output of the hidden layer is fed forward as inputs to the output layer, which is calculated according to Eq. (3.30), replacing h with y and x with h. Here, one can choose to omit the use of an activation function, in which case the output is calculated as a linear equation of the hidden layer output.

3.3.2 Backpropagation

Lets say we have a dataset with N input vectors of the form $\mathbf{x} = [x_1, x_2, \ldots x_m]$ for which we know the desired output values $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \ldots \hat{y}_p]$. We may now apply supervised learning using the backpropagation algorithm. The first step is to calculate the error E between the ANN output and the desired output. There are many ways to express the error, e.g. percentage deviation and mean absolute deviation, but the most common error function is the mean squared error (MSE),

$$E = \text{MSE} = \frac{1}{p}\sum_{j=1}^{p}(y_j - \hat{y}_j)^2. \tag{3.31}$$

The goal is to optimize the ANN on the given dataset by computing the weights and biases that minimize the error. How sensitive the error E is to a given weight $W_{ij}^{(k)}$ can be calculated by the derivative $\partial E/\partial W_{ij}^{(k)}$, where E descends fastest in the direction $-\partial E/\partial W_{ij}^{(k)}$. Hence, the weight can be updated according to

$$W_{ij}^{(k)} = W_{ij}^{(k)} - s\frac{\partial E}{\partial W_{ij}^{(k)}}, \tag{3.32}$$

where s is known as the step size or learning rate, defining how much the weight will be changed. The bias $b_j^{(k)}$ is updated similarly. All weights and biases must be updated according to Eq. (3.32), which is done efficiently with the backpropagation algorithm. The derivative of the error with respect to the weights connecting to the output layer can be calculated by applying the chain rule as follows

$$\frac{\partial E}{\partial W_{ij}^{(2)}} = \frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial z_j}\frac{\partial z_j}{\partial W_{ij}^{(2)}}. \tag{3.33}$$

The first derivative $\frac{\partial E}{\partial y_j}$ is calculated from Eq. (3.31), giving $2(y_j - \hat{y}_j)$. The second derivative $\frac{\partial y_j}{\partial z_j}$ is the derivative of the activation function with respect to z. Taking the hyperbolic tangent function as an example, we have $y_j = \tanh(z)$, and $\frac{\partial y_j}{\partial z_j} = 1 - \tanh^2(z) = 1 - y_j^2$. Finally, the derivative $\frac{\partial z_j}{\partial W_{ij}^{(2)}}$ is simply h_i, so that Eq. (3.33) becomes

$$\frac{\partial E}{\partial W_{ij}^{(2)}} = 2(y_j - \hat{y}_j)(1 - y_j^2)h_i. \tag{3.34}$$

The weights $W_{ij}^{(2)}$ can now be updated according to Eq. (3.32)

At this point I would like to inform the reader that remainder of this section is very technical and can be difficult to follow. Finding the derivatives to earlier layers in the

network is slightly more complicated, as each output node depends on all nodes in the previous layer. For instance, the weight $W_{11}^{(1)}$ influences the node h_1, which in turn influences all outputs $[y_1, y_2, \ldots y_p]$, as illustrated in Fig. 3.6. Here, the derivatives we are interested in are

$$\frac{\partial E}{\partial W_{ij}^{(1)}} = \frac{\partial E}{\partial h_j}\frac{\partial h_j}{\partial z_j}\frac{\partial z_j}{\partial W_{ij}^{(1)}}. \tag{3.35}$$

The last two derivatives in this equation are evaluated similarly to the last to derivatives in Eq. (3.33), giving $\frac{\partial h_j}{\partial z_j} = 1 - h_j^2$ and $\frac{\partial h_j}{\partial W_{ij}^{(1)}} = x_i$. For the first derivative, we can apply the general formula for finding the derivative of the error with respect to a node $h_j^{(L-1)}$ in layer $(L-1)$ in the network

$$\frac{\partial E}{\partial h_j^{(L-1)}} = \sum_{k=1}^{n_L} \frac{\partial E}{\partial h_k^{(L)}}\frac{\partial h_k^{(L)}}{\partial z_k^{(L)}}\frac{\partial z_k^{(L)}}{\partial h_j^{(L-1)}}, \tag{3.36}$$

For our specific case, the layer (L-1) is the hidden layer and (L) is the output layer. Hence, Eq. (3.36) becomes

$$\frac{\partial E}{\partial h_j} = \sum_{k=1}^{p} \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial z_k}\frac{\partial z_k}{\partial h_j}. \tag{3.37}$$

Note that Eq. (3.37) is almost identical to Eq. (3.33), except that this is now a sum over all output nodes, and the last derivative is with respect to the node h_j rather than the weight connecting h_j to the output layer. This, means that many of the terms involved in computing Eq. (3.36) were already calculated in the previous step (i.e. Eq. (3.33)). This is the power of the backpropagation algorithm, where we avoid doing duplicate calculations, allowing for efficient optimization of the ANN. This extends to a network with several hidden layers, where the necessary number of calculations only increase linearly with the number of layers.

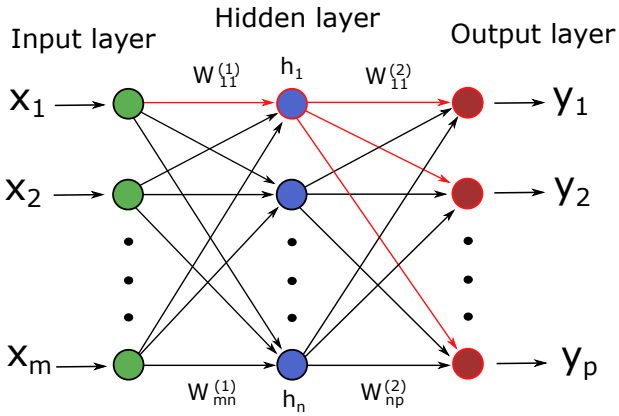

Figure 3.6: A feedforward neural network identical to the network presented in Fig. 3.4. The parts of the ANN that are influenced by the weight $W_{11}^{(1)}$ is indicated by the color red.

3.3.3 Applying an ANN for error correction

Now, we want to apply a feedforward neural network to correct multiple scattering errors in LISST-VSF measurements. A LISST-VSF measurement of the VSF contains 168 measurement points corresponding to the VSF over angles $0.09 - 150^{\circ}$. Hence, both the input layer and the output layer should have 168 nodes, where each node corresponds to the VSF value at a specific measurement angle. The input layer is the erroneous VSF measurement, while the output layer is the VSF corrected for multiple scattering errors. The size of the network (number of hidden layers and nodes per hidden layer) can be found through trail and error, i.e., testing different network sizes. Once the network is defined in terms of structure, we need a training dataset containing many pairs of input values $\mathbf{x} = [x_1, x_2, \ldots x_{168}]$ and corresponding desired output values $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \ldots \hat{y}_{168}]$. Here, the desired output values are the "true" VSF, i.e., the VSF that would be measured if no multiple scattered photons were detected. An example of an $\mathbf{x}$-$\hat{\mathbf{y}}$ pair is plotted in Fig. 3.7 for a simulated VSF. We may then optimize the ANN by calculating the error for the training dataset according to Eq. (3.31), and apply the backpropagation algorithm. This is an iterative process, were each iteration brings the output $\mathbf{y}$ of the ANN closer to the desired output $\hat{\mathbf{y}}$.

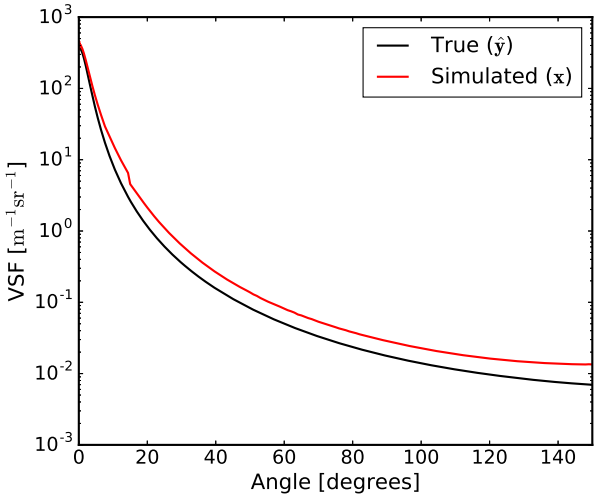

Figure 3.7: Semi-log plot of a simulated VSF $\mathbf{x}$ and the "true" VSF $\hat{\mathbf{y}}$, i.e., the single scattering VSF. The measurement points in the simulated VSF can be fed as inputs to an ANN, while the "true" VSF is the desired output of the ANN.

The goal is to train a generalized ANN model, meaning that the model can apply good corrections to data that is not in the training data. In order to achieve this, we need a large training dataset containing many VSFs of varying shape and magnitude. This presents a problem, as the "true" VSF $\hat{\mathbf{y}}$ is generally not known. When measuring the VSF at low particle concentrations, we can assume negligible multiple scattering errors. Thus, one way to obtain $\mathbf{x}$ and $\hat{\mathbf{y}}$ pairs is to make measurements on a sample with high particle concentration to obtain the erroneous measurement $\mathbf{x}$. Then, the sample can be

diluted to low particle concentration and measured again to obtain a measurement of the same particles, but now without multiple scattering errors. The estimated "true" VSF $\hat{\mathbf{y}}$ is then found by scaling the low concentration measurement with the dilution factor. While this method could work in theory, the process of obtaining a large and varied training dataset would be extremely tedious and time consuming. In addition, VSF measurements may contain significant errors that are not related to multiple scattering.

The solution to this problem, is to use the Monte Carlo simulation presented in Section 3.2.1. The simulation takes the "true" VSF $\hat{\mathbf{y}}$ as input, and simulates the LISST-VSF measurement $\mathbf{x}$, i.e., the reverse process of the ANN. Thus, the simulation can be used to generate a large training dataset in a relatively short amount of time. The input VSF to the simulation is user defined and can in principle be of any shape and magnitude. However, in order to efficiently train the model for purpose of correcting VSF measurements of natural samples, the training dataset should only contain VSFs that are similar to these. The VSFs can be computed from analytical phase functions, such as those presented in Section 2.1.1, and they can be obtained from measurements of the VSFs using the LISST-VSF (or other instruments). The main drawback with this method is that the Monte Carlo simulation does not perfectly simulate LISST-VSF measurements. Hence, any deviation between the simulation and the instrument will propagate to the ANN model.

Chapter 4

Statement of the problem

This thesis mainly revolves around multiple scattering errors in underwater optical measurements. Hence, this chapter is dedicated to a thorough description of the phenomena, and its effect on LISST-VSF measurements.

4.1 Multiple scattering

Multiple scattering is a known problem for underwater optical instrumentation, but is often considered negligible due to the typically short pathlength in the instruments. The literature provides several ranges for when the single scattering assumption is valid, i.e. negligible multiple scattering errors. The optical depth $\tau = cL$ has previously been used to define this range, where c is the attenuation constant and L is the pathlength. The single scattering condition $\tau < 1$ is given by *Hulst and van de Hulst* (1981), further stating that double scattering may become significant in the range $0.1 < \tau < 0.3$. In a study by *Koestner et al.* (2018), the single scattering condition is given as $\tau << 1$, while *Hu et al.* (2019) operate with the less restrictive $\tau < 1$. The single scattering condition is also sometimes given in terms of $\tau(1-g)$, where *Jonasz and Fournier* (2011) give $\tau(1-g) << 1$ as a general condition for single scattering. The term g is an asymmetry parameter (see Eq.3.24 for a definition of g), where a larger g means a more forward-peaked phase function. Hence, the condition $\tau(1-g) << 1$ also considers the shape of the VSF, saying that more forward peaked VSFs (larger g) give less multiple scattering errors.

4.2 Multiple scattering in the LISST-VSF

There are several sources of errors in LISST-VSF measurements, such as ambient light and laser drift. However, in turbid waters, multiple scattering of light is usually the dominant source of error. When processing LISST-VSF measurements, one assumes that light is only scattered once before it is detected. All light that is either absorbed or scattered away from the path to the detector is considered lost, and is added back to the signal by dividing the detected power with the attenuation factor τ (see Eq. (3.7) and Eq. (3.8)). However, light assumed lost due to scattering is not necessarily lost, as it may re-scatter (multiple scattered) and end up in one of the detectors (see Fig. 4.1). Hence, this light is both counted as lost (thus added back to the signal) and counted

as a real signal in a detector. The resulting error is observed as an increase in VSF magnitude and altered phase function shape. The magnitude of the error depends on both the scattering coefficient and phase function of the medium, which combined with the complex detection geometry, makes it a very difficult error to correct.

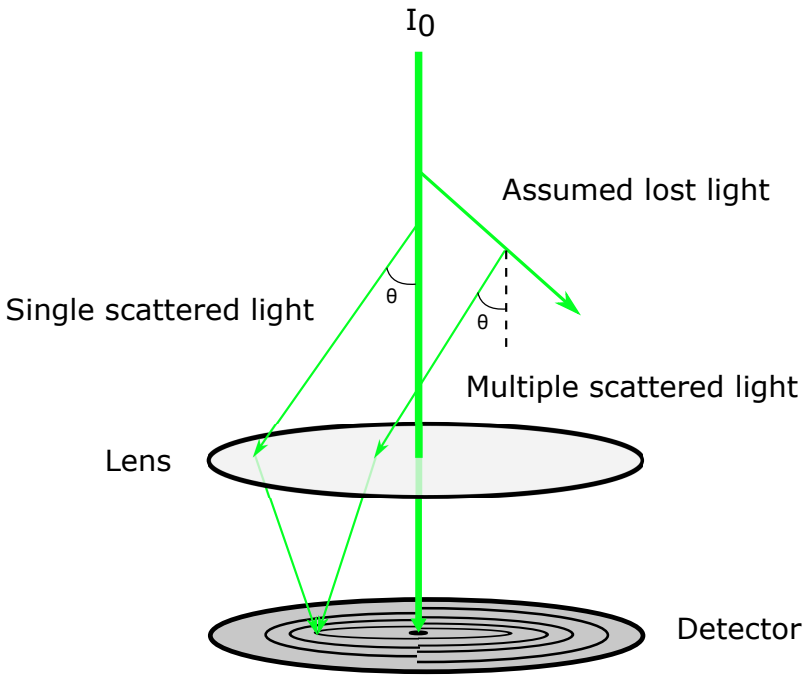

Figure 4.1: Illustration of multiple scattering in the LISST-VSF instrument. Light assumed lost due to scattering is re-scattered, hitting the lens at an incident angle θ and is refracted to the detector.

The effects of multiple scattering can be documented through controlled experiments in the lab. By employing a monodisperse solution of polystyrene beads as the scattering agent, a theoretical phase function can be calculated from Mie theory. Thus, a direct comparison between measured and theoretical VSF can be made. The results from a measurement series using polystyrene beads with narrow size distribution centered at 508 nm are presented in Fig. 4.2. The measurements are performed with increasing particle concentration, and compared to theoretical VSFs calculated from Mie theory.

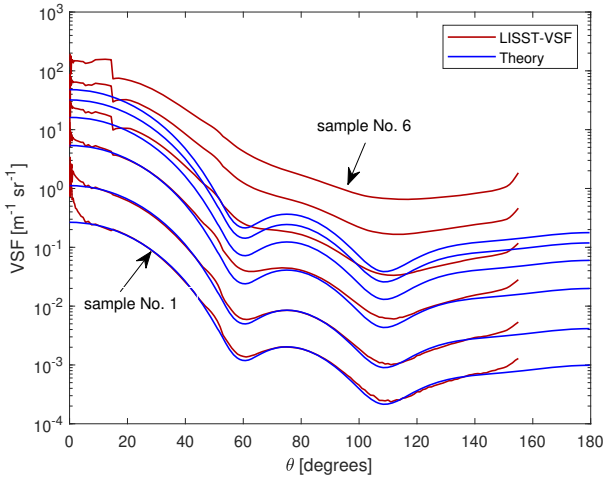

Figure 4.2: Semi-log plot of the VSF measured by the LISST-VSF (red) and calculated from Mie theory (blue). The VSF is plotted for six different particle concentrations of 508 nm *polystyrene beads, going from low (sample No. 1) to high (sample No. 6) particle concentration.*

At low particle concentration (sample No. 1), there is a good match between theory and measurement, with some deviation in the near forward ($\theta < 10^\circ$) and backward ($\theta > 140^\circ$) scattering angles. The overall deviation is observed to increase with increasing particle concentration, where a jump in the VSF at 15° can be observed from sample No. 3 and onward. At the highest particle concentration (sample No. 6), the measured VSF is much larger than the theoretical prediction, being over an order of magnitude larger at the dips located at $\theta \sim 60^\circ$ and $\theta \sim 110^\circ$. It can also be observed that the VSF shape becomes less similar to the original shape as the concentration increases, where the dips have almost completely vanished at the highest concentration.

Further investigation of multiple scattering errors can be done by using the Monte Carlo simulation described in Section 3.2.1. By simulating two different phase functions, we can observe how VSF shape influences multiple scattering errors. In Fig. 4.3, the results from simulating a Henyey-Greenstein (HG) and a Fournier-Forand (FF) phase function with a scattering coefficient of $b = 6\,\mathrm{m}^{-1}$ are shown. Both the theoretical $\mathrm{VSF}_{\mathrm{the}}(\theta)$ and simulated $\mathrm{VSF}_{\mathrm{sim}}(\theta)$ are shown, together with the angle dependent relative error calculated as

$$\mathrm{Error}(\theta) = \frac{\mathrm{VSF}_{\mathrm{sim}}(\theta) - \mathrm{VSF}_{\mathrm{the}}(\theta)}{\mathrm{VSF}_{\mathrm{the}}(\theta)} \tag{4.1}$$

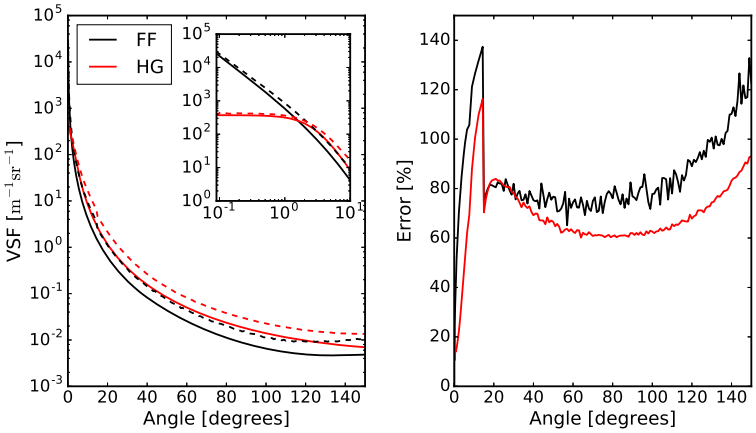

Figure 4.3: (a) Semi-log plot of the theoretical (solid) and simulated (dashed) VSF for a Fournier-Forand (black) and Henyey-Greenstein (red) phase function. The VSF is simulated for a scattering coefficient of $b = 6\,\mathrm{m}^{-1}$. The inset shows a log-log plot for angles $\theta < 10°$. (b) Error of the simulated VSF relative to the theoretical VSF, calculated according to Eq. (4.1)

As observed in Fig. 4.3, the VSFs simulated from both the FF and HG phase functions are severely influenced by multiple scattering at $b = 6\,\mathrm{m}^{-1}$. They are different in terms of shape, where the main difference is in the near forward part (see inset), where the HG plateaus while the FF keeps rising. This has a large impact on the error (Fig. 4.3(b)), where the FF VSF suffers from a larger error over the entire angular range, with the exception of the range $\theta \sim 15-30°$. For both VSFs, the error peaks at the last detection angle ($\theta = 14.38°$) in the ring detector, but drops quickly as θ approaches $0°$. The largest difference in error between the two VSFs is observed in the near forward part ($\theta = 0.09-6°$), where the error in the FF VSF is over twice the error in the HG VSF for most of this angular range.

With the simulation at hand, we may look further into the details of multiple scattering by looking at how multiple scattered photons are distributed over the LISST-VSF angular range. Some statistics are presented for the FF VSF in Fig. 4.4. Due to the large difference in signal over the angular spectrum, the statistics are divided into three parts: Ring detector (top row), eyeball detector for $\theta = 15-90°$ (middle row), and eyeball detector for $\theta = 90-150°$ (bottom row). The first column shows the percentage of detected photons per number of scattering events before detection, i.e. percentage of detected single scattered photons, percentage of detected double scattered photons, etc. Looking only at detected multiple scattered photons, the angular distribution after the first scattering event θ_1 and the last scattering event θ_d (detected scattering angle) is shown in the second column, and the corresponding change in angle $\Delta\theta = \theta_d - \theta_1$ between these are shown in the third column.

From the data presented in the first column in Fig. 4.4, one can see that the signal from multiple scattered photons accounts for a significant part of the total signal for all three intervals. The percentage of the total signal coming from multiple scattered photons is smallest for the ring detector, and largest for the backscattering region of the eyeball detector, where it is almost 50%.

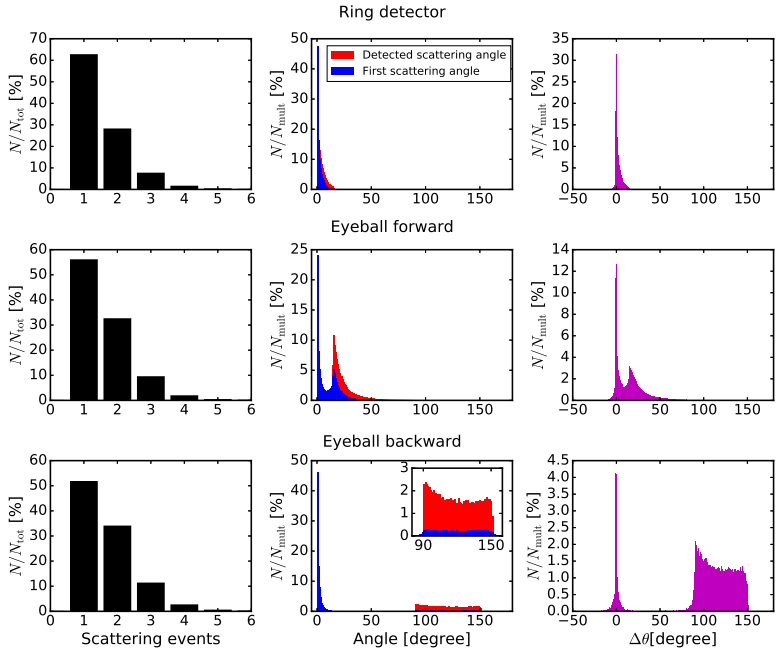

Figure 4.4: Statistics for a Fournier-Forand VSF simulated with a scattering coefficient of $b = 6\,\mathrm{m}^{-1}$. The angular range is divided into three intervals: ring detector (top row, $\theta = 0.09^\circ - 15^\circ$), eyeball forward (middle row, $\theta = 0.09^\circ - 15^\circ$), and eyeball backward (bottom row, $\theta = 0.09^\circ - 15^\circ$). The first column shows the number of detected photons N as a percentage of the total number of detected photons N_{tot} per scattering event. The second row shows the the number of multiple scattered photons N per angle as percentage of the total number of multiple scattered photons N_{mult} within the interval. Scattering angle after the first scattering event is shown as blue, and the detected angle (angle after last scattering event) is shown as red. The corresponding change in angle $\Delta\theta$ between the first and last scattering event is shown in the third column.

The second column in Fig. 4.4 shows the angular distribution of detected multiple scattered photons after the first scattering event and the last scattering event (detected scattering angle). For the ring detector, it can be seen that the angular distribution is very forward peaked after the first scattering event, which is expected from the very forward peaked FF phase function. This distribution becomes broader, as the photons are scattered more times. This is an expected result that agrees well with the study by *Piskozub and McKee* (2011), showing that multiple scattering generally results in a flattening of the phase function. The broadening of the peak is also evident from the third column, where $\Delta\theta$ is mostly positive, meaning that the detected scattering angle is larger than the first scattering angle. While the majority of detected multiple scattered photons have an original scattering angle $\theta < 15^\circ$, there is also a very small contribution (not visible in the plot) from photons originating from larger scattering angles.

Things become more interesting for the eyeball detector, where the angular distri-

bution after the first scattering event is divided into two main parts. For both forward and backward detected scattering angles, the largest peak is in the near forward direction, meaning that photons are first scattered at a very small angle and continues almost parallel to the incident laser beam. Subsequent scattering events redirect the photons toward the eyeball detector, resulting in the rightmost $\Delta\theta$ peak in the third column. The second peak (see inset in the third row for backscattering angles) is photons that are originally scattered in the general direction of the eyeball detector. Subsequent scattering events only change the scattering angle by a small amount, resulting in the leftmost $\Delta\theta$ peak in the third column.

The same statistics are given for the HG VSF in Fig. 4.5. Comparing to the statistics for FF VSF presented in Fig. 4.4, we see that the single scattering contribution is larger for the HG VSF in all parts of the spectrum. Unlike the FF VSF, the HG VSF has the highest contribution of multiple scattered photons in the forward part of the eyeball detector. Looking at the angular distribution of scattering angles, we see the same patterns as observed in Fig. 4.4. One key difference is that the peaks in the distribution are generally broader for the HG VSF, owing to the flatter peak in the forward direction of the phase function (see inset Fig. 4.3(a)). Also, the two peaks in the plot of $\Delta\theta$ (third column) for the eyeball detector are more comparable in height than for the FF VSF.

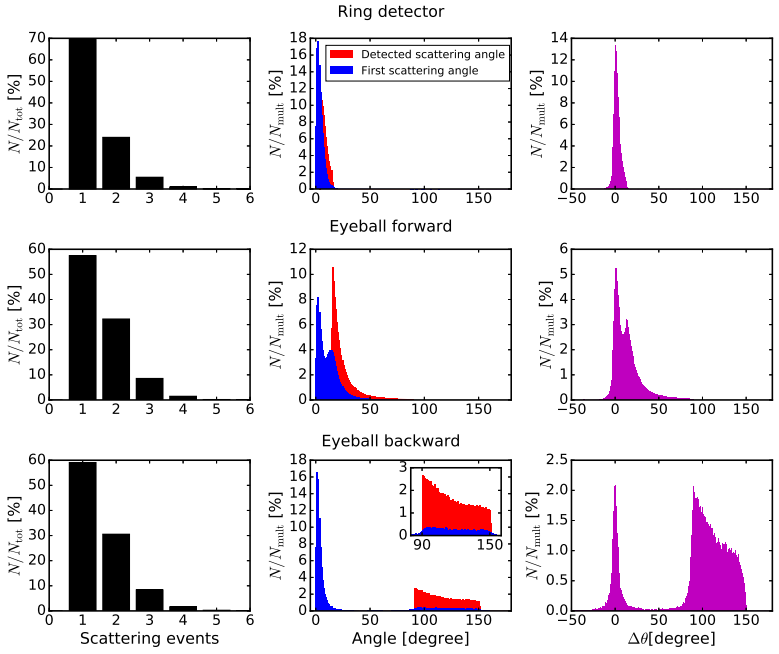

Figure 4.5: Statistics for a Henyey-Greenstein VSF simulated with a scattering coefficient of $b = 6\,\mathrm{m}^{-1}$. See caption in Fig. 4.4 for further description.

Chapter 5

Introduction to the papers

Paper I: **Analysis of multiple scattering errors in LISST-VSF volume scattering function measurements using Monte Carlo simulations and experimental data**

Ugulen, H. S., Sandven, H., Hamre, B., Kristoffersen, A. S., and Sætre, C. (2021), Optics Express, 29(8).

Paper I investigates multiple scattering effects in LISST-VSF measurements. The study was motivated by discrepancies seen in measured data, where multiple scattering was the suspected source. A Monte Carlo simulation was developed to simulate LISST-VSF measurements in order to test this hypothesis.

The Monte Carlo simulation was verified by comparison with controlled measurements performed in the lab, where 508 nm polymer beads and Arizona test dust was measured at varying concentrations. The results showed that the Monte Carlo simulation accurately reproduced LISST-VSF measurements at all concentrations, even at concentrations far into the multiple scattering regime. Hence, the Monte Carlo model was validated, and we confirmed that the observed discrepancies in LISST-VSF measurements were indeed due to multiple scattering.

In this paper, I coded the Monte Carlo simulation with some starting help from my supervisor Børge Hamre. I contributed to the lab measurements, which were mainly conducted by Håkon Sandven. Analyzes and interpretation of the results were done in cooperation with co-authors. I also wrote the majority of the manuscript, with inputs and revisions from co-authors, and made all the figures.

Paper II: **Efficient Monte Carlo simulation reveals significant multiple scattering errors in underwater angular scattering measurements**

Ugulen, H. S., Sandven, H., Hamre, B., Kristoffersen, A. S., and Sætre, C. (2022), Optics Express, 30(7).

Paper II builds on the results from Paper I. The first part of the study revolved around making the Monte Carlo simulation more efficient by implementing a number of variance reduction methods. The most important of these is known as detector directional importance sampling, which increases the probability of photons being scattered towards the detector in the simulation. The results showed an over 10x increase in effi-

ciency for scattering coefficients $b < 6\,\mathrm{m}^{-1}$.

In the second part of the study, we used the Monte Carlo simulation to look further into how multiple scattering affect LISST-VSF measurements. By using a variety of Henyey-Greenstein and Fournier-Forand phase functions we showed that multiple scattering errors can vary significantly depending on the phase function of the sample.

In this study, I did all the coding for the improved Monte Carlo simulation, and ran the simulations. Interpretation and analysis were done in cooperation with co-authors. I wrote the majority of the manuscript, with comments and revisions from co-authors, and made all the figures.

Paper III: A neural network approach for correction of multiple scattering errors in the LISST-VSF instrument

Ugulen, H. S., Koestner, D., Sandven, H., Hamre, B., Kristoffersen, A. S., and Sætre, C. (2023), Optics Express, (Submitted, May 2023).

In Paper III, we developed a feedforward neural network with 3 hidden layers for correction of multiple scattering errors in LISST-VSF measurements. The neural network was trained on a dataset generated with the Monte Carlo simulation developed in Paper I and Paper II. The neural network was initially tested with high concentration samples obtained from two different stations. For these samples, estimated "true" VSFs were obtained from dilution corrections. The results showed that the neural network corrected VSF was very similar to the dilution corrected VSFs, both in shape and in magnitude.

Further testing was performed with samples for which we did not have dilution corrections. In this case, the neural network was evaluated by comparing the neural network estimated scattering coefficient b and the measured attenuation coefficient c. The results showed that for $b > 4$ the measured b was larger than the measured c, meaning that the absorption coefficient $a = c - b$ is negative. On the other hand, the neural network estimated b was always lower than the measured c, resulting in a positive a.

My contribution to this study was coding the neural network, generating the training data set, and testing and validating the neural network. I wrote the majority of the manuscript and made all the figures. All experimental data used for training and testing of the neural network was collected by co-authors.

Bibliography

Aggarwal, C. C., et al. (2018), Neural networks and deep learning, *Springer*, *10*(978), 3. 2.3

Ali, A. T., H. S. Abdullah, and M. N. Fadhil (2021), Voice recognition system using machine learning techniques, *Materials Today: Proceedings*, pp. 1–7. 2.3

Binzoni, T., T. S. Leung, A. H. Gandjbakhche, D. Ruefenacht, and D. Delpy (2006), The use of the henyey–greenstein phase function in monte carlo simulations in biomedical optics, *Physics in Medicine & Biology*, *51*(17), N313. 2.1.1

Buras, R., and B. Mayer (2011), Efficient unbiased variance reduction techniques for monte carlo simulations of radiative transfer in cloudy atmospheres: The solution, *Journal of quantitative spectroscopy and radiative transfer*, *112*(3), 434–447. 3.2.2

Caldeira, K., and M. E. Wickett (2003), Anthropogenic carbon and ocean ph, *Nature*, *425*(6956), 365–365. 1

Carleo, G., I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová (2019), Machine learning and the physical sciences, *Reviews of Modern Physics*, *91*(4), 045,002. 2.3.2

Carlson, J., S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. Schmidt, and R. B. Wiringa (2015), Quantum monte carlo methods for nuclear physics, *Reviews of Modern Physics*, *87*(3), 1067. 2.2.1

Celebi, M. E., and K. Aydin (2016), *Unsupervised learning algorithms*, vol. 9, Springer. 2.3.1

Chandolikar, N., C. Joshi, P. Roy, A. Gawas, and M. Vishwakarma (2022), Voice recognition: A comprehensive survey, in *2022 International Mobile and Embedded Technology Conference (MECON)*, pp. 45–51, IEEE. 2.3

Chandrasekhar, S. (2013), *Radiative transfer*, Courier Corporation. 2.2

Chebud, Y., G. M. Naja, R. G. Rivero, and A. M. Melesse (2012), Water quality monitoring using remote sensing and an artificial neural network, *Water, Air, & Soil Pollution*, *223*, 4875–4887. 2.3.2

Chen, S., C. Hu, B. B. Barnes, R. Wanninkhof, W.-J. Cai, L. Barbero, and D. Pierrot (2019), A machine learning approach to estimate surface ocean pco2 from satellite measurements, *Remote Sensing of Environment*, *228*, 203–226. 2.3.2

Cortes, C., and V. Vapnik (1995), Support-vector networks, *Machine learning*, *20*, 273–297. 2.3

Costanza, R. (1999), The ecological, economic, and social importance of the oceans, *Ecological economics*, *31*(2), 199–213. 1

Doxaran, D., E. Leymarie, B. Nechad, A. Dogliotti, K. Ruddick, P. Gernez, and E. Knaeps (2016), Improved correction methods for field measurements of particulate light backscattering in turbid waters, *Optics express*, *24*(4), 3615–3637. 2.2.1

Fournier, G. R., and J. L. Forand (1994), Analytic phase function for ocean water, in *Ocean Optics XII*, vol. 2258, pp. 194–201, SPIE. 2.1.1

Fradkov, A. L. (2020), Early history of machine learning, *IFAC-PapersOnLine*, *53*(2), 1385–1390. 2.3

Gabriel, C., M.-A. Khalighi, S. Bourennane, P. Léon, and V. Rigaud (2013), Monte-carlo-based channel characterization for underwater optical communication systems, *Journal of Optical Communications and Networking*, *5*(1), 1–12. 2.2.1

Haltrin, V. I. (2002), One-parameter two-term henyey-greenstein phase function for light scattering in seawater, *Applied Optics*, *41*(6), 1022–1028. 2.1.1

Hammond, B. L., W. A. Lester, and P. J. Reynolds (1994), *Monte Carlo methods in ab initio quantum chemistry*, vol. 1, World Scientific. 2.2.1

Hansen, J., R. Ruedy, M. Sato, and K. Lo (2010), Global surface temperature change, *Reviews of Geophysics*, *48*(4). 1

Harmel, T., J. Agagliate, M. Hieronymi, and P. Gernez (2021), Two-term reynolds–mccormick phase function parameterization better describes light scattering by microalgae and mineral hydrosols, *Optics Letters*, *46*(8), 1860–1863. 2.1.1

Henyey, L. G., and J. L. Greenstein (1941), Diffuse radiation in the galaxy, *Astrophysical Journal, vol. 93, p. 70-83 (1941).*, *93*, 70–83. 2.1.1

Hu, L., X. Zhang, Y. Xiong, and M.-X. He (2019), Calibration of the lisst-vsf to derive the volume scattering functions in clear waters, *Optics Express*, *27*(16), A1188–A1206. 3.1.2, 4.1

Hulst, H. C., and H. C. van de Hulst (1981), *Light scattering by small particles*, Courier Corporation. 4.1

Jasman, F., and R. Green (2013), Monte carlo simulation for underwater optical wireless communications, in *2013 2nd international workshop on optical wireless communications (IWOW)*, pp. 113–117, IEEE. 2.2.1

Jayne, S. R., D. Roemmich, N. Zilberman, S. C. Riser, K. S. Johnson, G. C. Johnson, and S. R. Piotrowicz (2017), The argo program: present and future, *Oceanography*, *30*(2), 18–28. 1

Jonasz, M., and G. Fournier (2011), *Light scattering by particles in water: theoretical and experimental foundations*, Elsevier. 2.1.1, 4.1

Jordan, M. I., and T. M. Mitchell (2015), Machine learning: Trends, perspectives, and prospects, *Science*, *349*(6245), 255–260. 2.3.1

Juan, N. P., and V. N. Valdecantos (2022), Review of the application of artificial neural networks in ocean engineering, *Ocean Engineering*, *259*, 111,947. 2.3.2

Koestner, D., D. Stramski, and R. A. Reynolds (2018), Measurements of the volume scattering function and the degree of linear polarization of light scattered by contrasting natural assemblages of marine particles, *Applied Sciences*, *8*(12), 2690. 4.1

Levin, L. A., B. J. Bett, A. R. Gates, P. Heimbach, B. M. Howe, F. Janssen, A. McCurdy, H. A. Ruhl, P. Snelgrove, K. I. Stocks, et al. (2019), Global observing needs in the deep ocean, *Frontiers in Marine Science*, *6*, 241. 1

Lou, R., Z. Lv, S. Dang, T. Su, and X. Li (2021), Application of machine learning in ocean data, *Multimedia Systems*, pp. 1–10. 2.3.2

McKee, D., J. Piskozub, and I. Brown (2008), Scattering error corrections for in situ absorption and attenuation measurements, *Optics Express*, *16*(24), 19,480–19,492. 2.2.1

Medsker, L. R., and L. Jain (2001), Recurrent neural networks, *Design and Applications*, *5*, 64–67. 2.3.1

Metropolis, N., and S. Ulam (1949), The monte carlo method, *Journal of the American statistical association*, *44*(247), 335–341. 2.2.1

Mie, G. (1908), Beiträge zur optik trüber medien, speziell kolloidaler metallösungen, *Annalen der physik*, *330*(3), 377–445. 2.1.1

Noebauer, U. M., and S. A. Sim (2019), Monte carlo radiative transfer, *Living Reviews in Computational Astrophysics*, *5*, 1–103. 2.2.1

Piskozub, J., and D. McKee (2011), Effective scattering phase functions for the multiple scattering regime, *Optics express*, *19*(5), 4786–4794. 4.2

Preisendorfer, R. W. (1986), Secchi disk science: Visual optics of natural waters 1, *Limnology and oceanography*, *31*(5), 909–926. 1

Reynolds, L., and N. McCormick (1980), Approximate two-parameter phase function for light scattering, *JOSA*, *70*(10), 1206–1212. 2.1.1

Rogers, D. (2006), Fifty years of monte carlo simulations for medical physics, *Physics in Medicine & Biology*, *51*(13), R287. 2.2.1

Rosenblatt, F. (1957), *The perceptron, a perceiving and recognizing automaton Project Para*, Cornell Aeronautical Laboratory. 2.3

Rosenblatt, F. (1958), *Two theorems of statistical separability in the perceptron*, United States Department of Commerce Washington, DC, USA. 2.3

Rosenblatt, F. (1960), Perceptron simulation experiments, *Proceedings of the IRE*, *48*(3), 301–309. 2.3

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1985), Learning internal representations by error propagation, *Tech. rep.*, California Univ San Diego La Jolla Inst for Cognitive Science. 2.3

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Learning representations by back-propagating errors, *nature*, *323*(6088), 533–536. 2.3

Sagawa, T., Y. Yamashita, T. Okumura, and T. Yamanokuchi (2019), Satellite derived bathymetry using machine learning and multi-temporal satellite images, *Remote Sensing*, *11*(10), 1155. 2.3.2

Sandven, H., A. S. Kristoffersen, Y.-C. Chen, and B. Hamre (2020), In situ measurements of the volume scattering function with lisst-vsf and lisst-200x in extreme environments: evaluation of instrument calibration and validity, *Optics Express*, *28*(25), 37,373–37,396. 3.1.2

Schmidhuber, J. (2015), Deep learning in neural networks: An overview, *Neural networks*, *61*, 85–117. 2.3.1

Sharaf El Din, E., Y. Zhang, and A. Suliman (2017), Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework, *International Journal of Remote Sensing*, *38*(4), 1023–1042. 2.3.2

Slade, W. H., and E. Boss (2015), Spectral attenuation and backscattering as indicators of average particle size, *Applied optics*, *54*(24), 7264–7277. 2.1.1

Stamnes, K., G. E. Thomas, and J. J. Stamnes (2017), *Radiative transfer in the atmosphere and ocean*, Cambridge University Press. 2.2

Svozil, D., V. Kvasnicka, and J. Pospichal (1997), Introduction to multi-layer feed-forward neural networks, *Chemometrics and intelligent laboratory systems*, *39*(1), 43–62. 2.3.1

Tripathi, B. K. (2017), On the complex domain deep machine learning for face recognition, *Applied Intelligence*, *47*(2), 382–396. 2.3

Wei, J., X. Chu, X.-Y. Sun, K. Xu, H.-X. Deng, J. Chen, Z. Wei, and M. Lei (2019), Machine learning in materials science, *InfoMat*, *1*(3), 338–358. 2.3.2

Werdell, P. J., L. I. McKinna, E. Boss, S. G. Ackleson, S. E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, et al. (2018), An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing, *Progress in oceanography*, *160*, 186–212. 1

Wiering, M. A., and M. Van Otterlo (2012), Reinforcement learning, *Adaptation, learning, and optimization*, *12*(3), 729. 2.3.1

Chapter 6

Scientific results

Paper I

Analysis of multiple scattering errors in LISST-VSF volume scattering function measurements using Monte Carlo simulations and experimental data

Håvard S. Ugulen, Håkon Sandven, Børge Hamre, Arne S. Kristoffersen, and Camilla Sætre
Optics Express, **29(8)** (2021)

Research Article | Vol. 29, No. 8 / 12 April 2021 / Optics Express 12413
Optics EXPRESS

Analysis of multiple scattering errors in LISST-VSF volume scattering function measurements using Monte Carlo simulations and experimental data

HÅVARD S. UGULEN,* HÅKON SANDVEN, BØRGE HAMRE, ARNE S. KRISTOFFERSEN, AND CAMILLA SÆTRE

Department of Physics and Technology, University of Bergen, Allegaten 55, 5007 Bergen, Norway
**havard.ugulen@uib.no*

Abstract: A Monte Carlo algorithm has been developed to investigate the effects of multiple scattering on the volume scattering function measured by the LISST-VSF instrument. The developed algorithm is compared to experimental results obtained from bench-top measurements using 508 nm spherical polystyrene beads and Arizona test dust as scattering agents. The Monte Carlo simulation predicts measured volume scattering functions at all concentrations. We demonstrate that multiple scattered light can be a major contributor to the detected signal, resulting in errors in the measured volume scattering function and its derived inherent optical properties. We find a relative error of 10% in the scattering coefficient for optical depths ~0.4, and it can reach 100% at optical depths ~2.

1. Introduction

Multiple scattering may be a significant source of systematic errors when measuring the inherent optical properties (IOPs) of particle-rich natural waters [1,2]. IOPs include the absorption coefficient and the volume scattering function (VSF) with its derived properties, such as the scattering coefficient b, and the backscattering coefficient b_b, all of which may vary spectrally. Turbid particle-rich waters are characteristically found in coastal and estuarine regions, which are of increased interest to the optical oceanography community due to the ecological and economic importance of these regions [3–7]. Increased knowledge of IOPs in such optically complex waters would be beneficial for remote sensing observations, environmental monitoring and underwater optical communication and visibility studies. The assumption of single scattering, where each photon is only scattered once, is an important approximation when measuring the VSF or other IOPs, as this negates the use of extensive radiative transfer calculations. However, this assumption will no longer be valid if the amount of multiple scattering is too high. In a study by Chami et al. [8], it was found through radiative transfer modeling that multiple scattering may contribute to as much as ~ 94% of the radiance reflectance when the ratio of backscattering to absorption is larger than 0.3.

The VSF $\beta(\theta)$ is defined as the radiant intensity dI scattered per elemental volume dV in the direction θ per unit incident irradiance E,

$$\beta(\theta) = \frac{\mathrm{d}I}{E\,\mathrm{d}V}\ [\mathrm{m}^{-1}\mathrm{sr}^{-1}]. \tag{1}$$

Here, no azimuthal dependency is assumed. From this expression one can derive the scattering coefficient,

$$b = 2\pi \int_0^\pi \beta(\theta) \sin\theta d\theta. \tag{2}$$

#419116 https://doi.org/10.1364/OE.419116
Journal © 2021 Received 11 Jan 2021; revised 19 Feb 2021; accepted 19 Feb 2021; published 7 Apr 2021

Optics EXPRESS

The scattering phase function is given as

$$p(\theta) = \frac{\beta(\theta)}{b}. \tag{3}$$

The VSF has been sparingly measured *in situ*, especially at large scattering angles, due to lack of available instrumentation. The LISST-VSF (Sequoia Scientific) is a recently developed commercial instrument for measuring the VSF *in situ* or in the laboratory for angles 0.1-150° at a 515 nm wavelength [9]. The instrument uses a ring detector for measuring the light scattered from the incident laser beam (similar to other LISST-instruments) up to a scattering angle of 14.4°, while a rotating eyeball detector is used for angles 15-150°. While the geometries of the two detectors are different, both make the assumption that all detected light has been scattered solely from the beam. Light lost due to absorption or secondary scattering along the path of the scattered light is corrected for (see Section 2.2.2), but no additional light is assumed to be scattered into the detectors. It has been shown that applying an absolute calibration of the eyeball detector, using a method first presented by Hu et al. [10], yields a distinct discontinuity between the ring detector VSF and eyeball detector VSF in turbid waters [1]. In addition, the scattering values seem to be unreasonably elevated compared to parallel LISST-200X measurements, for which the sample chamber has a much shorter optical path length of 2.5 cm, compared to the 15 cm in the LISST-VSF. Microscopic polystyrene beads of precise and accurate size and concentration enable direct comparison of Mie theory and LISST-VSF measurements. Thus, polystyrene beads has been used to calibrate the LISST-VSF instrument [10–13]. In earlier laboratory measurements with 190 and 508 nm diameter polystyrene beads, we have seen similar VSF discontinuities at larger bead concentrations, as well as a non-linear relationship between scattering and particle concentration for 25 μm beads [1]. The discrepancies seen in both field and laboratory measurements motivate a deeper inquiry into the multiple scattering effects on the LISST-VSF measurements.

The single scattering approximation greatly simplifies the computation of IOPs from *in situ* measurements. The optical depth $\tau = cL$ has previously been used to state whether the single scattering assumption is a good approximation or not. Here, c is the attenuation coefficient, and L is the distance travelled in the medium. In the case of the LISST-VSF, this distance is $L = 0.15$m, which is the characteristic path length of the instrument. The closely related scaled optical depth $\tau^* = cL(1-g) = \tau(1-g)$, is also used. Here, g is known as the asymmetry factor, calculated as the mean cosine of the scattering angle of the phase function [14]. In a paper by Koestner et al. [12], the general condition for single scattering is first given as $\tau << 1$. They also state that $\tau^* << 1$ can be used to determine the regime where the single scattering approximation is valid. However, these two conditions are quite different, as the asymmetry factor g is typically between 0.89 and 0.95 for natural waters [14]. In a study by Hu et al. [10], $c = 5$ m^{-1} is stated as the upper limit for single scattering, which translates to $\tau = 0.75$ for the LISST-VSF instrument. By contrast, van de Hulst [15] state a significantly stricter general condition for single-scattering: For an optical depth $\tau < 0.1$ (corresponding approximately to $\tau^* < 0.01$), single scattering can be assumed. For $0.1 < \tau < 0.3$, double-scattering corrections may be necessary, and for $\tau > 0.3$ multiple scattering must be taken into account. In a study by Agrawal and Mengüç [16], unnormalized VSF measurements were made using a nephelometer (with a fixed path length) over a wide range of mono- and polydispersed polystyrene bead concentrations. Here, the single-scattering condition was shown to agree well with the measurements up to an optical depth of 0.1, and by including an analytical double-scattering term, there was good agreement between theoretical results and measurements up to $\tau = 1.0$. Similar measurement series with increasing bead concentrations were performed by Chae and Lee [17], where a non-linear trend could be seen when scattering measurements were compared with optical thickness for $\tau > 0.1$, although the authors concluded that the relationship was reasonably linear.

Optics EXPRESS

Monte Carlo (MC) simulations offer an excellent tool for investigating optical instrumentation due to the straight-forward implementation of geometry in the simulations. Analytical solutions to light transport can be difficult, if not impossible, to find without the use of approximations. MC simulations, are on the other hand numerical solutions and can usually be implemented without approximations. MC simulations have previously been used to investigate a variety of optical instruments. McKee et al. [18] applied a MC simulation in order to investigate the scattering collection performance of the AC-9 (WET Labs, now Sea-Bird Scientific) reflecting tube absorption meter. Based on the MC simulation, they developed an iterative scattering error correction procedure, which was later improved upon [19]. In Doxaran et al. [20] and Vadakke-Chanat et al. [21], the uncertainty and effects of absorption was analyzed and quantified for back scattered light in the ECO-BB9 (WET Labs, now Sea-Bird Scientific) sensor. Other examples can be found in [22–24]. In this paper, we present a MC simulation to investigate the effects of multiple scattering in LISST-VSF measurements. The aim of the work is to explain the observed features in the measured VSF, and to quantify errors in the phase function and scattering coefficient originating from multiple scattering.

2. Methods

2.1. LISST-VSF measurements

Two different scattering agents are used in this study: Polystyrene beads and Arizona test dust. The polystyrene beads solution is monodisperse with a particle size distribution centered at 508nm and with a full-width half-maximum of 16 nm. The Arizona test dust solution was made from dry powder with original particle sizes ranging from 0-50 µm. The mixed solution was left to stabilize overnight so that larger particles would settle, leaving only the smaller particles suspended in the solution. Consequently, the particle size distribution for the Arizona test dust measurements is not known but is expected to be dominated by the smaller particles.

All VSF measurements were performed in the laboratory using the LISST-VSF instrument in its benchtop mode. Each measurement series commenced with 1630 mL of milli-Q water being poured into the instrument sample chamber. After waiting one hour for bubbles to dissipate, a blank measurement was made. Subsequently, the scattering agent was added from a master solution to the sample chamber in predetermined amounts using pipettes for precise concentrations. The milli-Q water used in the blank was also used in the following measurements, yielding an effective subtraction of pure water scattering and other possible optical losses. A series of measurements were made with subsequent additions of scattering agent solution, going from low to very high concentrations. The concentrations used in this study are given in Table 1 (Section 3.1). To minimize uncertainties connected to the particle concentrations, the VSF measurements were scaled by the ratio of theoretical (Mie theory) and measured (LISST-VSF) attenuation,

$$\beta_{\mathrm{corr}}(\theta) = \beta_{\mathrm{uncorr}}(\theta)\frac{c_{\mathrm{Mie}}}{c_{\mathrm{LISST}}}, \tag{4}$$

as the LISST-VSF attenuation measurements have been shown to have high accuracy above ~0.1 m^{-1} and are virtually unaffected by multiple scattering [1,23]. This correction was only applied to the polystyrene beads measurements due to the lack of a theoretical attenuation coefficient for the Arizona test dust samples.

2.2. Monte Carlo simulation

Here, we present a Monte Carlo simulation based on the random sampling of variables. The variables are expressed as probability distributions, so that the value of a certain variable can be sampled by generating a random number $\xi \in [0, \quad 1]$. The geometry of the simulation is presented in Fig. 1. The simulation boundary is defined by a cylinder with a length stretching

Optics EXPRESS

from $z = 0$ cm to $z = 15$ cm, and a radius of 2 cm. The light source is positioned at the center of the top end of the cylinder ($z = 15$ cm) and the photons are emitted along the z-axis towards the bottom end of the cylinder. The eyeball detector is positioned at the cylinder wall at $z = 5$ cm, so that the distance from the light source is 10 cm along the z-axis. The ring detectors are positioned at $z = 0$ cm and covers the entire bottom area of the cylinder. The cylinder wall act as a perfect absorber, so that all photons crossing the cylinder boundaries are immediately eliminated. It is desirable to minimize the computation time of the simulation without compromising the precision. Thus, the cylinder radius of 2 cm is chosen, as it is considered statistically highly unlikely for a photon that has crossed this cylinder wall boundary to be scattered back into the sample volume and into a detector.

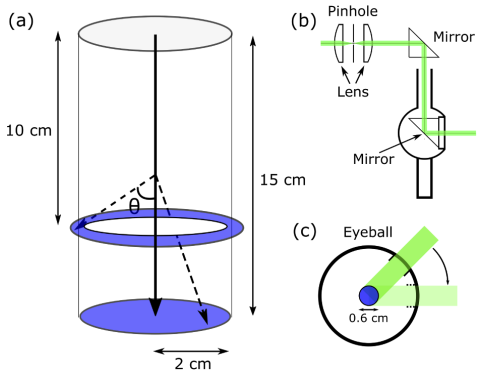

Fig. 1. Illustration of the simulation geometry. (a) The cylindrical sample volume with ring detector in the bottom end of the cylinder and eyeball detector as a ring torus at z=5 cm, both colored in blue. Forward scattered light collected by the blue disk (collecting lens) are sorted by their angle of incidence onto the disk. In the instrument, all photons arriving at a particular angle of incidence are directed to a corresponding ring detector. (b) Eyeball detector principle in the LISST-VSF instrument. (c) Eyeball detection principle in the simulation with the detection sphere colored in blue.

The individual photon's trajectory through the sample volume is calculated following the steps outlined in the flow chart presented in Fig. 2. The distance between photon-particle interaction will here be referred to as step size, where the step size distribution can be derived from the attenuation of a beam,

$$I = I_0 e^{-cl}. \tag{5}$$

Here, I_0 is the initial intensity of the beam, c is the attenuation coefficient, l is the distance over which the beam is attenuated, and I is the intensity after traveling the distance l. A random step size can then be sampled from the distribution where the sampled step size s can be calculated as

$$s = \frac{-\ln(\xi)}{c}. \tag{6}$$

The new position of the photon is calculated based on the traveling direction and the sampled step size, and the scattering angle distribution is obtained from the scattering probability density

Optics EXPRESS

function (PDF) of the sample in question. The scattering PDF is given as

$$f(\theta) = 2\pi p(\theta) \sin\theta \tag{7}$$

and integration yields

$$\int_0^\pi f(\theta) d\theta = 2\pi \int_0^\pi p(\theta) \sin\theta d\theta = 1, \tag{8}$$

where $p(\theta)$ is the scattering phase function. The scattering angle θ is randomly sampled by generating a new $\xi \in [0, \quad 1]$. The sampled scattering angle θ_{sc} is then found from the relation

$$2\pi \int_0^{\theta_{sc}} p(\theta) \sin\theta d\theta = \xi. \tag{9}$$

A lookup table containing the values of θ_{sc} for different ξ is made prior to the simulation. The azimuthal angle ϕ is sampled from a uniform distribution from 0 to 2π. When performing a simulation for the 508 nm spherical polystyrene beads, a theoretical phase function computed from Mie theory is used as input. For Arizona test dust, the input phase function is obtained from a LISST-VSF measurement at low concentration, so that one can assume that the single scattering approximation is valid.

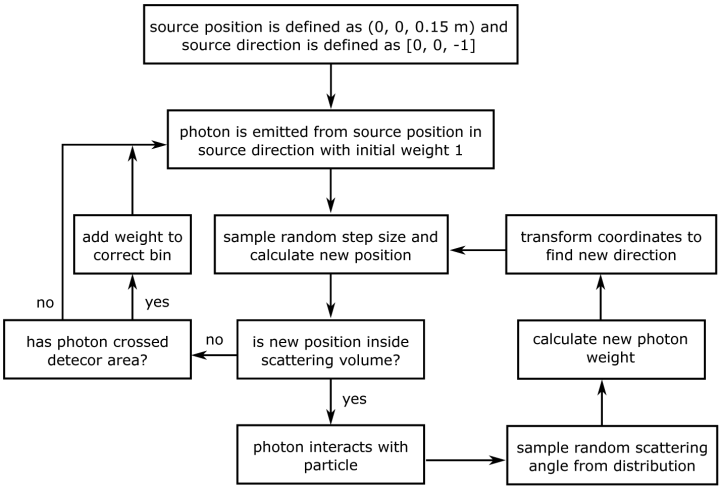

Fig. 2. Flowchart showing the steps in the Monte Carlo algorithm.

In order to account for absorption by the sample volume, we give the photons an initial weight $w = 1$, and at each scattering event, the weight is reduced according to

$$w_{i+1} = w_i \frac{b}{c} \tag{10}$$

where w_{i+1} is the weight after scattering, w_i is the weight before scattering, c is the attenuation coefficient, and b is the scattering coefficient. This avoids having to initiate a new photon every time one is absorbed. As the LISST-VSF measurement is calibrated so that the absorption and scattering from the water itself is zero, there is no need to implement the attenuation from water in the simulation.

Optics EXPRESS

2.2.1. Simulation of detectors

Ring detectors In order to detect the scattered intensity in the range $0.09 - 14.4°$, the LISST-VSF uses a set of ring detectors, covering different parts of the range. A lens is positioned above the ring detectors, focusing the light at the different rings depending on the angle of incidence. This detection scheme is simulated by simply defining a flat circle surface, corresponding to the lens, and registering the angle of incidence onto this surface, as illustrated in Fig. 1(a) . The registered photons are then binned according to their angle of incidence, matching the bin sizes of the LISST-VSF.

Eyeball detector The eyeball detection scheme is illustrated in Fig. 1(b). The light passes through an opening in the eyeball, leading to a set of reflecting mirrors. The mirrors guide the light to a spatial filter consisting of two lenses and a pinhole, resulting in an acceptance angle of $2\theta_a = 0.9°$. The total power reaching the detector is thus dependent on the opening of the eyeball and the acceptance angle. While the LISST-VSF eyeball rotates in order to measure the angle dependency of the scattered light, the simulation can detect scattered light from all angles θ simultaneously. This is achieved by defining a sphere positioned at the center of the eyeball, where passing through this sphere is the equivalent of passing through the eyeball opening and being within the acceptance angle of the spatial filter. The principle behind this design is illustrated in Fig. 1(c), and while this is a major simplification of the eyeball detection system, the results presented in this paper demonstrate that it works very well. A sphere radius of 3 mm was found to give a good fit to the experimental results. The simulation also allows us to extend this eyeball design around the cylinder, creating a ring torus (see Fig. 1(a)). This is the equivalent of having eyeball detectors covering the entire circumference of the cylinder, as opposed to the LISST-VSF which only has one eyeball detector at a single location. Thus, the simulation can detect scattered photons for all azimuth angles.

2.2.2. Processing simulated data

To obtain a VSF from the MC simulation, the raw data, i.e. number of photons detected at each angle, must be processed. The VSF is calculated based on the method presented in [25]. The number of scattered photons detected by the ring detectors can be calculated for each bin (ring) as

$$N_i = 2\pi e^{-cl} x N_0 \beta_i(\theta) \phi (\cos\theta_{i,l} - \cos\theta_{i,h}), \tag{11}$$

where e^{-cl} is the attenuation factor and l is the total distance traveled from the laser injection point to the detector, assuming single scattering. The factor x is the length of the laser beam contributing to the signal (see Fig. 3), N_0 is the total number of simulated photons, equivalent to the laser power, $\beta_i(\theta)$ is the VSF for bin i, and $\theta_{i,h}$ and $\theta_{i,l}$ is the high and low limit of the bin, respectively. The symbol ϕ does here refer to the fraction of a full circle covered by the ring detectors. For practical reasons, the ring detectors in the LISST-VSF only cover $1/6$ of the full circle, meaning that $\phi = 1/6$. In the simulation however, we can use the full circle so that $\phi = 1$. Rearranging Eq. (11), we get the following expression for the VSF

$$\beta_i(\theta) = \frac{N_i}{N_0} \frac{e^{cl}}{2\pi\phi x(\cos\theta_{i,l} - \cos\theta_{i,h})}. \tag{12}$$

Here, the factor e^{cl} accounts for photons lost along the path to the detector. For the ring detector, l does not vary much with detection angle, and is set to $l = L = 0.15$ m. An identical correction for attenuation is done for the LISST-VSF measurements. For angles smaller than $\arctan(r/L)$, where r is the radius of the lens focusing light at the ring detectors, the entire beam contributes to the signal, so that $x = L = 0.15$ m. For angles larger than this, only a fraction of

Optics EXPRESS

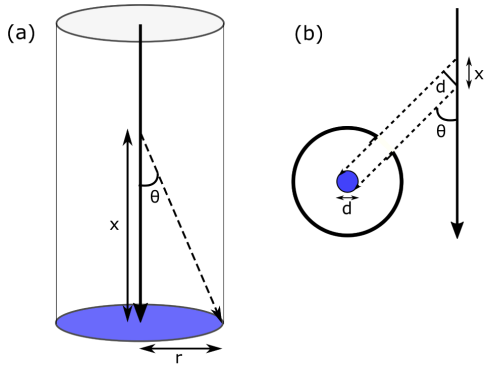

Fig. 3. Geometric illustration of the contributing length x of the laser beam for (a) the ring detector and (b) the eyeball detector.

the beam contributes as the photons scattered from the top end of the beam does not hit the lens (see Fig. 3(a)). In this case, x is calculated as $x = r/\tan\theta_i$, such that Eq. (12) becomes

$$\beta_i(\theta) = \frac{N_i}{N_0} \frac{e^{cl} \tan\theta_i}{2\pi\phi r(\cos\theta_{i,l} - \cos\theta_{i,h})}, \tag{13}$$

where θ_i is the mean angle for bin i.

The same approach is used to calculate the VSF from the signal obtained by the eyeball detector. In this case, only a small fraction of the beam contributes to the signal, where x is calculated as $x = d/\sin\theta_i$ (see Fig. 3(b)). Here, $d = 6$ mm is the diameter of the eyeball detection sphere defined in Section 2.2.1.2. The VSF for the eyeball detector can then be calculated for bin i as,

$$\beta_i(\theta) = \frac{N_i}{N_0} \frac{e^{cl} \sin\theta_i}{2\pi\phi d(\cos\theta_{i,l} - \cos\theta_{i,h})}. \tag{14}$$

Here, the path length l is dependent on the angle of detection, and can be calculated as

$$l = l_{\text{eye}} - R\cot\theta + \frac{R}{\sin\theta}, \tag{15}$$

where $l_{\text{eye}} = 0.10$ m is the distance from the top end of the cylinder to the eyeball detector along the z-axis, and R is the distance from the laser beam to the eyeball detector measured at a right angle. Again, we can set $\phi = 1$ due to the full circle coverage of the eye detector.

In order to ensure a similar signal for all samples, the simulation is set to run until 10^7 photons are detected by the eye detector, i.e. $\sum N_{i,eye} = 10^7$. The same limit is set for the ring detector, at which the detector stops counting detected photons. The final signal in the ring detector is then calculated as $N_i = N \cdot N_{i,\text{stop}}/N_{\text{stop}}$, where N_i is the adjusted number of photons in bin i, N is the total number of photons simulated, $N_{i,\text{stop}}$ is the number of photons in bin i at which the limit of 10^7 detected photons is reached, and N_{stop} is the total number of simulated photons when the number of detected photons reaches 10^7.

3. Results and discussion

3.1. Monte Carlo simulations compared with experimental results

The results obtained using the polystyrene beads as the scattering agent are presented in Fig. 4. The simulation input parameters b and c are given in Table 1, along with the corresponding

Optics EXPRESS

experimental particle concentrations C and optical depths τ. Measurements by the LISST-VSF are compared with single-scattering theoretical values computed from Mie theory, and simulation results from the Monte Carlo algorithm. At low concentrations, the theoretical, measured, and simulated VSF are in overall good agreement. For increasing particle concentrations, the measured VSF increasingly deviates from the theoretical VSF, both in magnitude and shape. For the highest bead concentration, the amount of measured scattering at large angles is more than an order of magnitude larger than predicted from single scattering, and the oscillating features of the 508 nm bead phase function have vanished. Another prominent feature at higher concentrations is the discontinuity at 15°. This angle is the boundary between where the VSF is measured by the ring detector and where the eye detector is used.

Table 1. Simulation input parameters b and c, and corresponding particle concentrations C (or relative concentrations C_r) and optical depths τ for the 508 nm beads and Arizona test dust samples.

	508 nm beads				Arizona test dust			
Sample No.	C [μL/L]	b [m^{-1}]	c [m^{-1}]	τ	C_r	b [m^{-1}]	c [m^{-1}]	τ
1	0.050	0.182	0.183	0.027	1	0.084	0.088	0.013
2	0.10	0.365	0.365	0.055	3	0.253	0.263	0.039
3	0.21	0.766	0.767	0.12	10	0.842	0.876	0.13
4	0.51	1.86	1.86	0.28	30	2.53	2.63	0.39
5	1.0	3.68	3.69	0.55	100	8.41	8.75	1.3
6	2.0	7.43	7.44	1.1	296	24.9	25.9	3.9
7	3.0	11.1	11.1	1.7	-	-	-	-
8	4.0	14.7	14.7	2.2	-	-	-	-
9	6.0	21.9	22.0	3.3	-	-	-	-
10	8.9	32.8	32.9	4.9	-	-	-	-

By contrast, the simulation predicts a VSF with closer resemblance to the LISST-VSF measurements, even at the highest scattering coefficient $b \approx 33$ m^{-1}. Some minor deviations can be seen. At small scattering angles (<7°) and low concentrations, the LISST-VSF measured enhanced forward scattering relative to theory, whereas the simulated and theoretical curves are almost indistinguishable. The reason is not entirely clear, but is likely attributed to bead flocs, i.e., beads clumped together and effectively acting as larger particles [1,26]. Increased forward scattering for larger particles is consistent with Mie theory. The bead flocs may then deflocculate (break apart) when mixing with water during the experiments, resulting in the spike gradually vanishing. The concentrated polystyrene bead solution was also shaken before each addition of beads to the LISST-VSF sample chamber, so later additions of beads were more thoroughly shaken. Thus, the later additions of beads may have contained less flocculated beads, adding to the reduction of the forward spike. There are other possible explanations for the observed spike in forward scattered intensity. Local heating of the sample volume in the vicinity of the laser beam might cause a temperature gradient to arise, which in turn results in density gradients and deviations in the refractive index [27]. Such deviations can cause scattering of light, and thereby an increase in the measured forward scattering. However, if this were the case, one would also expect to measure increased attenuation, which is not observed in our experiments. Thus, this explanation is considered unlikely. The discontinuity at 15° is also apparent in the simulated VSF, but appears larger in the experimental data for sample Nos. 6-9. After inspection of the experimental raw data, we find that this is due to saturation of the eye detector sensor at the smallest angles (15 – 25°). At the highest concentration, the saturation is quenched by increased attenuation of the light (which is compensated for in the data processing), hence the improved agreement between the experiment and simulation for sample No. 10. From 140° and

Optics EXPRESS

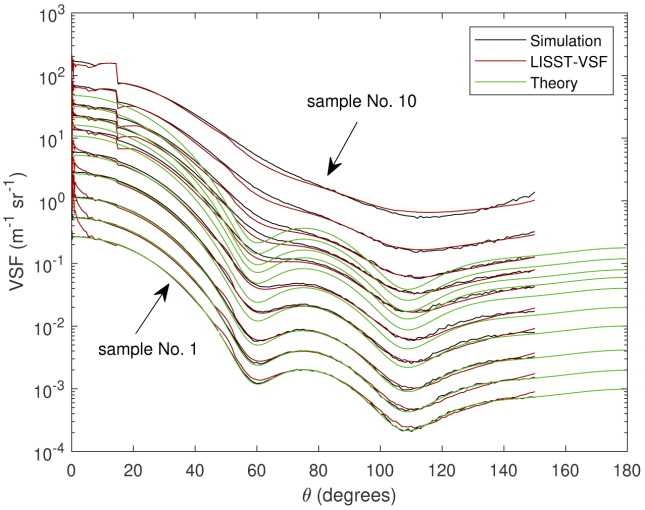

Fig. 4. Volume scattering functions measured by the LISST-VSF (red line) for varying 508 nm bead concentrations (see Table 1), are plotted with theoretically predicted results assuming single scattering (green line), and results from the Monte Carlo simulations of the LISST-VSF measurements (black line). The abrupt jump in the VSF seen at 15° is due to the longer optical path for the ring detector than for the eyeball detector.

onward, the simulated backscattering increases relative to the experimental backscattering, with increasing particle concentration. This is most likely due to the simulated eyeball detection being a simplified version of the instrument eyeball detection. Consequently, the two different eyeball detection schemes might probe the sample volume slightly different, resulting in deviations in the VSF when the turbidity becomes very high.

The Arizona test dust results are presented in Fig. 5. The simulation input parameters b and c are given in Table 1, with the corresponding experimental relative particle concentrations (specific particle concentration is not known) and optical depths τ. Here, the theoretical VSF is acquired from a low concentration sample and scaled according to the relative concentrations C_r presented in Table 1. The results follow the same trend as for the polystyrene beads, showing clear deviations between the predicted values for single scattering and the measurements at high concentrations, and close agreement between simulated and experimental results. A clear difference in the two phase functions, are the oscillatory features seen in the 508 nm beads phase function (see Fig. 4), as opposed to Arizona test dust phase function, where such features are not present. The oscillatory features appear due to constructive and destructive interference of the electromagnetic field scattered by individual particles [12,28]. The angular positions of constructive and destructive interference are dependent on particle size and refractive index, such that the monodisperse 508 nm beads samples will display oscillations in the phase function, while the polydisperse Arizona test dust samples will not.

Results seen in Figs. 4 and 5 validate the developed Monte Carlo simulation of the LISST-VSF, and offer robust explanations of instrument artifacts seen in both laboratory and field measurements. Multiple scattering causes both erroneously large detected scattering and an altered phase function. The discontinuity at 15° arises due to the sudden decrease in path length

Optics EXPRESS

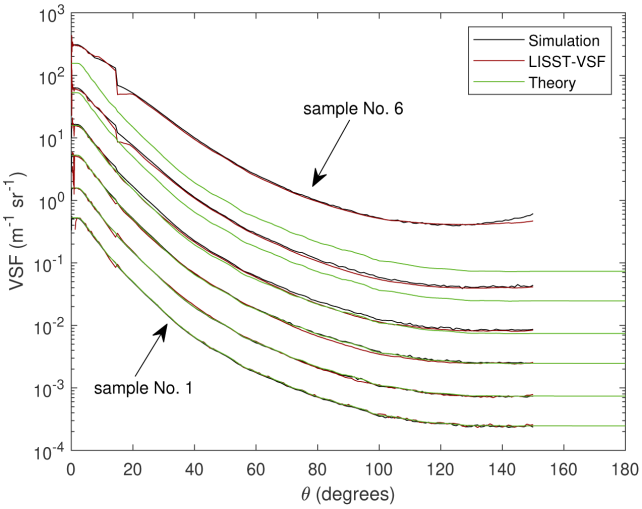

Fig. 5. Volume scattering functions measured by the LISST-VSF (red line) for varying Arizona test dust concentrations (see Table 1), are plotted with theoretically predicted results assuming single scattering (green line), and results from the Monte Carlo simulations of the LISST-VSF measurements (black line).

going from ring detectors to eyeball detector (see Fig. 1). This decrease in path length has a double effect on the measured VSF. Firstly, longer path lengths result in more multiple scattered photons. Secondly, the attenuation factor e^{cl} is larger for the ring detectors (Eq. (12)) than for the eye detector (Eq. (14)) in the forward direction. Thus, the increase in number of detected photons N_i due to multiple scattering contributes more to the absolute error in the VSF measured by the ring detector compared to the one measured by the eyeball detector. However, this does not impact the relative error. Saturation of the eyeball detector may also be a significant factor in waters dominated by scattering processes.

3.2. Error analysis

3.2.1. Convergence test

A convergence test has been performed for the 508 nm beads sample No. 1 (see Table 1 for details about the sample). Convergence is tested for the scattering coefficient b, calculated from Eq. (2), and the results are presented in Fig. 6. For the test, the simulation is set to run until a specific number of photons N_{eye} are detected by the eye detector, starting at 1221 photons and doubling for each simulation up til 10^7 (same as for the results presented in Figs. 4 and 5). The simulation is repeated 10 times for each N_{eye}, providing enough data to calculate a coefficient of variance (CV), also known as relative standard deviation. The CV is calculated as $c_v = \frac{\sigma}{\mu} \times 100\%$, where σ is the standard deviation of the mean μ. The calculated scattering coefficients b from the simulations are presented in Fig. 6(a), and the CV of b is plotted in Fig. 6(b), in addition to a theoretical fit calculated as $c_v = C/\sqrt{N_{\text{eye}}}$, where C is a proportionality constant.

From Fig. 6, one can see that b converges relatively fast. In fact, the CV is below 1% when N_{eye} is larger than 10^4. While the simulated VSF obtained from a signal of 10^4 detected photons would be very noisy, the scattering coefficient can be accurately predicted as it is obtained from

Optics EXPRESS

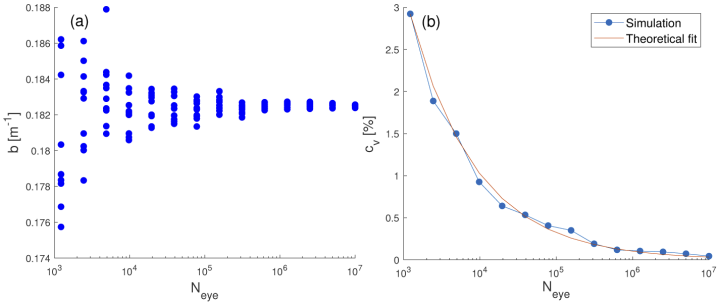

Fig. 6. (a) Scattering coefficient b calculated from simulation for increasing number of detected photons N_{eye} for sample No. 1. (b) Coefficient of variance c_v for increasing number of detected photons N_{eye} and a theoretical fit calculated as $c_v = C/\sqrt{N_{\text{eye}}}$, with $C = 102$.

integrating the VSF over the angels $0.09^\circ - 150^\circ$ (see Eq. (2)). At 10^7 detected photons, the CV is approximately 0.04%, giving a very high precision. A similar convergence test was performed for sample number 5, and no significant difference in CV was found. Thus, it was concluded that the optical depth has little to no influence on the CV for the scattering coefficient.

3.2.2. Multiple scattering error

The validity of the single-scattering approximation may readily be assessed by comparing the Monte Carlo simulation results with the predicted results from assuming single-scattering. For the error analysis, the theoretical scattering coefficients and phase functions are calculated from the VSFs labeled "Theory" in Fig. 4 and 5. The theoretical, experimental, and simulated scattering coefficients, denoted b_t, b_e, and b_s, respectively, are given in Table 2. The scattering coefficients are calculated by integrating from 0.09° to 150°, so that theory, experiment and simulation are considered over the same angular spectrum. Thus, the theoretical scattering coefficient is not the same as the input scattering coefficient given in Table 1.

Table 2. Theoretical b_t, experimental b_e, and simulated b_s scattering coefficients for the 508 nm beads and Arizona test dust samples. The scattering coefficients are calculated by integrating from 0.09° to 150°.

	508 nm beads			Arizona test dust		
Sample No.	b_t [m^{-1}]	b_e [m^{-1}]	b_s [m^{-1}]	b_t [m^{-1}]	b_e [m^{-1}]	b_s [m^{-1}]
1	0.182	0.186	0.182	0.086	0.084	0.084
2	0.363	0.360	0.367	0.258	0.257	0.254
3	0.763	0.750	0.781	0.858	0.837	0.871
4	1.85	1.84	1.97	2.58	2.64	2.83
5	3.67	3.83	4.14	8.57	11.7	12.5
6	7.40	8.56	9.54	25.4	85.2	89.0
7	11.0	14.5	16.25	-	-	-
8	14.6	22.5	24.8	-	-	-
9	21.8	47.2	49.8	-	-	-
10	32.7	119.7	123.7	-	-	-

Optics EXPRESS

From Table 2, one can see that a difference in the theoretical single-scattering scattering coefficient b_t and the experimental scattering coefficient b_e can be quite dramatic for high particle concentrations. As the LISST-VSF attenuation measurements have been shown to have a high accuracy and are virtually unaffected by multiple scattering [1,23], one would see a similar dramatic difference in the measured absorption a, calculated as $a = c - b$. In fact, in some cases, when $b>c$, one would get a negative absorption coefficient. The relative errors due to multiple scattering for the scattering coefficient and phase function are plotted in Fig. 7. Three different relative errors are presented. The desired output from a LISST-VSF measurement is the single-scattering phase function and scattering coefficient, and are thus taken as true values when computing the error in the LISST-VSF measurements. This is also the case when comparing simulated results to single scattering. When comparing simulation and experiment, the experimental results are taken as true values, as the goal of the simulation is to reproduce the LISST-VSF measurements. The scattering coefficient relative error e_b is computed from

$$e_b = \frac{|b_{\text{meas}} - b_{\text{true}}|}{b_{\text{true}}} \times 100\%. \tag{16}$$

For the phase function, where the error may vary significantly with angle, the mean relative error e_p is calculated, using

$$e_p = \frac{1}{n} \sum_{i}^{n} \frac{|p_{\text{meas}}(\theta_i) - p_{\text{true}}(\theta_i)|}{p_{\text{true}}(\theta_i)} \times 100\%. \tag{17}$$

The measured phase function is measured at log-spaced angles at $<15^\circ$. To avoid bias towards forward scattering effects, the data was interpolated to evenly spaced angles.

The scattering coefficient and phase function errors for the 508 nm beads samples are plotted in Figs. 7(a) and (b), respectively, and the errors for the Arizona test dust samples are plotted in Figs. 7(c) and (d). The error is plotted against the dimensionless parameter τ. As seen in Fig. 7, the simulation predicts a scattering coefficient and phase function close to those derived from the LISST-VSF measurements. For the scattering coefficients, the relative error is always $<13\%$, and $<10\%$ for the phase function. As discussed in section 3.1, a large portion of the error in the simulation relative to the LISST-VSF measurements can be attributed to measurement errors in the experimental results. Comparing simulation and experiment to the single-scattering theoretical values, the errors in the scattering coefficients arise solely due to the added intensity from multiple scattered photons. The phase function, however, is normalized so that the addition of multiple scattered photons does not alone cause the error. Rather, it is the angular distribution of multiple scattered photons that causes the error. Due to the phase functions being highly forward peaked, the photons scattered more than once tend to be distributed to angles close to the original scattering angle. As a consequence, phase functions that varies rapidly with scattering angle is more subjected to multiple scattering errors in the phase function, as the relative increase in signal becomes large when a low signal part of the spectrum neighbors a high signal part. This is especially evident looking at the relative error in the phase functions (Figs. 7(b) and (d)), where the error is significantly larger for the 508 nm beads than for the Arizona test dust. This is due to the low signal dips in the 508 nm beads phase function, approximately at 60° and 110°, gaining a relatively large quantity of multiple scattered photons from the neighboring higher signal parts. Ultimately, as the optical depth approaches infinity, the angular distribution of photons becomes completely random, resulting in a uniform VSF.

Comparing simulation to theory, the error in the scattering coefficient is calculated to be $e_b = 0.4\%$ for the 508 nm at an optical depth of $\tau = 0.027$. The error increases with optical depth reaching $e_b = 29\%$ for $\tau = 1.1$, and finally $e_b = 278\%$ for $\tau = 4.9$. As the CV was calculated to be $c_v = 0.04\%$, the numerical uncertainty in the simulated results can be considered

Optics EXPRESS

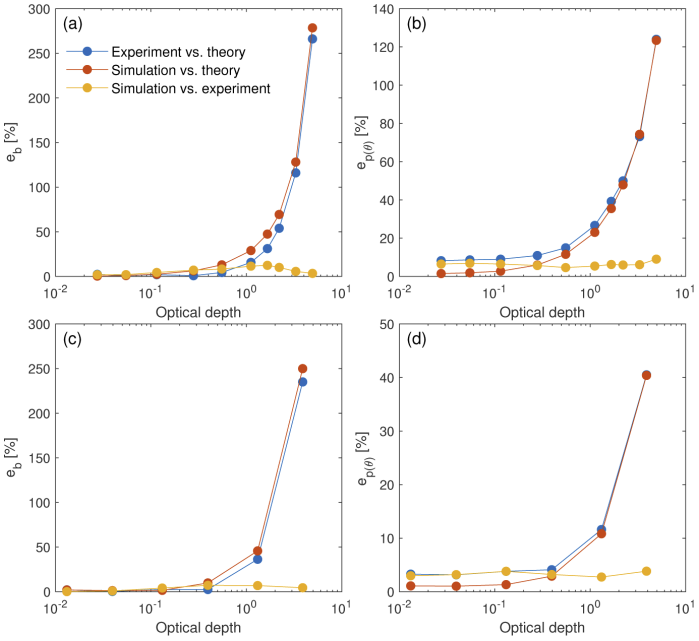

Fig. 7. Relative errors calculated from comparing experiment vs. theory (blue), simulation vs. theory (red) and simulation vs. experiment (yellow). Relative error of (a) the scattering coefficient e_b and (b) the phase function $e_{p(\theta)}$ for the 508 nm beads samples. Relative error of (c) the scattering coefficient e_b and (d) the phase function $e_{p(\theta)}$ for the Arizona test dust samples. The errors are plotted against the dimensionless optical depth τ.

negligible, except for at the smallest optical depths. The error in the phase function is, for the most part, significantly smaller than for the scattering coefficient. The exceptions are for the three smallest optical depths. This is probably due noise being the dominating source of error when the difference between simulated and theoretical VSF becomes very small. The scattering coefficient is not subjected to noise in the same way, as the scattering coefficient is calculated by integrating the VSF over the entire angular range (see Eq. (2)). At the smallest optical depth $\tau = 0.027$, the error is calculated to be $e_p = 2\%$, increasing to $e_p = 123\%$ for the largest optical depth of $\tau = 4.9$. For the Arizona test dust, the error in the scattering coefficient goes from $e_b = 2\%$ to $e_b = 250\%$ when the optical depth goes from $\tau = 0.013$ to $\tau = 3.9$. The error in the phase function goes from $e_p = 1\%$ to $e_p = 40\%$ for the same values of τ. An overview of the optical depths at which the scattering coefficient and phase function reach specific relative errors is presented in Table 3. The values presented are calculated by linear interpolation of the results labeled as "Simulation vs. theory" in Fig. 7.

As seen in the results presented here, a 15 cm path length is sufficient to cause considerable errors in the measured VSF and its derived inherent properties, here represented by the scattering coefficient and phase function. Thus, it becomes interesting to compare the LISST-VSF with other optical equipment and measurements. For instance, the historical Petzold measurements are often used as a reference to *in situ* data [29]. For these measurements, two different instruments

Optics EXPRESS

Table 3. Optical depths at which the 508 nm beads and Arizona test dust reach specific levels of relative error in the scattering coefficient *b* and phase function $\tilde{p}(\theta)$.

	Scattering coefficient		Phase function	
Relative error	508 nm beads	Arizona test dust	508 nm beads	Arizona test dust
1%	0.055	NaN	NaN	NaN
2%	0.10	0.15	0.062	0.24
5%	0.23	0.24	0.23	0.64
10%	0.43	0.40	0.48	1.2
30%	1.1	0.91	1.4	3.0
50%	1.7	1.4	2.3	NaN
100%	2.8	2.0	4.1	NaN

were used, depending on the scattering angle. The instrument used to measure the low angle scattering (approximately $0.06^\circ - 0.4^\circ$) had a path length of 50 cm, while the range $10^\circ - 170^\circ$ was measured at a path length of 18.8 cm. The VSF was measured for a range of different ocean waters, with an attenuation coefficient peaking at 2.190, which translates to an optical depth of 1.1, for the longest path length of 50 cm. From Table 3, one can see that this results in a relative error of approximately 30% in the measured scattering coefficient, assuming the equipment is subjected to multiple scattering similarly to the LISST-VSF.

4. Summary and conclusion

We have demonstrated that the errors originating from multiple scattering in the LISST-VSF measurements can be significant for large optical depths, which largely explains *in situ* LISST-VSF measurement artifacts seen in turbid waters measurements. The developed Monte Carlo algorithm has been validated by employing 508 nm polystyrene beads and Arizona test dust as scattering agents, and has proven to be an accurate tool for analyzing errors in both the measured phase function and scattering coefficient. A convergence test was performed for the MC algorithm, from which we found a coefficient of variance of $c_v \leq 0.04\%$ for the scattering coefficient. Comparing simulation to experiment, the relative error is <13% for the scattering coefficient and <10% for the phase function, for the concentrations and scattering agents tested in this study. A large fraction of these errors can be attributed to errors in the LISST-VSF measurements, caused by particle flocculation, or saturation of the eyeball detector.

Comparing the simulated VSF to the single-scattering VSF, the relative error in the scattering coefficient reaches 10% for an optical depth of $\tau \approx 0.4$ for both scattering agents, while an error of 100% is reached for $\tau \approx 2.8$ and $\tau \approx 2.0$ for the 508 nm beads and Arizona test dust, respectively. The error in the phase function was found to be significantly larger for the 508 nm beads than for the Arizona test dust. A 10% error is found for $\tau \approx 0.48$ for the 508 nm beads, reaching 100% for $\tau \approx 4.1$. The Arizona test dust, however, does not reach a 10% error before $\tau \approx 1.2$ and does not go above 50% for the concentrations investigated. The large difference in error between the two phase functions is attributed to the finer details in the 508 nm beads phase function. Thus, we have shown that errors in both the scattering coefficient and phase function are dependent on the phase function, but to what extent is still not fully explored. Furthermore, both sample sets investigated in this paper are heavily dominated by scattering, increasing the ratio of absorption to scattering is expected to influence the error when plotted against optical depth.

Acknowledgments. We would like to thank one of our anonymous reviewers who brought to our attention the possibility of multiple scattering errors in the Petzold measurements.

Optics EXPRESS

Disclosures. The authors declare no conflicts of interest.

References

1. H. Sandven, A. S. Kristoffersen, Y.-C. Chen, and B. Hamre, "In situ measurements of the volume scattering function with LISST-VSF and LISST-200X in extreme environments: evaluation of instrument calibration and validity," Opt. Express **28**(25), 37373–37396 (2020).
2. Y. C. Agrawal and H. C. Pottsmith, "Instruments for particle size and settling velocity observations in sediment transport," Mar. Geol. **168**(1-4), 89–114 (2000).
3. C. B. Mouw, S. Greb, D. Aurin, P. M. DiGiacomo, Z. Lee, M. Twardowski, C. Binding, C. Hu, R. Ma, T. Moore, W. Moses, and S. E. Craig, "Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions," Remote. Sens. Environ. **160**, 15–30 (2015).
4. M. Chami, E. B. Shybanov, G. A. Khomenko, M. E.-G. Lee, O. V. Martynov, and G. K. Korotaev, "Spectral variation of the volume scattering function measured over the full range of scattering angles in a coastal environment," Appl. Opt. **45**(15), 3605–3619 (2006).
5. M. Soja-Woźniak, M. Baird, T. Schroeder, Y. Qin, L. Clementson, B. Baker, D. Boadle, V. Brando, and A. D. L. Steven, "Particulate backscattering ratio as an indicator of changing particle composition in coastal waters: Observations from Great Barrier Reef waters," J. Geophys. Res.: Oceans **124**(8), 5485–5502 (2019).
6. V. G. Harharasudhan and P. Shanmugam, "Modelling the particulate backscattering coefficients of turbid and productive coastal waters," Ocean Sci. J. **54**(2), 147–164 (2019).
7. C. Nima, Ø. Frette, B. Hamre, S. R. Erga, Y.-C. Chen, L. Zhao, K. Sørensen, M. Norli, K. Stamnes, and J. J. Stamnes, "Absorption properties of high-latitude Norwegian coastal water: The impact of CDOM and particulate matter," Estuarine, Coastal Shelf Sci. **178**, 158–167 (2016).
8. M. Chami, D. McKee, E. Leymarie, and G. Khomenko, "Influence of the angular shape of the volume-scattering function and multiple scattering on remote sensing reflectance," Appl. Opt. **45**(36), 9210–9220 (2006).
9. W. H. Slade, Y. C. Agrawal, and O. A. Mikkelsen, "Comparison of measured and theoretical scattering and polarization properties of narrow size range irregular sediment particles," in *2013 OCEANS-San Diego*, (IEEE, 2013), pp. 1–6.
10. L. Hu, X. Zhang, Y. Xiong, and M.-X. He, "Calibration of the LISST-VSF to derive the volume scattering functions in clear waters," Opt. Express **27**(16), A1188–A1206 (2019).
11. T. Harmel, M. Hieronymi, W. Slade, R. Röttgers, F. Roullier, and M. Chami, "Laboratory experiments for inter-comparison of three volume scattering meters to measure angular scattering properties of hydrosols," Opt. Express **24**(2), A234–A256 (2016).
12. D. Koestner, D. Stramski, and R. A. Reynolds, "Measurements of the volume scattering function and the degree of linear polarization of light scattered by contrasting natural assemblages of marine particles," Appl. Sci. **8**(12), 2690 (2018).
13. D. Koestner, D. Stramski, and R. A. Reynolds, "Polarized light scattering measurements as a means to characterize particle size and composition of natural assemblages of marine particles," Appl. Opt. **59**(27), 8314–8334 (2020).
14. M. Jonasz and G. Fournier, *Light Scattering by Particles in Water: Theoretical and Experimental Foundations* (Elsevier, 2007).
15. H. C. van de Hulst, *Light Scattering by Small Particles* (Courier Corporation, 1981).
16. B. M. Agrawal and M. P. Mengüç, "Forward and inverse analysis of single and multiple scattering of collimated radiation in an axisymmetric system," Int. J. Heat Mass Transfer **34**(3), 633–647 (1991).
17. S.-K. Chae and H. S. Lee, "Determination of radiative transport properties of particle suspensions by a single-scattering experiment," Aerosol Sci. Technol. **18**(4), 389–402 (1993).
18. D. McKee, J. Piskozub, and I. Brown, "Scattering error corrections for in situ absorption and attenuation measurements," Opt. Express **16**(24), 19480–19492 (2008).
19. D. McKee, J. Piskozub, R. Röttgers, and R. A. Reynolds, "Evaluation and improvement of an iterative scattering correction scheme for in situ absorption and attenuation measurements," J. Atmospheric Ocean. Technol. **30**(7), 1527–1541 (2013).
20. D. Doxaran, E. Leymarie, B. Nechad, A. Dogliotti, K. Ruddick, P. Gernez, and E. Knaeps, "Improved correction methods for field measurements of particulate light backscattering in turbid waters," Opt. Express **24**(4), 3615–3637 (2016).
21. S. Vadakke-Chanat, P. Shanmugam, and B. Sundarabalan, "Monte Carlo simulations of the backscattering measurements for associated uncertainty," Opt. Express **26**(16), 21258–21270 (2018).
22. D. Stramski and J. Piskozub, "Estimation of scattering error in spectrophotometric measurements of light absorption by aquatic particles from three-dimensional radiative transfer simulations," Appl. Opt. **42**(18), 3634–3646 (2003).
23. J. Piskozub, D. Stramski, E. Terrill, and W. K. Melville, "Influence of forward and multiple light scatter on the measurement of beam attenuation in highly scattering marine environments," Appl. Opt. **43**(24), 4723–4731 (2004).
24. J. T. O. Kirk, "Monte Carlo modeling of the performance of a reflective tube absorption meter," Appl. Opt. **31**(30), 6463–6468 (1992).
25. Y. C. Agrawal, "The optical volume scattering function: Temporal and vertical variability in the water column off the New Jersey coast," Limnol. Oceanogr. **50**(6), 1787–1794 (2005).
26. W. H. Slade, E. Boss, and C. Russo, "Effects of particle aggregation and disaggregation on their inherent optical properties," Opt. Express **19**(9), 7945–7959 (2011).

Optics EXPRESS

27. O. A. Mikkelsen, T. G. Milligan, P. S. Hill, R. J. Chant, C. F. Jago, S. E. Jones, V. Krivtsov, and G. Mitchelson-Jacob, "The influence of schlieren on in situ optical measurements used for particle characterization," Limnol. Oceanogr.: Methods **6**(3), 133–143 (2008).
28. M. J. Berg, C. M. Sorensen, and A. Chakrabarti, "Explanation of the patterns in Mie theory," J. Quant. Spectrosc. Radiat. Transfer **111**(5), 782–794 (2010).
29. T. J. Petzold, "Volume scattering functions for selected ocean waters," Tech. rep., Scripps Institution of Oceanography La Jolla Ca Visibility Lab (1972).

Paper II

Efficient Monte Carlo simulation reveals significant multiple scattering errors in underwater angular scattering measurements

Håvard S. Ugulen, Håkon Sandven, Børge Hamre, Arne S. Kristoffersen, and Camilla Sætre
Optics Express, **30(7)** (2022)

Research Article | Vol. 30, No. 7 / 28 Mar 2022 / Optics Express 10802
Optics EXPRESS

Efficient Monte Carlo simulation reveals significant multiple scattering errors in underwater angular scattering measurements

HÅVARD S. UGULEN,* HÅKON SANDVEN, BØRGE HAMRE, ARNE S. KRISTOFFERSEN, AND CAMILLA SÆTRE

Department of Physics and Technology, University of Bergen, Allegaten 55, 5007 Bergen, Norway
**havard.ugulen@uib.no*

Abstract: Multiple scattering can severely affect the accuracy of optical instrumentation. Variance reduction methods have been implemented to improve a Monte Carlo model developed to simulate volume scattering functions measured by LISST-VSF instruments. The implemented methods can result in more than a tenfold increase in efficiency. The simulation is used to analyze multiple scattering errors for a range of Fournier-Forand (FF) phase functions. Our results demonstrate significant errors in the scattering coefficient, backscattering coefficient and phase function, where multiple scattering errors may only be considered negligible (<10%) for scattering coefficients <1 m^{-1}. The errors depend strongly on the scattering coefficient but also increase when phase functions become more forward-peaked.

1. Introduction

Scattering is a fundamental optical process which largely governs radiative transport processes within the visible spectrum in the ocean and the atmosphere [1]. In optical oceanography, measuring scattering properties of the water column can give valuable information about size and composition of particles submerged in the water mass, such as phytoplankton or colloidal mineral particles [2–9]. Both the scattering and backscattering coefficient have been found to have a strong relationship with suspended particulate matter concentration [2–4]. These properties tell us how much light is scattered in the backward direction (backscattering) and in total (scattering). The ratio between the backscattering and scattering coefficient yields an estimate of the bulk refractive index, which indicates whether inorganic or organic particles are dominating [5]. The full angular distribution of scattering, namely the volume scattering function (VSF), has not been widely measured due to difficulties with designing instrumentation and conducting measurements *in situ*, but may tell us much more about the particulate properties [6,7]. The VSF at angles smaller than 15° is now routinely used for estimating the particle size distribution between approximately 1 and 500 μm [8,9]. The VSF is also an important property to parameterize correctly in radiative transfer modelling of the ocean, which is critical for achieving accurate ocean color remote sensing observations of the oceanic and coastal environments around the globe [10,11]. Ecologically and economically important coastal environments are of considerable interest to measure with *in situ* instrumentation, as these environments are still difficult to observe accurately using satellite observations due to optical and spatio-temporal complexity [12]. Scattering properties in coastal and estuarine environments have a large dynamic range, for instance, the scattering coefficient can span from approximately 0.05 to 30 m^{-1} [2], which puts extra demands on instrumentation.

The LISST-VSF (Sequoia Scientific) is a commercially developed instrument able to measure the volume scattering function (VSF) from 0.1° to 150°. From the VSF one can also derive other

#446045 https://doi.org/10.1364/OE.446045
Journal © 2022 Received 15 Oct 2021; revised 29 Nov 2021; accepted 1 Dec 2021; published 16 Mar 2022

Optics EXPRESS

important inherent optical properties such as the scattering coefficient and phase function. Recent studies [13,14] have demonstrated that, for sufficiently large optical depths, multiple scattering results in significant errors in the VSF measured by the LISST-VSF. This error arises from the assumption that all detected light is only scattered once. The detected signal is corrected for light lost, either to scattering or absorption, along the path from laser to detector by multiplying the detected intensity with the factor e^{cl}. Here, c is the attenuation coefficient and l is the path length from laser to detector. However, not all detected light is scattered only once. Some of the light which is assumed to be lost due to scattering along the path, may still reach a detector. Thus, these photons add to the single-scattering signal, resulting in elevated VSFs, scattering coefficients b, and backscattering coefficients b_b, as well as distorted phase functions.

Monte Carlo simulations have been widely used to investigate measurement errors in optical instruments, as it is relatively easy to implement complex geometries without significant approximations [15–21]. For instance, scattering errors in spectrophotometric absorption meters have been assessed using Monte Carlo methods in several studies including Kirk [15], Stramski and Piskozub [16], and McKee et al. [17]. Piskozub et al. [18] found that multiple scattering has a negligible effect on attenuation measurements. Some recent research efforts with Monte Carlo modelling have been directed towards backscattering sensors. Doxaran et al. [19] simulated two backscattering sensors and found that the ECO-BB instrument (Sea-Bird Sci) is less affected by scattering errors than HydroScat instruments (HOBI Labs), due to a shorter path length (0.02 m, compared to 0.10 m). Vadakke-Chanat et al. [20] also conducted an assessment of the ECO-BB instrument with Monte Carlo modelling and reproduced many of the findings in Doxaran et al. [19], but the effective path length found for the instrument differs between both studies and the user manual. Recently, another study combining direct modeling of light propagation and Monte Carlo simulations of the HydroScat and ECO-BB instruments was published [21]. Here, angular response functions were derived for both instruments, and ECO-BB was found to have an especially large angular full width at half maximum, which can be a significant error source, while multiple scattering due to long path lengths is the main uncertainty for the HydroScat instruments. Monte Carlo modelling has also been applied to other aspects of marine and atmospheric optics, including radiative transfer, underwater optical wireless communication, and LIDAR [22–24].

The effect of multiple scattering on VSF instruments has been studied less than the error for backscattering sensors. In a recent study [14], we presented a Monte Carlo simulation developed to analyze and quantify multiple scattering errors in the LISST-VSF. The simulation was compared to experimental measurements using 508 nm polystyrene beads and Arizona test dust as the scattering agents. The simulation was found to accurately predict experimental measurements for the samples investigated. When comparing simulation to experimental results, the relative error did not exceed 13% for the scattering coefficient and 10% for the phase function for optical depths < 4.9, where a considerable fraction of these errors can be attributed to errors in the experimental measurements, e.g. detector saturation. Comparing simulated and theoretical VSF, i.e the relative measurement error originating from multiple scattering, the error was found to depend on the phase function of the sample volume. The relative error in the scattering coefficient reached 100% at an optical depth of 2.8 and 2.0 for the 508 nm polystyrene beads and Arizona test dust samples, respectively.

While the simulation is proven accurate, it suffers from long computation time due to the fact that most phase functions in natural waters are very forward peaked. Thus, a significant number of photons must be simulated in order to achieve low variance results in the backward direction. In this study, we present three variance reduction methods (VRMs) that have been implemented to increase the efficiency of the simulation. The VRMs are called mean free path transform (MFPTR), detector directional importance sampling (DDIS) and splitting, and are presented in section 2.2. Additionally, the simulation is used to further investigate the effects of multiple scattering on LISST-VSF measurements, especially for realistic phase functions. By

Optics EXPRESS

using a range of different Fournier-Forand (FF) phase functions [25,26], we analyze how the errors originating from multiple scattering depend on the phase function, scattering coefficient, and absorption coefficient of the sample volume. This is presented in section 3.2.

2. Methods

2.1. Monte Carlo simulation

The Monte Carlo simulation presented in [14] is a straight forward algorithm where the individual photons are tracked through the sample volume until they cross the sample volume border or reach a detector. There are two separate detectors in the LISST-VSF instrument. The geometry of the simulation is presented in Fig. 1, and a more detailed description of the simulation geometry is given in [14]. A set of ring detectors are located at the bottom of the sample volume chamber, detecting light at scattering angles 0.01-15°, while scattering angles 15-150° are detected by a rotating eyeball detector. The ring detectors are simulated as a circular surface, where photons crossing this surface are detected and sorted according to their angle of incidence. The eyeball detector is simulated as a ring torus, where crossing the surface area corresponds to detection by the eyeball. The path length between each scattering event is sampled from a distribution derived from the probability density function (PDF) calculated from the beam attenuation $f(l)$, and the scattering angle θ is sampled from a distribution derived from a PDF calculated from the phase function $f(\theta)$.

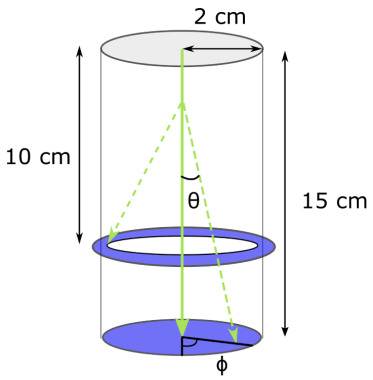

Fig. 1. Simulation geometry. The sample volume is limited to the cylinder. The detector areas are colored in blue, where the disk represents the set of ring detectors, and the ring torus represents the eyeball detector. The solid green arrow represent incident laser beam, while the dashed green arrows represent scattered light.

2.2. Variance reduction methods

The VRMs presented here are based on manipulating the natural scattering behaviour of the photons. Photons are given a statistical weight, initially $w_0 = 1$, which can be adjusted in order to reflect the alteration in the natural scattering behavior. A common example of this is to account for absorption of the sample volume by scaling the weight at each scattering event to reflect the probability of absorption. In this case, the weight after a scattering event is calculated as $w_{i+1} = w_i \cdot b/c$, where w_i is the weight before the scattering event. The first VRM presented in this study is called mean free path transform (MFPTR). Usually, this method is used to stretch the

Optics EXPRESS

path length in order to probe deeper into a medium (e.g. in Mainegra-Hing and Kawrakow [27]). Here, the method is applied to limit the path length such that all photons are scattered within the sample volume, i.e. the transmission is zero. Thus, all photons interact with the sample volume, increasing the simulated signal at the detectors. The second VRM presented, addresses the issue of poor statistics in the backwards direction. The method is based on the work presented in a study by Buras and Mayer [22], and is called detector directional importance sampling (DDIS). In essence, DDIS increases the number of photons scattered towards the detector by giving each photon a probability of being turned directly towards the detector prior to sampling a scattering angle. A consequence of this method is the possibility of photons accumulating weight, which can result in spikes in the detected signal. Thus, the third method presented in this study is implemented to avoid large photon weights and is called splitting [22,27,28]. When the photon weight becomes larger than a threshold value, the photon is split into N identical parts which are propagated individually from the position of splitting. Thus, the weight is distributed to more than one detection angle, avoiding spikes in the signal.

2.2.1. Mean free path transform (MFPTR)

Depending on the scattering coefficient of the sample in question, the fraction of transmitted photons can be large, meaning that a lot of simulation time is spent initiating photons that do not contribute to the detected signal. Here, we have applied MFPTR to the path length distribution before the first scattering event so that all photons are scattered within the sample volume. This is achieved by defining the path length PDF as

$$f_{\text{mod}}(l) = \frac{1}{L} \quad \Rightarrow \quad \int_0^L f_{\text{mod}}(l)dx = 1, \tag{1}$$

where $L = 15$ cm is the length from top to bottom of the cylindrical sample volume and the path length l is sampled from the uniform distribution $[0, L]$. Now, we need to modify the weight of the photon to reflect the change in the path length distribution. The initial weight of the photon is $w_0 = 1$ and the new weight w is calculated as

$$w = w_0 \frac{f(l)}{f_{\text{mod}}(l)} = w_0 \frac{ce^{-cl}}{1/L}, \tag{2}$$

where $f(l) = ce^{-cl}$ is the PDF according to Beer's law. Thus, the new weight is simply the initial weight times the ratio of probabilities of sampling the path length l. After the first scattering event, all path lengths are sampled according to Beer's law.

2.2.2. Detector directional importance sampling (DDIS)

In general, the signal is much higher for the ring detector than for the eyeball detector. Thus, DDIS is only applied to the eyeball detector, i.e. the probability of being scattered towards the eyeball detector is increased and the probability of being scattered towards the ring detector is decreased. Due to the complexity of the detector geometry in this simulation, the implementation of DDIS is not straight forward. First, a random azimuth angle ϕ is sampled from a uniform distribution $[0, 2\pi]$. Once an azimuth angle has been sampled, all possible scattering directions lies in a plane defined by the sampled ϕ. By determining where this plane intersects the eyeball detector (ring torus), we find the angles at which the photon is scattered directly towards the eyeball detector. This is illustrated in Fig. 2(a) for a photon at scattering position P and direction prior to scattering indicated by the red arrow. The angles between the red arrow and the two yellow arrows in Fig. 2(a) indicate the scattering angles $\theta_{\text{rot},1}$ and $\theta_{\text{rot},2}$ giving two possible scattering directions directly towards the detector. Now, instead of letting it follow the natural path, the photon is given a probability ϵ of being redirected towards the eyeball detector through

Optics EXPRESS

either the angle $\theta_{\text{rot},1}$ or $\theta_{\text{rot},2}$ (chosen with equal probability). Then the scattering angle θ_{ddis} is sampled according to a PDF $f_{\text{ddis}}(\theta_{\text{ddis}})$ tailored for the specific geometry in this simulation (see discussion below). Thus, the final scattering angle becomes

$$\theta = \theta_{\text{rot}} + \theta_{\text{ddis}}. \tag{3}$$

By contrast, with a probability of (1-ϵ), the photon is scattered naturally, i.e. according to the phase function of the sample volume. In order to account for this alteration of the natural scattering behaviour, the weight of the photon must be scaled according to the function

$$w_{i+1} = w_i \frac{f(\theta)}{(1-\epsilon)f(\theta) + \epsilon f_{\text{ddis}}(\theta_{\text{ddis}})}, \tag{4}$$

where w_i and w_{i+1} is the weight of the photon before and after scattering, respectively. The numerator of the fraction in Eq. (4) is the PDF corresponding to natural scattering, while the denominator is the PDF corresponding to scattering by the use of DDIS. Thus, the new photon weight w_{i+1} is simply the photon weight before scattering multiplied by the ratio of probabilities between natural scattering and scattering with DDIS. For instance, if the probability of scattering to a given angle θ using DDIS is twice the probability of naturally scattering to the same angle, the weight of the photon after scattering to this angle becomes $w_{i+1} = \frac{1}{2}w_i$. Now we need to find a suitable PDF $f_{\text{ddis}}(\theta_{\text{ddis}})$ for our specific geometry. One could for instance use a PDF calculated from a very forward peaked phase function, so that those photons that are turned towards the detector and scattered according to this PDF would have a very high probability of being scattered towards the detector. However, in order to simplify the VRM, we choose a PDF that is constant within a small interval $\Delta\theta$ in the forward direction and zero everywhere else. Thus, the PDF can be expressed as

$$f_{\text{ddis}}(\theta_{\text{ddis}}) = \frac{1}{\Delta\theta}, \tag{5}$$

so that

$$\int_0^{\Delta\theta} f_{\text{ddis}}(\theta_{\text{ddis}})d\theta = \int_0^{\Delta\theta} \frac{1}{\Delta\theta}d\theta = 1 \tag{6}$$

Next, we want to choose $\Delta\theta$ so that we maximize the number of photons being scattered towards the detector, without compromising the total intensity (weight) detected. Since we know the distance D from the photon position P to the eyeball detector, we can calculate the maximum angle at which the photon can scatter away from the center of the detector and still be within the detector area, see Fig. 2(b),

$$\theta_{\text{max}} = \tan^{-1}\frac{d}{2D}, \tag{7}$$

where d is the diameter of the eyeball detector. Thus, choosing $\Delta\theta = 2\theta_{\text{max}}$, ensures that the photon is scattered towards the detector (see Eq. (3)), where the scattering angle θ_{ddis} is sampled from the uniform distribution $[-\theta_{\text{max}}, \theta_{\text{max}}]$. This is also the smallest $\Delta\theta$ we can choose. While a smaller $\Delta\theta$ also ensures that the photon is scattered towards the detector, the weight becomes smaller, see Eq. (4). Thus, you get the same number of photons scattered towards the detector, but with a smaller average weight. Choosing a larger $\Delta\theta$ will give the correct result, as the increase in average weight is balanced by the decreased probability of being scattered towards the detector. However, this results in less detected photons and poorer statistics. There is also a chance that the sampled azimuth angle ϕ, results in a scattering plane that does not intersect with the ring detector, i.e. there are no scattering angles θ resulting in scattering towards the eyeball detector. In this case, we simply choose $\Delta\theta = \pi$ and sample the scattering angle from the uniform distribution $[0, \pi]$. Choosing a probability $\epsilon = 0.2$ of being turned towards the eyeball detector prior to scattering, has proven to work well for all samples investigated in this study.

Optics EXPRESS

This results in a considerable increase in photons sent towards the detector, while maintaining a large enough portion to explore the sample volume, providing good statistics for the multiple scattered photons.

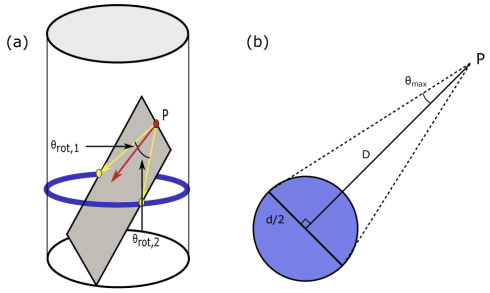

Fig. 2. Illustration of the geometry applied to find (a) the scattering direction resulting in the photon being sent directly towards the eyeball detector (blue ring torus), and (b) the maximum angle the photon can scatter away from the center of the detector while still being detected.

2.2.3. Splitting

As mentioned, applying DDIS may result in photons accumulating weight, causing spikes in the detected signal. In order to avoid this, a photon is split into N identical parts when the photon reaches a weight larger than a threshold value. When the photon is split, the weight of each part can be calculated as

$$w_{\text{split}} = w/N. \tag{8}$$

The N parts are then propagated individually from the position where the original photon was split and in the same scattering direction. For the results presented in Fig. 3, the threshold weight was chosen to be $w_{\max} = 1$, and N was chosen to be 2 times the floor of w ($N = 2\lfloor w \rfloor$), so that the photon is split into a minimum of two parts. The threshold value can be set to $w_{\max} = 1$ because the MFPTR results in reduced weight before the first scattering event. However, this assumption may fail when the optical depth becomes very large.

2.3. Analysing simulated VSFs

Using the improved MC simulation described in section 2.2, a range of different sample volumes were simulated using FF phase functions with scattering coefficients b ranging from 0.05 to 20 m^{-1}. The wide range of different phase functions used enables the investigation of the multiple-scattering errors dependency on the phase function. The FF phase function is an approximate analytic expression of the phase function for an ensemble of spherical particles [25,26]. It has been extensively used for modeling the VSF of natural waters, usually providing a better fit than the simple one-term Henyey-Greenstein phase function. The function takes two parameters n and m, where n is the refractive index of the particles and m defines the particle size distribution according to the inverse power-law $f(D) \propto D^{-m}$, where D is particle diameter. The quantitative investigation of the multiple-scattering errors is performed by calculating the relative error for the scattering coefficient b, backscattering coefficient b_b and phase function

Optics EXPRESS

$p(\theta)$. The relative error in the scattering coefficient is calculated as

$$e_b = \frac{|b_{\text{sim}} - b_{\text{true}}|}{b_{\text{true}}}, \tag{9}$$

where b_{sim} and b_{true} are calculated from the simulated VSF and the single-scattering calculated VSF, respectively. In order to better compare simulated and true scattering coefficient, they are both calculated for the angular measurement range of the LISST-VSF, so that

$$b = 2\pi \int_{0.1^\circ}^{150^\circ} \beta(\theta) \sin\theta d\theta. \tag{10}$$

Thus, b_{true} is in fact slightly smaller than the true scattering coefficient, but equal to the scattering coefficient one would measure given a perfect measurement, i.e. no multiple scattering. The relative error e_{b_b} in the backscattering coefficient is calculated similarly to the error in the scattering coefficient, where the backscattering coefficient b_b is calculated according to Eq. (10), but over the range 90-150°.

The relative error of the phase function $p(\theta) = \beta(\theta)/b$ is calculated as the mean relative error across the range 0.1-150°,

$$e_p = \frac{1}{K} \sum_i^K \frac{|p_{\text{sim}}(\theta_i) - p_{\text{true}}(\theta_i)|}{p_{\text{true}}(\theta_i)}. \tag{11}$$

The simulated phase functions are measured at log-spaced angles for θ_i<15°, similar to LISST-VSF measurements. To avoid bias in the forward scattering direction, the data was interpolated to evenly spaced angles with a step size of 1°, so that $K = 150$.

The errors are analyzed in terms of the optical depth $\tau = cL$ and asymmetry parameter g. Here, $L = 15$ cm is the length of the cylindrical sample volume. As all samples are simulated with absorption $a = 0$ m^{-1}, we have $\tau = cL = bL$. The asymmetry parameter g is a measure of how forward peaked the phase function is, and is calculated as the mean cosine of the phase function [29]

$$g = \langle \cos\theta \rangle = 2\pi \int_0^\pi p(\theta) \cos\theta \sin\theta d\theta. \tag{12}$$

The asymmetry parameter ranges from $g = 0$ for isotropic scattering and approaches $g = 1$ for very forward scattering phase functions.

3. Results and discussion

3.1. Efficiency of VRMs

The efficiency of the VRMs is tested by running the simulation for 1000 s, both with and without the VRMs. A FF phase function with parameters $n = 1.05$ and $m = 3.5$ was used for this purpose, simulated with scattering coefficients $b = 0.05$, 0.17, 0.55, 1.8, 6, and 20 m^{-1}. This phase function is used due to its large asymmetry parameter (low scattering in the backward direction), so that the increase in efficiency is expected to be large. In order to evaluate the efficiency of the different VRMs, the VSF obtained with the different VRMs are compared to a reference VSF, which is obtained using no VRMs and a very long simulation time (t>24 hours). The resulting simulated VSFs from the efficiency test are presented in Fig. 3. The parameters used for the simulations with VRMs are $\epsilon = 0.2$ and $w_{\text{max}} = 1$.

It becomes immediately clear from looking at the plots that the applied VRMs result in major noise reduction. This is especially evident in the backward direction and for low scattering coefficients. Also, the new algorithm (with VRMs) reproduces the results obtained with the old algorithm (without VRMs). As the old algorithm has been validated in a previous study (see [14]), the new algorithm is indirectly validated. The discontinuity in the simulated VSFs seen at

Optics EXPRESS

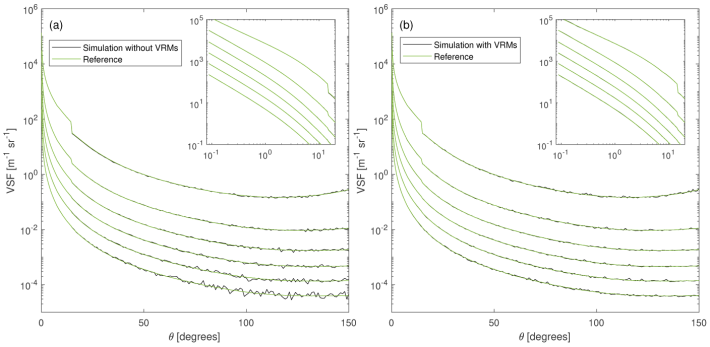

Fig. 3. Simulated VSF for a total simulation time of 1000s (black) compared to a reference VSF which is simulated without any VRMs for a very long time (t>24 hours). The VSF is plotted for six scattering coefficients b, increasing from 0.05 (bottom) to 20 (top). The insets show the volume scattering function plotted from $0 - 20^\circ$ where both the x- and y-axis are in log scale. (a) Simulation without VRMs. (b) simulation with VRMs.

$\theta = 15^\circ$ and large scattering coefficients is due to the abrupt decrease in single-scattering path length going from the ring detector ($l = 15$ cm) to eyeball detector ($l = 10.3$ cm), where a longer path length results in more multiple scattered photons and larger errors [14].

In order to quantify the improvement, we calculate the average relative error in the VSFs over the entire measurement range as

$$e_\beta = \frac{1}{K}\sum_{i}^{K}\frac{|\beta_{\text{sim}}(\theta_i) - \beta_{\text{ref}}(\theta_i)|}{\beta_{\text{ref}}(\theta_i)}, \tag{13}$$

where K is the number of measurement points (same as LISST-VSF), and β_{sim} and β_{ref} are simulated and reference VSF, respectively. The VRMs are compared by running the simulation with the different methods applied. The results are plotted in Fig. 4, where e_β is plotted against the optical depth $\tau = cL$. In addition to the 1000s simulations, VSFs simulated with no VRMs and for 10000s (10x the time) are included to better quantify the improvement.

From Fig. 4(a), one can see that the DDIS is by far the most important method for reducing the variance in the simulated VSFs. In fact, comparing the simulation with DDIS to the simulation without any VRMs and a computation time of 10000s, we see that applying DDIS results in over 10x increased efficiency for the five smallest optical depths. At the smallest optical depths, applying DDIS results in much more than 10x efficiency, while it is similar to 10x efficiency at an optical depth of 3. When no VRMs are applied, the relative error decreases with increasing optical depth. This is due to the increased number of photons interacting with the sample volume, and the increased number of multiple scattered photons which increases the probability of scattering to larger angles, resulting in better statistics in the backward direction. Applying only MFPTR, the relative error is significantly decreased at small optical depths compared to the VSF with no VRMs. This is due to the large portion of photons that would pass straight through the sample volume without interaction. As the optical depth increases, this VRM becomes less efficient.

A cropped version of Fig. 4(a) is presented in Fig. 4(b), making it easier to visually differentiate the plots where DDIS is applied. Comparing splitting to MFPTR (combined with DDIS), one can see that MFPTR is more efficient at small optical depths, approximately halving the relative

Optics EXPRESS

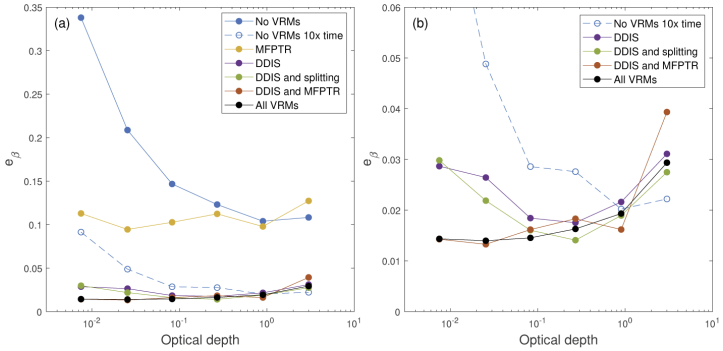

Fig. 4. Average relative error between simulated VSFs with a running time of 1000s and a low-variance simulation used as reference. (a) Comparison of VSFs simulated with no VRMs (blue), MFPTR (yellow), DDIS (purple), DDIS and splitting (green), DDIS and MFPTR (orange), and with all VRMs (black). (b) Zoomed in on the lowest error plots.

error for the smallest optical depth compared to when only DDIS is applied. As the optical depth increases, the splitting method becomes more important for the overall efficiency. This is an expected result, as larger optical depths lead to increasing number of scattering events per photon, hence more weight accumulation. Applying all VRMs results in the overall best results, having the lowest or close to lowest relative error for all optical depths. The same values for the parameters ϵ and $w_{\max}$ are used for all optical depths shown in Fig. 4. Thus, the VRMs are not optimized and the efficiency of the different methods is expected to improve by tailoring these parameters to the different optical depths. The phase function is also expected to influence the efficiency of the different VRMs and optimal values of the VRM parameters. For instance, the efficiency improvement is expected to be smaller for phase functions with more scattering in the backward direction, as the simulation without VRMs would perform better.

3.2. Multiple scattering errors

Two examples of simulated phase functions are presented in Fig. 5, representing opposite ends of the spectrum of g for the simulations performed in this study. In both examples, it can be seen that the deviation between simulated and true VSFs increases with increasing scattering coefficient (from bottom spectrum to top spectrum), where the deviation is solely due to multiple scattering.

3.2.1. Phase function dependency

In order to analyze how the multiple scattering error depends on the optical depth and phase function, the errors e_b and e_{b_b} and e_p are plotted as a function of τ and g. A contour plot for e_b is presented in Fig. 6, where all FF phase functions investigated are included. A second x-axis is added to the top of the graph, displaying the x-values in terms of the scattering coefficient b. The phase functions simulated have parameters m ranging from 3.5 to 4.5 in steps of 0.5, and parameters n ranging from 1.05 to 1.35 in steps of 0.05, resulting in a total of 21 phase functions. The phase functions are simulated with the same scattering coefficients as for the VRMs test.

From Fig. 6, one can see how the relative error increases diagonally from small τ and g to large τ and g. As expected, the multiple scattering error increases with increasing τ, but there is also a considerable increase with g as well. For the largest g, corresponding to $n = 1.05$ and

Optics EXPRESS

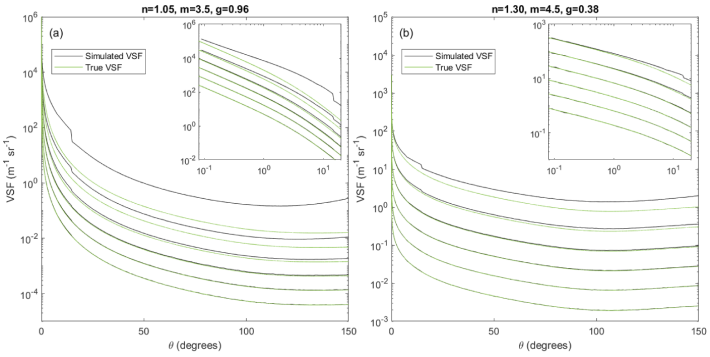

Fig. 5. VSF plotted for simulation with all VRMs applied (black) and true values (green). The VSF is plotted for six scattering coefficients b, increasing from 0.05 (bottom) to 20 (top). (a) FF phase function with parameters $n = 1.05$, $m = 3.5$ (b) FF phase function with parameters $n = 1.30$ and $m = 4.5$. The insets show the volume scattering function plotted from $0 - 20°$ where both the x- and y-axis are in log scale.

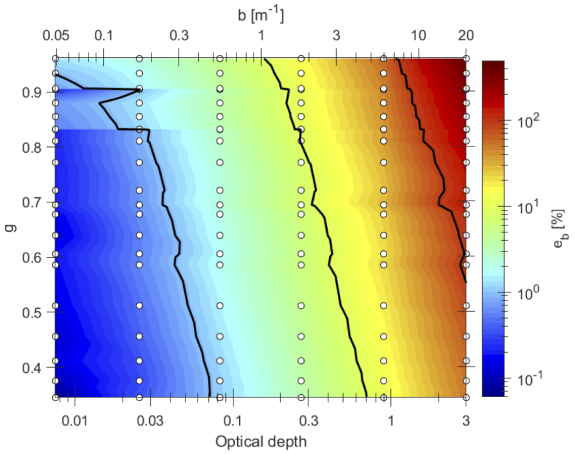

Fig. 6. Contour plot of the relative error e_b in the scattering coefficient. The error is plotted as a function of optical depth τ (or scattering coefficient b) and asymmetry parameter g. The x- and z-axis (color-axis) are plotted in log scale, while the y-axis is linear. The simulation data is marked as white dots. The black lines indicate a 1%, 10% and 100% relative error, respectively, going from left to right.

Optics EXPRESS

m = 3.5, the relative error in b increases from 1.4% at τ = 0.0075 (b = 0.05) to 494% at τ = 3 (b = 20). An error of 10% is observed at approximately τ = 0.15 (b = 1) while the error is above 100% for τ>1 (b>7). Looking at the dependency on g we can see that, for τ = 3, the relative error increases from 55% at g = 0.34 to 494% at g = 0.96. The dependency on g can largely be explained by the LISST-VSF geometry [14]. Because the radius of the cylindrical sample volume is only 2 cm, photons scattered at large angles may exit the sample volume with relative short path lengths. On the other hand, photons scattered at small angles may travel along the z-axis, which has a length of 15 cm. Thus, very forward peaked phase functions will have more photons traveling a longer path length within the sample volume, resulting in more multiple scattered photons and larger errors.

Some distortions can be observed in the top-left corner, caused by a sudden increase in the relative error for some of the phase functions. This is especially apparent at g = 0.83 and g = 0.91, where sharp horizontal lines can be seen. Here, there are two different phase functions but with a similar g, such that they cannot be visually separated in the plot. It is not clear why this abrupt change in relative error only occurs for small optical depths. However, on closer inspection we find that all the phase functions with increased relative error have the smallest parameter m = 3.5, where a small m is related to a large VSF in the extreme forward direction. This suggests that extremely forward peaked phase functions are more sensitive to multiple scattering errors at small optical depths. Similarly to Fig. 6, contour plots are plotted in Fig. 7 and 8 for e_{b_b} and e_p, respectively.

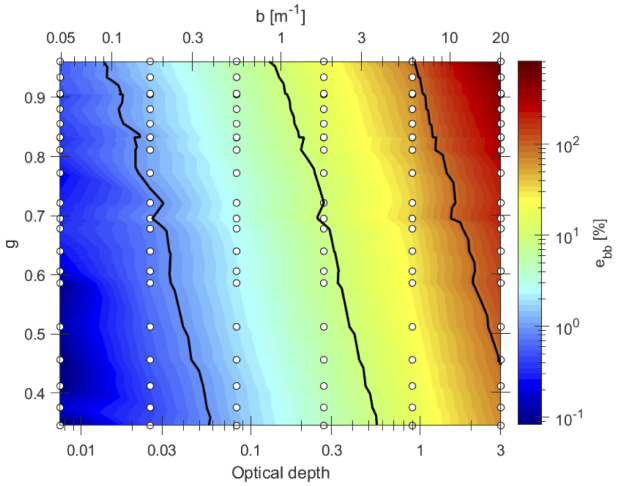

Fig. 7. Contour plot of the relative error e_{b_b} in the backscattering coefficient. The error is plotted as a function of optical depth τ (or scattering coefficient b) and asymmetry parameter g. The x- and z-axis (color-axis) are plotted in log scale, while the y-axis is linear. The simulation data is marked as white dots. The black lines indicate a 1%, 10% and 100% relative error, respectively, going from left to right.

The plot of e_{b_b} displays the same trend seen for e_b in Fig. 6, with e_{b_b} increasing diagonally from small τ and g to large τ and g. For the largest g, the relative error increases from 0.6% to 828%, while for the largest optical depth, the error increases from 74% to 828%. Generally, the error e_{b_b} is larger than e_b, which can be explained by the large difference in phase function in

Optics EXPRESS

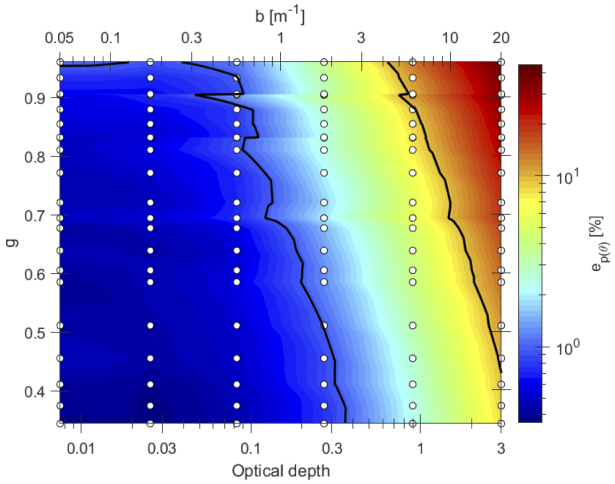

Fig. 8. Contour plot of the relative error e_p in the phase function. The error is plotted as a function of optical depth τ (or scattering coefficient b) and asymmetry parameter g. The x- and z-axis (color-axis) are plotted in log scale, while the y-axis is linear. The simulation data is marked as white dots. The black lines indicate a 1% and 10% relative error, respectively, going from left to right.

the forward and backward direction. As more photons are scattered multiple times, the angular distribution becomes more random. Thus, multiple scattered photons are more evenly distributed over the angular spectrum than the single scattered photons. Adding a photon to an angle where the single-scattering signal is low results in a larger relative increase in signal than adding a photon to an angle where the single-scattering signal is high. Thus, the relative increase in detected photons is larger in the backwards direction, resulting in larger backscattering errors. The distortions seen in Fig. 6 are not observed for the backscattering coefficient, indicating that the distortions are indeed related to forward scattering.

The same diagonal trend is observed for the error $e_{p(\theta)}$, as for e_b and e_{b_b}. In general, the error e_p is much smaller than the errors in e_b and e_{b_b}. This is because the phase functions is normalized, such that the error is only related to the shape of the VSF and not the magnitude. For the largest g, the relative error increases from 1% to 45%, while for the largest optical depth, the error increases from 9% to 45%. For τ<0.1, the relative error e_p does not increase much, which is due to noise being the dominating source of error when the difference between simulated and true VSF becomes small. This is not the case for e_b and e_{b_b}, as both b and b_b are calculated as integrals over an angular range. Thus, fluctuations in the signal are evened out. The plot in Fig. 8 also displays the same distortions seen in Fig. 6, most notably at $g = 0.91$.

In order to better demonstrate the effect of the phase function on the multiple scattering error, the relative error e_b is plotted in Fig. 9 as a function of only g at the largest optical depth $\tau = 3$. Here, we have also included a set of Henyey-Greenstein phase functions, which are much less peaked in the extreme forward direction [29].

From the plot, we can see that the different FF phase functions follow slightly different trend lines, depending on the parameter m. For the phase functions investigated here, the phase functions with $m = 4.0$ have the smallest error for similar asymmetry parameters g, while phase

Optics EXPRESS

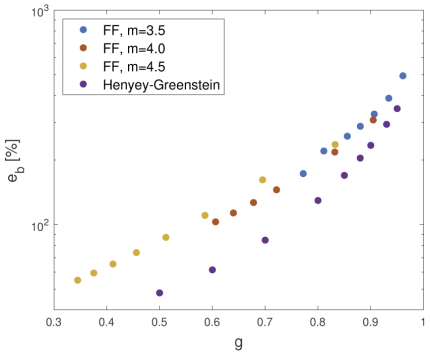

Fig. 9. Relative error e_b in the scattering coefficient plotted as a function of asymmetry parameter g. The error is plotted for a variety of FF phase functions with parameter $m = 3.5$ (blue), $m = 4.0$ (red) and $m = 4.5$ (yellow), in addition to a set of Henyey-Greenstein functions (purple).

functions with $m = 3.5$ and $m = 4.5$ seem to have similar errors. However, this cannot be concluded due to the small overlap in g for the $m = 3.5$ and $m = 4.5$ phase functions. Comparing the FF to the Henyey-Greenstein phase functions, we see that the errors are significantly smaller for the latter. This is due to the very forward peaked nature of the FF phase functions, compared to Henyey-Greenstein phase functions. The difference in error between Henyey-Greenstein and FF phase functions is largest for small g, and decreases with increasing g. The results presented here demonstrates that, while the asymmetry parameter g is a good indicator for the size of the error, details in the phase function may influence the error significantly.

3.2.2. Absorption dependency

The results presented in the previous section were all obtained with an absorption coefficient of $a = 0$ m^{-1}. In Fig. 10, the error e_b is plotted for a phase function simulated with $a = 0c$, $a = 0.2c$ and $a = 0.4c$, in order to investigate the effect of absorption. The same set of attenuation coefficients, i.e. optical depths, is used for each absorption coefficient. Thus, the only difference between each simulation set is the ratio of scattering to absorption.

In Fig. 10(a), the error in the scattering coefficient e_b is plotted against the attenuation coefficient c. Here, we see that e_b decreases with increasing absorption, which is expected as the scattering decreases, hence the number of multiple scattered photons decreases. The difference in error between the different values of a increase with increasing optical depths. At the smallest optical depth, the errors at $a = 0c$ and $a = 0.4c$ are 1.43% and 1.28%, respectively, while at the largest optical depth, the errors are 494% and 180%. In Fig. 10, e_b is plotted against the scattering coefficient b. Here, it is clearly demonstrated that the multiple scattering error is only dependent on b. Thus, when evaluating the multiple scattering error in a measurement, it is better to use the scattering coefficient, or the size bL, rather than the attenuation coefficient or the optical depth $\tau = cL$.

Optics EXPRESS

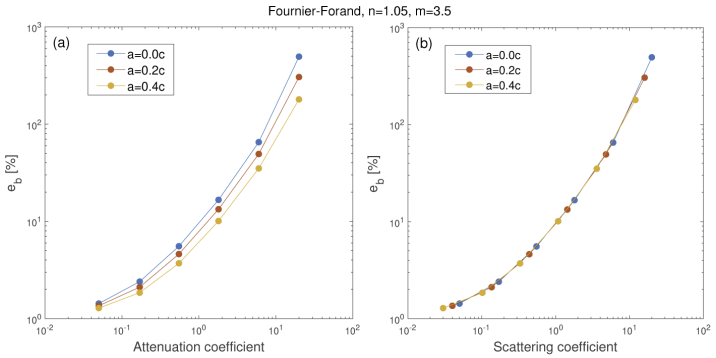

Fig. 10. The relative error e_b in the scattering coefficient plotted as function of (a) attenuation coefficient c, and (b) scattering coefficient b. The error is plotted for an absorption of $a = 0c$ (blue), $a = 0.2c$ (orange) and $a = 0.4c$ (yellow).

4. Summary and conclusion

We have performed an extensive investigation into the errors originating from multiple scattering for the LISST-VSF instrument. Several different Henyey-Greenstein and FF phase functions have been simulated, with scattering coefficients b ranging from 0.05 to 20 m^{-1}, and with asymmetry parameters g ranging from 0.34 to 0.96. The results show that errors in the scattering coefficient b, backscattering coefficient b_b, and phase function $p(\theta)$, depend heavily on both the scattering coefficient and asymmetry parameter. We find that multiple scattering errors may only be considered negligible (<10%) when the scattering coefficient is smaller than 1 m^{-1}. The largest error is observed for the backscattering coefficient, where the error ranges from 74% to 828% for the different phase functions at an optical depth of $\tau = 3$. For the scattering coefficient, the error ranges from 55% to 494%, and the phase function from 9% to 45%. The results also show significant differences between Henyey-Greenstein and FF phase functions, where the error is considerable larger for the FF phase functions for similar asymmetry parameters. We conclude that this is due to the extremely forward peaked shape of the FF functions compared to the Henyey-Greenstein functions, resulting in a longer average distance traveled within the sample volume and more multiple scattered photons. This observation demonstrates the importance of the shape of the phase function when measuring the VSF with the LISST-VSF instrument. The effect of absorption was also investigated. We found, when the error is plotted against optical depth, the error decreases with increasing absorption. However, when plotted against the scattering coefficient, the error was independent of absorption, i.e. the error is only dependent on the scattering coefficient for a given phase function.

The LISST-VSF has an angle-dependent path length varying between 10.3 and 17.5 cm, which accentuates multiple scattering effects on the measured VSF in turbid waters, in particular after applying the absolute calibration [13]. This does not mean other VSF meters are not affected by multiple scattering. The MASCOT and MVSM instrument, two VSF meters that have been widely used for modern era VSF measurements, have reported constant path lengths of 20.0 cm [30,31]. The longer path lengths imply that the multiple scattering errors can be significant for b>1 m^{-1}. This can have implications for historical data sets in waters with moderate-to-high turbidity. However, we should note that the results in this study are instrument-specific to the LISST-VSF.

Optics EXPRESS

A set of variance reduction methods has been applied to the previously developed Monte Carlo simulation, resulting in drastically reduced computation time. This enables a fast method for analysing errors in the measured VSFs. The most important of these methods is detector directional importance sampling (DDIS), where the photons have an increased probability of being scattered towards the eyeball detector. The applied methods result in over 10 times increased efficiency for optical depths below 0.9, and similar to 10 times increased efficiency up to an optical depth of $\tau = 3$. The efficiency is expected to be further enhanced by optimizing the parameters used for the variance reduction methods. The efficiency is also expected to depend on the phase function, where the improvement in efficiency is expected to increase with increasing asymmetry parameter. This is because the simulation without variance reduction methods achieves a very low signal in the backward direction when the asymmetry parameter is large. For small asymmetry parameters, many photons are naturally scattered in the backward direction, reducing the need for DDIS. The implementation of the detector directional importance sampling in this simulation is geometry specific, however, the method of implementation might be relevant for other cases.

For our future work we aim to correct the errors arising from multiple scattering in the LISST-VSF instrument. The main goal is to correct the VSF over the entire LISST-VSF measuring range, thus correcting both the scattering coefficient, backscattering coefficient, and phase function at the same time.

Funding. Universitetet i Bergen.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. K. Stamnes and J. J. Stamnes, *Radiative transfer in coupled environmental systems: An introduction to forward and inverse modeling* (John Wiley & Sons, 2016).
2. M. Babin, A. Morel, V. Fournier-Sicre, F. Fell, and D. Stramski, "Light scattering properties of marine particles in coastal and open ocean waters as related to the particle mass concentration," Limnol. Oceanogr. **48**(2), 843–859 (2003).
3. E. Boss, L. Taylor, S. Gilbert, K. Gundersen, N. Hawley, C. Janzen, T. Johengen, H. Purcell, C. Robertson, D. W. H. Schar, G. J. Smith, and M. N. Tamburri, "Comparison of inherent optical properties as a surrogate for particulate matter concentration in coastal waters," Limnol. Oceanogr.: Methods **7**(11), 803–810 (2009).
4. G. Neukermans, H. Loisel, X. Mériaux, R. Astoreca, and D. McKee, "In situ variability of mass-specific beam attenuation and backscattering of marine particles with respect to particle size, density, and composition," Limnol. Oceanogr. **57**(1), 124–144 (2012).
5. M. S. Twardowski, E. Boss, J. B. Macdonald, W. S. Pegau, A. H. Barnard, and J. R. V. Zaneveld, "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.: Oceans **106**(C7), 14129–14142 (2001).
6. X. Zhang, R. H. Stavn, A. U. Falster, D. Gray, and R. W. Gould Jr, "New insight into particulate mineral and organic matter in coastal ocean waters through optical inversion," Estuarine, Coastal Shelf Sci. **149**, 1–12 (2014).
7. D. Koestner, D. Stramski, and R. A. Reynolds, "Polarized light scattering measurements as a means to characterize particle size and composition of natural assemblages of marine particles," Appl. Opt. **59**(27), 8314–8334 (2020).
8. Y. Agrawal, A. Whitmire, O. A. Mikkelsen, and H. Pottsmith, "Light scattering by random shaped particles and consequences on measuring suspended sediments by laser diffraction," J. Geophys. Res.: Oceans **113**(C4), C04023 (2008).
9. E. Boss, N. Haëntjens, T. K. Westberry, L. Karp-Boss, and W. H. Slade, "Validation of the particle size distribution obtained with the laser in-situ scattering and transmission (LISST) meter in flow-through mode," Opt. Express **26**(9), 11125–11136 (2018).
10. A. Tonizzo, M. Twardowski, S. McLean, K. Voss, M. Lewis, and C. Trees, "Closure and uncertainty assessment for ocean color reflectance using measured volume scattering functions and reflective tube absorption coefficients with novel correction for scattering," Appl. Opt. **56**(1), 130–146 (2017).
11. P. J. Werdell, L. I. W. McKinna, E. Boss, S. G. Ackleson, S. E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, D. Stramski, J. M. Sullivan, M. S. Twardowski, M. Tzortziou, and Z. Xiaodong, "An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing," Prog. Oceanogr. **160**, 186–212 (2018).

Optics EXPRESS

12. D. Blondeau-Patissier, J. F. Gower, A. G. Dekker, S. R. Phinn, and V. E. Brando, "A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans," Prog. Oceanogr. **123**, 123–144 (2014).
13. H. Sandven, A. S. Kristoffersen, Y.-C. Chen, and B. Hamre, "In situ measurements of the volume scattering function with LISST-VSF and LISST-200X in extreme environments: Evaluation of instrument calibration and validity," Opt. Express **28**(25), 37373–37396 (2020).
14. H. S. Ugulen, H. Sandven, B. Hamre, A. S. Kristoffersen, and C. Sætre, "Analysis of multiple scattering errors in LISST-VSF volume scattering function measurements using Monte Carlo simulations and experimental data," Opt. Express **29**(8), 12413–12428 (2021).
15. J. Kirk, "Monte Carlo modeling of the performance of a reflective tube absorption meter," Appl. Opt. **31**(30), 6463–6468 (1992).
16. D. Stramski and J. Piskozub, "Estimation of scattering error in spectrophotometric measurements of light absorption by aquatic particles from three-dimensional radiative transfer simulations," Appl. Opt. **42**(18), 3634–3646 (2003).
17. D. McKee, J. Piskozub, R. Röttgers, and R. A. Reynolds, "Evaluation and improvement of an iterative scattering correction scheme for in situ absorption and attenuation measurements," J. Atmospheric Ocean. Technol. **30**(7), 1527–1541 (2013).
18. J. Piskozub, D. Stramski, E. Terrill, and W. K. Melville, "Influence of forward and multiple light scatter on the measurement of beam attenuation in highly scattering marine environments," Appl. Opt. **43**(24), 4723–4731 (2004).
19. D. Doxaran, E. Leymarie, B. Nechad, A. Dogliotti, K. Ruddick, P. Gernez, and E. Knaeps, "Improved correction methods for field measurements of particulate light backscattering in turbid waters," Opt. Express **24**(4), 3615–3637 (2016).
20. S. Vadakke-Chanat, P. Shanmugam, and B. Sundarabalan, "Monte Carlo simulations of the backscattering measurements for associated uncertainty," Opt. Express **26**(16), 21258–21270 (2018).
21. X. Zhang, E. Leymarie, E. Boss, and L. Hu, "Deriving the angular response function for backscattering sensors," Appl. Opt. **60**(28), 8676–8687 (2021).
22. R. Buras and B. Mayer, "Efficient unbiased variance reduction techniques for Monte Carlo simulations of radiative transfer in cloudy atmospheres: The solution," J. Quant. Spectrosc. Radiat. Transf. **112**(3), 434–447 (2011).
23. Z. Vali, A. Gholami, Z. Ghassemlooy, D. G. Michelson, M. Omoomi, and H. Noori, "Modeling turbulence in underwater wireless optical communications based on Monte Carlo simulation," J. Opt. Soc. Am. A **34**(7), 1187–1193 (2017).
24. L. R. Poole, D. D. Venable, and J. W. Campbell, "Semianalytic Monte Carlo radiative transfer model for oceanographic lidar systems," Appl. Opt. **20**(20), 3653–3656 (1981).
25. G. R. Fournier and J. L. Forand, "Analytic phase function for ocean water," in *Ocean Optics XII*, vol. 2258 (International Society for Optics and Photonics, 1994), pp. 194–201.
26. M. Jonasz and G. Fournier, *Light scattering by particles in water: Theoretical and experimental foundations* (Elsevier, 2007).
27. E. Mainegra-Hing and I. Kawrakow, "Variance reduction techniques for fast Monte Carlo CBCT scatter correction calculations," Phys. Med. Biol. **55**(16), 4495–4507 (2010).
28. M. A. Cooper and E. W. Larsen, "Automated weight windows for global Monte Carlo particle transport calculations," Nucl. Sci. Eng. **137**(1), 1–13 (2001).
29. L. G. Henyey and J. L. Greenstein, "Diffuse radiation in the galaxy," Astrophys. J. **93**, 70–83 (1941).
30. X. Zhang, M. Lewis, M. Lee, B. Johnson, and G. Korotaev, "The volume scattering function of natural bubble populations," Limnol. Oceanogr. **47**(5), 1273–1282 (2002).
31. J. M. Sullivan and M. S. Twardowski, "Angular shape of the oceanic particulate volume scattering function in the backward direction," Appl. Opt. **48**(35), 6811–6819 (2009).

Paper III

A neural network approach for correction of multiple scattering errors in the LISST-VSF instrument

Håvard S. Ugulen, Daniel Koestner, Håkon Sandven, Børge Hamre, Arne S. Kristoffersen, and Camilla Sætre
Submitted to Optics Express (May 2023)

A neural network approach for correction of multiple scattering errors in the LISST-VSF instrument

Håvard S. Ugulen*, Daniel Koestner, Håkon Sandven, Børge Hamre, Arne S. Kristoffersen and Camilla Saetre

[1]*Department of Physics and Technology, University of Bergen, Allegaten 55, 5007 Bergen, Norway*
havard.ugulen@uib.no

Abstract:

The LISST-VSF is a commercially developed instrument used to measure the volume scattering function (VSF) and attenuation coefficient in natural waters, which are important for remote sensing, environmental monitoring and underwater optical wireless communication. While the instrument has been shown to work well at relatively low particle concentration, previous studies have shown that the VSF obtained from the LISST-VSF instrument is heavily influenced by multiple scattering in turbid waters. High particle concentrations result in errors in the measured VSF, as well as the derived properties, such as the scattering coefficient and phase function, limiting the range at which the instrument can be used reliably. Here, we present a feedforward neural network approach for correcting this error, using only the measured VSF as input. The neural network is trained with a large dataset generated using Monte Carlo simulations of the LISST-VSF with scattering coefficients $b = 0.05 - 50\,\mathrm{m}^{-1}$, and tested on VSFs from measurements with natural water samples. The results show that the neural network estimated VSF is very similar to the expected VSF without multiple scattering errors, both in angular shape and magnitude. One example showed that the error in the scattering coefficient was reduced from 103% to 5% for a benchtop measurement of natural water sample with expected $b = 10.6\,\mathrm{m}^{-1}$. Hence, the neural network drastically reduces uncertainties in the VSF and derived properties resulting from measurements with the LISST-VSF in turbid waters.

1. Introduction

Accurate characterization of light-particle interactions in marine environments have become increasingly important in recent years due to growing interest in applications of optical measurements for environmental monitoring, remote sensing, and underwater wireless communication [1–6]. The volume scattering function (VSF) is an important inherent optical property (IOP), describing the angular distribution and intensity of scattered light. Measurements of the VSF β can be used to derive the scattering coefficient b through integration of the VSF over all scattering directions and the phase function $\tilde{\beta}$ through normalization of the VSF by b. Accurate measurements of the VSF of natural water samples have been difficult to obtain, requiring custom-built and complex instrumentation which are often challenging to deploy for *in situ* measurements. Consequently, the VSF has been sparingly measured, and common simplified models have not been extensively validated [7], despite its importance in radiative transfer modelling [2, 5, 8, 9]. Variations in the scattering properties of natural waters are primarily governed by suspended particles on the

micro- and nano-scale. The VSF may be utilized to extract considerable information on the particulate composition in the water mass [10–13].

The LISST-VSF (Sequoia Scientific) is a commercially available instrument, designed to perform *in situ* measurements of the VSF at a single light wavelength of 515 nm over angles $0.01 - 150^\circ$ with two detection systems; a series of 32 logarithmically-spaced ring detectors covering the range $0.01 - 15^\circ$ and a rotating eyeball detector covering the range $15 - 150^\circ$ in 1° increments. Additionally, the LISST-VSF can simultaneously measure the attenuation coefficient c, enabling the extraction of the absorption coefficient a trough the relation $c = a + b$. Although the LISST-VSF has been validated to perform well with relatively low concentrations of known particles [14–16], previous studies have also shown that large errors in the measured VSF arise in turbid waters [13, 17, 18]. In particular, Monte Carlo simulations of the LISST-VSF revealed elevated and distorted VSFs originating from multiple scattering. Consequently, the phase function and scattering coefficient derived from the VSF measurement are also subject to multiple scattering errors. The Monte Carlo simulation has been verified by laboratory measurements [17], where only small deviations were observed. Simulations show that, for an expected scattering coefficient of $b \sim 1.5\,\mathrm{m}^{-1}$, an overestimation of 10% can be observed in the measured scattering coefficient, reaching 100% at $b \sim 10\,\mathrm{m}^{-1}$. On the other hand, the attenuation coefficient c is not significantly affected by multiple scattering errors [13, 15, 19]. Hence, errors in the scattering coefficient also affect the derived absorption coefficient ($a = c - b$), which can become negative for sufficient errors in the measured scattering coefficient.

Multiple scattering is a known challenge for optical measurement systems seeking to characterize turbid media [10, 11, 20, 21]. Multiple scattering of light within the LISST-VSF instrument impacts the VSF measurement results in two ways: overestimation of VSF magnitude and incorrect angular shape (i.e., phase function). These issues relate to multiple scattered light which is presumed lost along the path between the light source and detector. Both scattering and absorption losses are accounted for along this path using an attenuation correction determined as e^{-cz} where z is pathlength. This correction utilizes a measurement of attenuation from the LISST-VSF with a very narrow acceptance angle ($<0.04^\circ$). However, a portion of the scattered light assumed lost is not truly lost as it continues through the water and is ultimately detected. For example, light which in a single scattering scenario would be detected at 20° is scattered a second or third time, resulting in detected light at either the same, or a different detection angle. This light is accounted for by the attenuation correction for the 20° detection angle, but is also observed as an actual signal in the detector. Hence, this light is effectively double-counted. Furthermore, multiple scattering generally results in a flattening of the typical forward-peaked phase function of marine particles (e.g., [22]). Combined, these effects result in an overestimation of VSF magnitude at all scattering angles and a flatter VSF than would be observed in a single scattering scenario.

The correction of these undesirable multiple scattering effects on light scattering measurements is nontrivial and depends on many factors including particle properties (e.g., concentration and phase function) and instrument geometry (e.g., pathlength and acceptance angle). A common solution has been to develop light scattering meters with relatively short pathlengths; however, this approach is not always feasible considering desired angular-resolution and retrieval of adequate signal for detection of small scattering signals common in oceanic environments. The HydroScat-6 instrument (HS-6; HOBI Labs, Inc.) has a 15 cm pathlengh, approximately the same as the LISST-VSF instrument, but only measures light scattered at 141º from incident direction. HOBI Labs recommend adjusting the scattering coefficient by a factor of 0.4 when deriving an attenuation coefficient for the attenuation correction (i.e., $c^* = a + 0.4b$, where c^* refers to an adjusted c for attenuation correction), under the assumption that only a portion of scattered light

is truly lost along the path [23]. Recently, Doxaran et al. [24] used light simulations of HydroScat instruments to show that adjustments to the scattering coefficient of factors of about 0.05–0.4 were actually warranted for different scenarios of optical properties. This type of approach may be sufficient for some cases with a single-angle scattering meter; however, more advanced approaches are necessary for the complexity of multiple scattering effects in the LISST-VSF instrument which includes two detector systems covering a wide angular range.

The Monte Carlo model developed in [17] is able to reliably reproduce a LISST-VSF measurement given a phase function and scattering coefficient as inputs. By simulating a large variety of different VSFs, we have shown that the magnitude of the multiple scattering error is dependent on both the scattering coefficient and phase function, i.e., magnitude and shape of the VSF [18]. Furthermore, the error varies with angle of detection, such that the percentage error in the measured VSF at 15° may be significantly different from the error at 150°, making it challenging to apply corrections to the measured VSF. Koestner et al. [13] developed a correction function based on empirical data, where two mineral-dominated samples were diluted to obtain benchtop measurements with attenuation coefficients ranging from $1.8\,\mathrm{m}^{-1}$ to $14\,\mathrm{m}^{-1}$. Assuming negligible multiple scattering errors for the smallest attenuation coefficient, an expected VSF could be calculated for the higher concentration samples affected by multiple scattering. Hence, by comparing expected VSF with measured VSF, a correction was found as a function of scattering angle and scattering coefficient. For applications, the scattering coefficient was estimated by $b = c_{\mathrm{LISST}} - a_{\text{ac-s}}$, where $a_{\text{ac-s}}$ is the absorption coefficient measured with an ac-s instrument (Sea-Bird Scientific) and c_{LISST} is the attenuation coefficient measured with the LISST VSF instrument. While this method accounts for the angle dependent error, it does not account for the effect of varying phase functions and requires an unbiased estimate of the absorption coefficient.

Recently, artificial neural networks (ANNs) have seen a rise in popularity due to their wide range of applications and generalization capabilities [25]. ANNs are being increasingly utilized in aquatic sciences, providing useful models for water quality monitoring [26–33], and the prediction of wave heights [34–37]. One of the primary attributes of ANNs are their ability to provide solutions to complex multivariate problems. In the current study, the desired outcome is to take a VSF measurement which is influenced by multiple scattering errors as input, and find the "true" VSF unaffected by multiple scattering errors representing what would be measured in a single scattering regime. To accomplish this, we recognized the multivariate capabilities of ANNs and developed a feedforward neural network to correct multiple scattering errors unique to LISST-VSF measurements. To train the ANN, a large set of varied data containing both VSFs with multiple scattering errors and the corresponding VSF unaffected by multiple scattering errors was needed. As such experimental datasets are not readily available on a large scale, a large training dataset was generated using 154 different phase functions representative of natural assemblages of aquatic particles and the previously developed Monte Carlo model of the LISST-VSF instrument [17]. In this study, we present comprehensive analysis of the ANN model development and testing to reliably correct for multiple scattering errors in VSF measurements made by the LISST-VSF instrument. We present results of ANN testing with two contrasting natural samples prepared using dilutions to produce various particle concentrations in a laboratory environment and we examine the application of the ANN to a large dataset of LISST-VSF measurements from natural seawater samples spanning highly turbid coastal waters to clear oceanic waters.

2. Methods

A description of the ANN is given in Section 2.1, which includes how the ANN works and how it is structured. The dataset used for training the ANN is presented in section 2.2, followed by a description of how data was treated for use in the ANN (Section 2.3). Details of the training process, such as training algorithm and loss function, are given in Section 2.4. Finally, Section 2.5 describes how the ANN was evaluated and the datasets used for this purpose.

Note that all optical properties presented in this study refer to the properties of the particles in the water, i.e., scattering and absorption by pure water is not included. Also, the primary LISST-VSF used in this study refers to a single wavelength light source of 515 nm, meaning that all quantified properties refer to the value at this particular wavelength. Some data presented also refer to measurements from a different LISST-VSF instrument which utilizes a 532 nm light source, however no distinction is made for the proposes of this study and expected differences between optical measurements with 515 nm or 532 nm light source are assumed minor.

2.1. ANN structure

The ANN employed in this study was a multi-layer feedforward neural network containing five layers in total (see Fig. 1) [38]. The model was built in Python using Keras [39], which runs on top of TensorFlow [40]. The first layer is called the input layer and simply contains the input data which is a measured VSF. The input layer is followed by three hidden layers, and the final layer is the output layer containing the VSF corrected for multiple scattering errors. All nodes in one layer are connected to all nodes in the neighboring layers, but the connections only work in the forward direction. After the input layer, each node in a following layer is calculated as a function of every node in the previous layer. For instance, the first node in the first hidden layer in Fig. 1 is calculated as

$$h_1^1 = F(W_{11}^1 x_1 + W_{21}^1 x_2 + ... + W_{n1}^1 x_n + b_1^1). \tag{1}$$

Here, x_i is the input VSF at measurement angle i, the weights W_{ij}^1 describes how much the node j in the hidden layer depends on node i in the input layer, i.e., the strength of the connection. The superscript in W_{ij}^1 refers to which layers the weights connect to (first hidden layer is 1, output layer is 4), and the superscript in h_1^1 refers to which hidden layer the node belongs to (see Fig. 1). The term b_1^1 is the bias for node h_1^1, which is a constant that is added to the product of inputs and weights. The function F is called an activation function, of which there are several alternatives to choose from. The hyperbolic tangent function is commonly used and was found to produce good results for the ANN developed in this study. Thus, Eq. (1) becomes

$$h_1^1 = \tanh(W_{11}^1 x_1 + W_{21}^1 x_2 + ... + W_{n1}^1 x_n + b_1^1). \tag{2}$$

The nodes in the first hidden layer are fed forward as inputs to the second hidden layer, and so on until reaching the output layer. The output layer is calculated without the activation function, so that the output corresponding to the first angle in the corrected VSF is given as

$$y_1 = W_{11}^4 h_1^3 + W_{21}^4 h_2^3 + ... + W_{r1}^4 h_r^3 + b_1^4. \tag{3}$$

The ANN developed in this study takes 158 inputs, corresponding to the VSF measured for angles 0.09 – 140°. The exclusion of data points above 140° is due to deviations between measured and simulated VSFs in this angular range for large scattering coefficients ($b > 20\,\mathrm{m}^{-1}$) [17]. The output layer contains 168 nodes, corresponding to the corrected VSF over the angles 0.09 – 150° (same as the LISST-VSF). Thus, the ANN estimated VSF for angles 140 – 150° is predicted based on the measured VSF at angles 0.09 – 140°.

The size of the network was determined based on a validation process, where the available data was split into a training dataset and a validation dataset. Due to limited data, only 20% of the data was used for validation, while the remaining 80% was kept for training. The phase functions used for validation were randomly selected, but were the same for all configurations. A total of 15 configurations of the ANN structure were tested, varying in both the number of layers and the number of nodes per layer. Based on the results, a network size of three hidden layers with 168 nodes per layer was chosen (same number of nodes as the output layer). The results from the validation process is presented in Section 3.1.

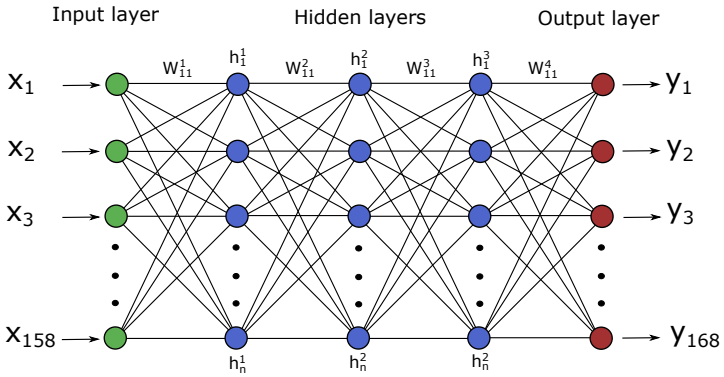

Fig. 1. Illustration of the ANN with a size of three hidden layers with n nodes per layer. The input values X are the original VSF values for angles 0.09 – 140°, nodes h in the hidden layers are calculated as function of W and b (weights and biases), and the output values Y are the estimated VSF values for angles 0.09 – 150°.

2.2. Training dataset

As mentioned, large experimental VSF datasets with high quality reference measurements are not readily available. However, by utilizing the previously developed Monte Carlo simulation of the LISST-VSF instrument, a large training dataset can be generated [17, 18]. The bulk of the training data consists of VSFs simulated with Henyey-Greenstein (HG), Fournier-Forand (FF), and two-term Reynolds-McCormick (TTRM) phase functions, which covers a wide range of phase functions with resemblance to those found in seawater [2, 7, 11]. In addition, some VSF were simulated with phase functions obtained from measurements with natural sample, so that more natural VSFs are represented in the training dataset. A total of 154 different phase functions were used, of which 18 were from natural samples. Each was simulated with scattering coefficients ranging from $b = 0.05 - 50\,\mathrm{m}^{-1}$, resulting in a total of 1386 VSFs. The particulate phase functions $\tilde{\beta}$ have backscattering ratios in the range $b_b/b = 0.0003$-0.3 with a median of

0.034, and are shown in Fig. 2.

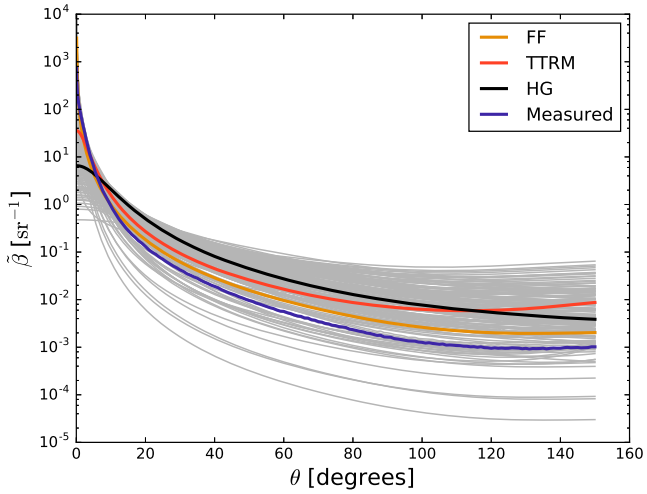

Fig. 2. All phase functions $\tilde{\beta}$ used to generate training data are shown in grey. A representative phase function for each type are shown for Fournier-Forand (yellow), two-term Reynolds-McCormick (Orange), Henyey-Greenstein (black), and from measurement (purple).

By using several types of phase functions, one obtains a training set of large variety. Typically, the TTRM and HG phase functions plateau when approaching 0°, while the FF phase functions continue to rise with decreasing angle (see Fig. 2). The TTRM phase function also offers more variety in terms of shape in the backwards direction, where the FF and HG phase functions are much more limited. The phase functions obtained from measurements of natural seawater samples generally provide more realistic training data, accounting for variations that are not adequately represented by the HG, FF, and TTRM phase functions.

2.3. Data transformation and standardization

Since VSF data can vary with several orders of magnitude, it was preprocessed to make the data suitable for application with the ANN using a log transformation and statistical standardization [41, 42]. The first step is to transform the values to the $\log_e$ domain, so that

$$x'_{i,j} = \log_e(x_{i,j}), \tag{4}$$

where $x_{i,j}$ is the original VSF data at input node i for sample j. Then the data is standardized by subtracting the mean and dividing by the standard deviation for each individual node in the input layer, giving

$$x''_{i,j} = \frac{x'_{i,j} - \overline{x'_i}}{s_{x'_i}}. \tag{5}$$

Here, $x''_{i,j}$ is the transformed and standardized value for input node i in training sample j, $\overline{x'_i}$ is the mean value of input node i over all training samples, and $s_{x'_i}$ is the standard deviation. This transformation gives a distribution of $x''_{i,j}$ with a mean of zero and a standard deviation of one. The true values $\hat{y}$ are transformed following the equivalent formula, and the output values y from the ANN can be transformed back to regular values by applying the reverse transformation.

2.4. Training the ANN

The aim of the training process is to compute weights W_{ij}^k that minimizes the deviation between the output of the ANN and the true values, and is expressed by a loss function. Here, the mean squared error (MSE) function, which is a commonly used loss function, was found to work well,

$$\text{loss} = \text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2. \tag{6}$$

Here, y_i is the predicted values, $\hat{y}_i$ is the true values and N is the total number of outputs in the training set, i.e. the number of VSFs times the number of measurement points per VSF. The MSE was calculated for the transformed data, and is likely insensitive to specific angular regions. The ANN was trained using the Adam algorithm, which is a built-in optimizer in Keras. The algorithm minimizes the loss with respect to its parameters W and b (weights and biases) trough stochastic gradient descent [43].

2.5. Evaluating the ANN

It is not possible to directly evaluate the ANN's performance on LISST-VSF measurements of turbid samples influenced by multiple scattering errors, since we do not know the "true" VSF (i.e., the VSF in the single scattering regime unaffected by multiple scattering errors). The preferred way to evaluate the ANN's performance is to obtain LISST-VSF measurements with both a turbid water sample and the same sample diluted to particle concentrations unlikely to induce multiple scattering errors. At a sufficient dilution, the measured VSF will have negligible multiple scattering error and the expected VSF for higher concentration samples can be calculated by multiplying with an appropriate dilution factor. The dilution-corrected VSF can thus be used as "true", or expected, VSF for evaluating the ANN performance. This has been done for two different samples, resulting in two sample sets that contain the original VSF measurements at different dilutions, in addition to the dilution-corrected VSFs. These sample sets are used for evaluation of the ANN in Section 3.2.

Further evaluation of the ANN model was carried out with 98 measurements on natural samples for which we do not have dilution corrections. The dataset presented in Section 3.3 contain a large variation of samples, with measured attenuation coefficients ranging from $c = 0.64 - 21.82\,\text{m}^{-1}$ and backscattering ratios ranging from $\tilde{b}_b = 0.006 - 0.027$. The dataset consist of LISST-VSF measurements collected from various places: Coastal waters of Southern California and Northern Alaska [13]; Coastal waters around Svalberd, including Rijåfjorden, Hinlopen and Isfjorden [15]; Norwegian fjords, including Gaupnefjorden [15] and an Emiliana Huxleyi algae bloom in Hardangerfjorden; Coastal measurements in Storfjorden, Svalbard [44].

Assuming that the LISST-VSF attenuation measurement is accurate, the performance of the ANN can be evaluated through the relation between b and measured c. Ideally, one would also have absorption measurements obtained simultaneously with a different instrument, so that a comparable scattering coefficient could be obtained indirectly through the relation $c = a + b$. However, such measurements were not performed for the data presented here.

Percentage deviation was used to quantify differences between values obtained from original uncorrected measurements or values obtained from application of the ANN, and reference values obtained from dilution correction. The percentage deviation is defined as

$$\frac{X_m - X_r}{X_r} \times 100\%, \tag{7}$$

and describes how much the value X_m deviates from the reference value X_m. The values X_m and X_r are specified in the text where Eq. (7) is applied.

3. Results and discussion

This section is divided into three parts. First, the results from the validation process used to determine the size of the network is presented. Then, evaluation of the ANN is performed with dilution-corrected VSFs obtained for two different sample sets. Finally, further evaluation is performed with VSF measurements obtained from natural samples, for which dilution-corrected values are not available. For this section, all scattering and attenuation measurements are performed with the LISST-VSF instrument, unless otherwise is stated.

3.1. Validation and determination of network size

An appropriate ANN size was determined by testing 15 different ANN configurations with hidden layers ranging from 1 to 5, and with either 84, 168 or 252 nodes per hidden layer. Each model was trained for 100,000 iterations and the results are presented in Fig. 3.

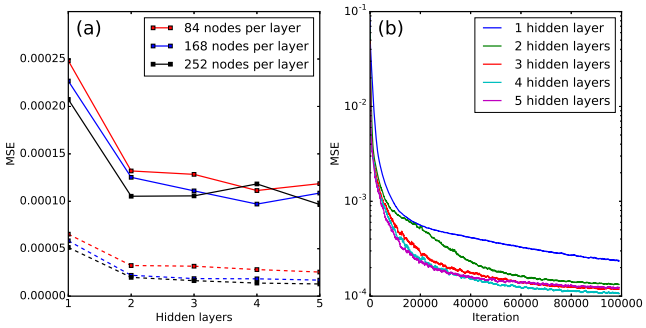

Fig. 3. (a) Mean squared error (MSE) after 100,000 iterations plotted as a function of hidden layers for an ANN with 84, 168 and 252 nodes per layer. The MSE is shown for both the validation set (solid lines) and the training set (dashed lines). (b) MSE of the validation set plotted as a function of iterations for ANNs with a number of hidden layers ranging from 1 to 5, and with 168 nodes per hidden layer.

From Fig. 3(a), it can be seen that the MSE of the training set (dashed) generally decreases with both increasing number of hidden layers and number of nodes per layer. For the validation set (solid), the largest improvement in performance is observed when going from 1 to 2 hidden layers. For both 84 and 168 nodes per layer, the MSE has a small downwards trend going from 2 to 4 layers, with a slight increase at 5 hidden layers. This could be a sign of overfitting, where the amount of parameters (weights and biases) becomes so large that the solution for the training data may not be a good general solution for data not included in the training. Additional evidence of overfitting is observed in Fig. 3(b), where it can be seen that the 5 hidden layer ANN is only the top performer in the early stages of the training. The 5 hidden layer ANN displays rapid improvement in MSE during the first 30,000 iterations at which point the 4 hidden layer network begins to outperform, followed by the 3 hidden layer network at approximately 55,000 iterations (Fig. 3b). Similar plots were made for the 84 and 252 nodes per layer configurations (not presented here), displaying similar trends as those seen in Fig. 3(b). For the 252 nodes per layer configurations, similar signs of overfitting were observed with 4 hidden layers. Based on these observations, an ANN with 3 layers and 168 nodes per layer was chosen as our model. The chosen ANN size results in low risk of overfitting, while providing sufficient complexity for generalization.

3.2. Testing with diluted samples

The results from the dilution-corrected sample sets are presented in Fig. 4, showing both the measured, dilution-corrected (expected), and ANN estimated VSF. Here, only the lowest concentration and the highest concentration samples are shown. The first sample set (Fig. 4a) was collected from Los Peñasquitos Lagoon, San Diego County, California, and will be referred to as the lagoon samples. The second sample set (Fig. 4b) was collected from glacial meltwater near King George Island, Antarctica [13], and will be referred to as the meltwater samples.

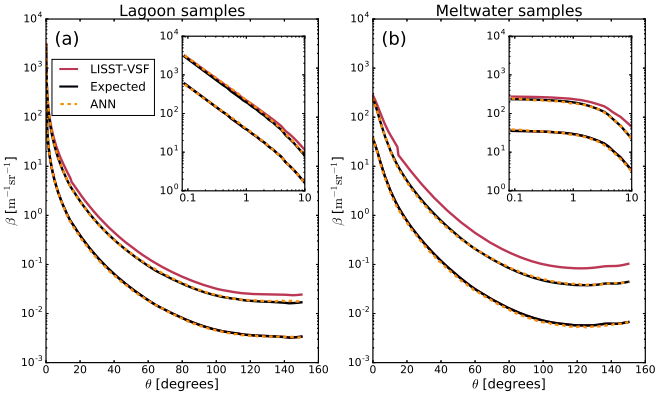

Fig. 4. VSFs obtained from LISST-VSF measurement (red), expected VSFs from dilution correction (black), and ANN estimated VSFs (dashed yellow). The insets show the log-log plot of the VSFs for angles $\theta \leq 10$. Highest (top black curve) and lowest concentration (lower black curve) samples shown here. (a) Lagoon samples. Dilution factor of 10. (b) Meltwater samples. Dilution factor of 6.6.

The ANN estimation of the true VSFs agrees well with the dilution corrected VSFs (Fig. 4). The two sample sets have noticeable differences in both the forward (<10°) and backward scattering angles, showcasing the generalization capabilities of the ANN. The ANN estimates deviate slightly from the expected VSFs, most noticeable for the lowest concentration sample in the meltwater samples where the ANN estimate is somewhat lower than the expected VSF. Here, the measured scattering coefficient is $b = 1.77\,\mathrm{m}^{-1}$, while the ANN's estimate is $b = 1.62\,\mathrm{m}^{-1}$. According to previous investigations into the effect of multiple scattering, the percentage error in the scattering coefficient at $b = 1.62\,\mathrm{m}^{-1}$ is likely around 10% [18]. In this case, the ANN estimate of b is about 9% lower than b measured by the LISST-VSF, which agrees with expectation. The multiple scattering errors observed in the LISST-VSF measurement at the lowest particle concentration will naturally propagate into the estimations of expected VSF at higher particle concentrations, partly explaining why the ANN estimated VSF is lower than the dilution corrected VSF for the meltwaters. For the lagoon samples, this problem is less severe, as the ANN estimated scattering coefficient is only $b = 0.99\,\mathrm{m}^{-1}$, which is likely to contain a multiple scattering error of about $\sim 3\%$ [18].

The performance of the ANN can be quantified by comparing the scattering coefficients b and backscattering coefficients b_b obtained from the original (uncorrected) LISST-VSF measurement, dilution-corrected (expected) LISST-VSF measurement, and VSF estimated from application of the ANN to the original LISST-VSF measurement. In Fig. 5, the scattering and backscattering coefficients are plotted against the expected values from the dilution correction for both sample sets. Here, the expected values are plotted with an adjustment for expected multiple scattering error in the LISST-VSF measurements with lowest particle concentration, accounting for $\sim 3\%$ and $\sim 10\%$ error in the lagoon and meltwater samples, respectively.

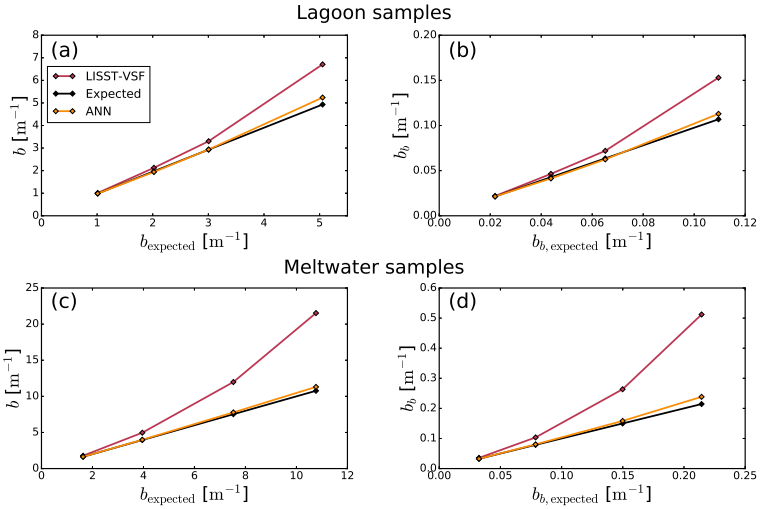

Fig. 5. Scattering coefficients and backscattering coefficients plotted against the expected values for the lagoon samples (a and b), and the meltwater samples (c and d). Values measured by LISST-VSF (red), expected from dilution correction (black, 1:1 line), and estimated by the ANN (yellow).

For increasing scattering coefficients, the measured scattering and backscattering coefficients deviate increasingly from the expected values for both sample sets, as seen in Fig. 5. This is most noticeable for the meltwater samples (see Fig. 5(c)), where the highest concentration sample has a measured scattering coefficient of $b = 21.5\,\text{m}^{-1}$, which is more than twice the expected value of $b = 10.6\,\text{m}^{-1}$. On the other hand, the ANN estimates a scattering coefficient of $b = 11.1\,\text{m}^{-1}$, corresponding to a reduction in percent deviation from 102.8% to 4.7%. Here, the percentage deviation is calculated according to Eq. (7), where the reference value X_r is b expected from dilution correction, while X_m is b obtained from LISST-VSF measurement and from application of ANN, respectively. In general, the ANN estimates are very similar to the expected values, with both sample sets showing an increase in deviation at the highest concentration for both b and b_b. The largest deviation is observed in the backscattering coefficient for the highest concentration sample in the meltwater samples. Here, the percentage deviation is 13%, which is still a large improvement from the deviation of 138.8% in the measured b_b.

There are several possible reasons for the observed trend in deviation between expected and ANN estimated VSF. For these samples specifically, the dilution factor is considered to be a significant source of uncertainty. For the meltwater samples, the dilution was done subsequently, meaning that any error in the first dilution would propagate to the next dilution. On the other hand, the samples in the lagoon samples were diluted individually. These differences in dilution methodology could explain why the ANN estimate for the lagoon samples is observed to be both below and above the expected value (random error), while the ANN estimate for the meltwater samples is always above the expected value (systematic error). Also, due to increased

measurement uncertainty in the very near forward direction ($< 1°$), measurement results in this angular region were treated using extrapolation and smoothing routines. In a previous study, we have shown that the multiple scattering error is dependent on the shape of the VSF [18]. Specifically, the extreme forward part seems to be of special importance. Hence, the VSF shape resulting from extrapolation and smoothing at $\theta < 1°$ may be slightly inconsistent with the multiple scattering error in these measurements, resulting in deviations between the expected and ANN estimated VSFs. Furthermore, the results presented here are particularly sensitive to errors in the lowest concentration samples, as the expected VSF is directly calculated from these.

Other possible reasons to the observed deviation are of a more general form. The ANN is trained on simulated data, meaning that the ANN can only be as good as the simulation. While the simulation has been shown to be generally accurate, there are noticeable deviations at large scattering coefficients [17]. Another likely reason is that these samples are not adequately represented in the training data. Only 18 of the 154 phase functions used for training are obtained from measurements, meaning that the majority of the training set is obtained from artificial phase functions, e.g. Henyey-Greenstein and Fournier-Forand. By including more VSFs obtained from measurements of natural samples, the performance of the ANN is very likely to improve. However, due to limited available VSF measurements, data for training and data for testing had to be balanced. As more *in situ* data becomes available, additional samples can be added to the training dataset. The final factor to consider is uncertainty in the measured data, where all individual sample measurements contain errors of varying degree, in addition to multiple scattering errors [11, 15].

3.3. Testing with measurements of natural seawater samples

A large dataset consisting of 98 VSFs obtained from measurements on natural samples are used for further evaluation of the ANN. In Fig. 6(a), both the measured and ANN estimated b is plotted against measured c. The same data is also presented in Fig. 6(b), but here plotted against the ANN estimated b.

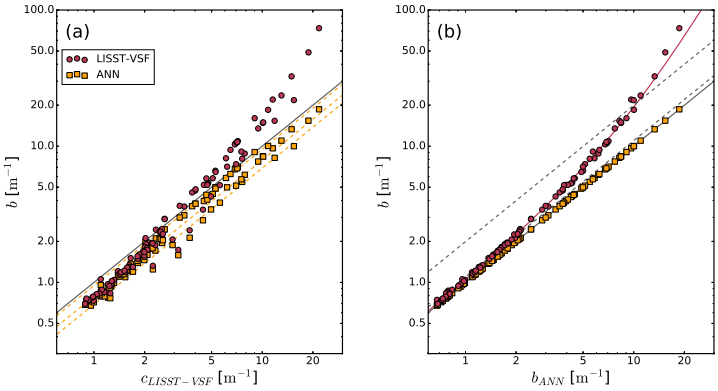

Fig. 6. (a) Measured (red) and ANN estimated (yellow) scattering coefficient b plotted against the measured attenuation coefficient c. The 1:1 line is plotted in solid grey, and the yellow dashed lines represent the median ratio of $b/c = 0.79$, together with the 10th percentile ($b/c = 0.69$) and 90th percentile ($b/c = 0.97$). (b) Measured (red) and ANN estimated (yellow) scattering coefficient b plotted against the ANN estimated scattering coefficient. The dashed grey lines show the 10% and 100% deviation between measured and ANN estimated b. The 1:1 line is plotted in solid grey.

From the relation $c = a + b$, we must always have $b < c$ to have a positive absorption coefficient. Thus, all measurement points plotted in Fig. 6(a) should be below the 1:1 line (grey). However, this is not the case for many of the LISST-VSF measurements, where $b > c$ can be observed for attenuation coefficients $c > 2\,\mathrm{m}^{-1}$. On the other hand, the ANN estimates of b are always below the 1:1 line. The ANN estimated b follows a linear trend with a median single scattering albedo b/c of 0.79. The observed variations around the median are expected due to natural variability in the optical properties of natural assemblages of marine particles. In Fig. 6(b), the same data are plotted as a function of the ANN estimated b. Here, it is observed a clear relationship between measured and ANN estimated b, where the percentage deviation follows a trend that can be approximated by a power law fit (see Fig. 7(a)).

The estimated adjustment after applying the ANN can be assessed by looking at the deviation between the original LISST-VSF measurement and ANN estimates. In Fig. 7, the percentage deviation is plotted for the scattering coefficient b, backscattering coefficient b_b and backscattering ratio $\tilde{b}_b = b_b/b$. The latter may be of special interest, as $\tilde{b}_b$ describes the VSF shape, rather than magnitude. Here, the percentage deviation is calculated according to Eq. (7), where X_r is the ANN estimate, and X_m is the measured value.

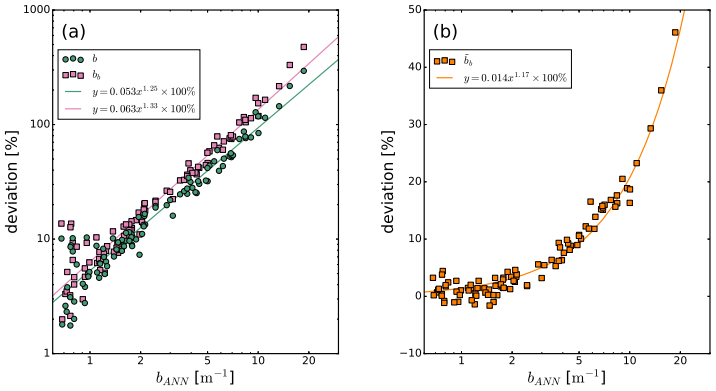

Fig. 7. Percentage deviation plotted (semi-log) against ANN estimated b for (a) the scattering coefficient b (green) and backscattering coefficient b_b (pink), and (b) the backscattering ratio (orange).

As seen in Fig. 7(a), the deviations in both b and b_b follow a trend that can be approximated by a power law fit, where the deviation in b_b is larger than for b. From the trend lines, the percentage deviation in b reaches 10% at approximately $b = 1.7\,\mathrm{m}^{-1}$ and 100% at $b = 10.5\,\mathrm{m}^{-1}$, while for b_b it is 10% at $b = 1.4\,\mathrm{m}^{-1}$ and 100% at $b = 8.0\,\mathrm{m}^{-1}$. These results are consistent with our previous study on multiple scattering errors in LISST-VSF measurements [18]. The results presented in Fig. 7(a) also illustrate the dramatic increase in measurement error when the scattering coefficient becomes large ($b > 10\,\mathrm{m}^{-1}$). The highest concentration sample in this dataset has an ANN estimated scattering coefficient of $b = 18.7\,\mathrm{m}^{-1}$, and the percent deviation is 294% for b and 476% for b_b.

From Fig. 7(b), the deviation in the backscattering ratio $\tilde{b}_b$ is less severe than that observed for b and b_b. Here, the deviation is 10% at $b \approx 5\,\mathrm{m}^{-1}$, reaching 46% for the highest concentration sample at $b = 18.7\,\mathrm{m}^{-1}$. The data points generally follow an exponential fit, but as with b and b_b, the fit is slightly too high for $b < 2\,\mathrm{m}^{-1}$. Here, a positive deviation in $\tilde{b}_b$ means that the increase in the measured VSF β due to multiple scattering is relatively larger in the backwards direction than in the forward direction. This is also true for the phase function, which is calculated as $\tilde{\beta} = \beta/b$.

4. Summary and conclusion

In this study, we have developed an artificial neural network (ANN) with the aim of correcting multiple scattering errors in VSFs measured by the LISST-VSF instrument. It is a traditional feed forward network, containing 3 hidden layers with 168 nodes layer. The ANN estimates the true VSF taking only the measured VSF as input, eliminating the necessity for additional measurements. Training data was generated using a previously developed Monte Carlo simulation of the LISST-VSF instrument, verified by laboratory measurements [17]. A large and varied training set was constructed, consisting of VSFs simulated using Henyey-Greenstein, Fournier-

Forand, and two-term Reynolds-McCormick phase functions, in addition to some phase functions obtained from *in situ* measurements. A total of 154 unique phase functions were simulated with 9 different scattering coefficients ranging from $b = 0.05 - 50\,\mathrm{m}^{-1}$, resulting in a training set consisting of 1386 VSFs.

A benchmark test of the ANN was performed using measurements with dilution corrected samples. The results showed that the VSFs estimated by the ANN were very similar to the expected VSFs in both shape and magnitude. The performance of the ANN was evaluated in terms of the scattering and backscattering coefficients. Both sample sets showed similar trends, where the deviation between LISST-VSF measurement and expected coefficients increased significantly with increasing particle concentration. On the other hand, the ANN estimates were very similar to the expected values, with a slight increase in deviation with increasing concentration. For the highest concentration sample, the expected scattering coefficient was $b = 10.6\,\mathrm{m}^{-1}$. In this case, applying the ANN correction reduced the scattering coefficient from a measured value of $b = 21.5\,\mathrm{m}^{-1}$ to the ANN estimated value of $b = 11.1\,\mathrm{m}^{-1}$, equivalent to a reduction in error from 103% to 5%.

Further testing was performed with a large database of natural seawater samples from various oceanic environments, for which measurements were made with original samples without any dilutions. The sample set contained 98 individual measurements with measured attenuation coefficients ranging from $c = 0.64 - 21.82\,\mathrm{m}^{-1}$. Here, the ANN was evaluated by calculating the scattering coefficients for the measured and ANN estimated VSF, and comparing them to the attenuation coefficients measured by the LISST-VSF. From the relation $c = a + b$, the expected b must be lower than the measured c. However, for samples with an attenuation coefficient greater than $c = 4\,\mathrm{m}^{-1}$, the measured b was consistently larger than the measured c, while instances of $b > c$ were observed at attenuation coefficients as low as $c = 2\,\mathrm{m}^{-1}$. The ANN was effective in reducing this discrepancy, with all estimates satisfying the criterion $b < c$. Furthermore, reductions in VSF following application of the ANN were more significant for the backscattering coefficient as compared with the scattering coefficient, typically resulting in a reduction of the backscattering ratio of about 5–20% for $b = 2 - 10\,\mathrm{m}^{-1}$.

The ANN developed in this study has been shown to correct multiple scattering errors in VSF measurements collected by the LISST-VSF for a wide range of phase functions and scattering coefficients. While undesirable multiple scattering errors arise for scattering coefficients $b > 2\,\mathrm{m}^{-1}$, the ANN provides reduced uncertainty in derived optical properties far above $b = 2\,\mathrm{m}^{-1}$. This significantly increases the turbidity range the LISST-VSF can reliably operate in, allowing for more accurate measurements in particle rich waters such as phytoplankton blooms, and coastal environments including river outlets, harbors and glacial meltwaters.

Acknowledgements

Data from Storfjorden, Svalbard were collected and processed in the project Useful Arctic Knowledge (project no. 274891), and made freely available by NERSC under projects INTAROS (GA No. 727890) and Norwegian Scientific Data Network (project no. 245967). The second author (DK) received support from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Marie Skłodowska- Curie grant agreement No. 101034309. Measurements from Hardangerfjorden and Gaupnefjorden were collected as part of the EcoSens project (project no. 303190).

Disclosures

The authors declare no conflicts of interest.

Data Availability Statement

The artificial neural network (ANN) will be made available upon acceptance. Other data, such as the training and validation data are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. P. J. Werdell, L. I. W. McKinna, E. Boss, S. G. Ackleson, S. E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, D. Stramski, J. M. Sullivan, M. S. Twardowski, M. Tzortziou, and Z. Xiaodong, "An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing," Prog. Oceanogr. **160**, 186–212 (2018).
2. J. Chowdhary, P.-W. Zhai, E. Boss, H. Dierssen, R. Frouin, A. Ibrahim, Z. Lee, L. A. Remer, M. Twardowski, F. Xu *et al.*, "Modeling atmosphere-ocean radiative transfer: A pace mission perspective," Front. Earth Sci. **7**, 100 (2019).
3. Y. Agrawal, A. Whitmire, O. A. Mikkelsen, and H. Pottsmith, "Light scattering by random shaped particles and consequences on measuring suspended sediments by laser diffraction," J. Geophys. Res. Ocean. **113** (2008).
4. B. Cochenour, K. Dunn, A. Laux, and L. Mullen, "Experimental measurements of the magnitude and phase response of high-frequency modulated light underwater," Appl. optics **56**, 4019–4024 (2017).
5. Z. Zeng, S. Fu, H. Zhang, Y. Dong, and J. Cheng, "A survey of underwater optical wireless communications," IEEE communications surveys & tutorials **19**, 204–238 (2016).
6. C. D. Mobley, *Light and water: radiative transfer in natural waters* (Academic press, 1994).
7. T. Harmel, J. Agagliate, M. Hieronymi, and P. Gernez, "Two-term reynolds–mccormick phase function parameterization better describes light scattering by microalgae and mineral hydrosols," Opt. Lett. **46**, 1860–1863 (2021).
8. C. Supplis, J. Dauchet, V. Gattepaille, F. Gros, T. Vourc'h, and J.-F. Cornet, "Radiative analysis of luminescence in photoreactive systems: Application to photosensitizers for solar fuel production," Plos one **16**, e0255002 (2021).
9. C. D. Mobley, L. K. Sundman, and E. Boss, "Phase function effects on oceanic light fields," Appl. optics **41**, 1035–1050 (2002).
10. T. J. Petzold, "Volume scattering functions for selected ocean waters," Tech. rep., Scripps Institution of Oceanography La Jolla Ca Visibility Lab (1972).
11. M. Jonasz and G. Fournier, "Theoretical and experimental foundations light scattering by particles in water," (2007).
12. X. Zhang, R. H. Stavn, A. U. Falster, D. Gray, and R. W. Gould Jr, "New insight into particulate mineral and organic matter in coastal ocean waters through optical inversion," Estuarine, Coast. Shelf Sci. **149**, 1–12 (2014).
13. D. Koestner, D. Stramski, and R. A. Reynolds, "Characterization of suspended particulate matter in contrasting coastal marine environments with angle-resolved polarized light scattering measurements," Appl. Opt. **60**, 11161–11179 (2021).
14. D. Koestner, D. Stramski, and R. A. Reynolds, "Measurements of the volume scattering function and the degree of linear polarization of light scattered by contrasting natural assemblages of marine particles," Appl. Sci. **8**, 2690 (2018).
15. H. Sandven, A. S. Kristoffersen, Y.-C. Chen, and B. Hamre, "In situ measurements of the volume scattering function with LISST-VSF and LISST-200X in extreme environments: Evaluation of instrument calibration and validity," Opt. Express **28**, 37373–37396 (2020).
16. L. Hu, X. Zhang, Y. Xiong, and M.-X. He, "Calibration of the LISST-VSF to derive the volume scattering functions in clear waters," Opt. Express **27**, A1188–A1206 (2019).
17. H. S. Ugulen, H. Sandven, B. Hamre, A. S. Kristoffersen, and C. Sætre, "Analysis of multiple scattering errors in LISST-VSF volume scattering function measurements using Monte Carlo simulations and experimental data," Opt. Express **29**, 12413–12428 (2021).
18. H. S. Ugulen, H. Sandven, B. Hamre, A. S. Kristoffersen, and C. Sætre, "Efficient monte carlo simulation reveals significant multiple scattering errors in underwater angular scattering measurements," Opt. Express **30**, 10802–10817 (2022).
19. J. Piskozub, D. Stramski, E. Terrill, and W. K. Melville, "Influence of forward and multiple light scatter on the measurement of beam attenuation in highly scattering marine environments," Appl. Opt. **43**, 4723–4731 (2004).
20. H. C. van de Hulst, "Multiple light scattering," (1980).

21. E. Dan Hirleman, "Modeling of multiple scattering effects in fraunhofer diffraction particle size analysis," Part. & Part. Syst. Charact. **5**, 57–65 (1988).
22. J. Piskozub and D. McKee, "Effective scattering phase functions for the multiple scattering regime," Opt. express **19**, 4786–4794 (2011).
23. R. A. Maffione and D. R. Dana, "Instruments and methods for measuring the backward-scattering coefficient of ocean waters," Appl. Opt. **36**, 6057–6067 (1997).
24. D. Doxaran, E. Leymarie, B. Nechad, A. Dogliotti, K. Ruddick, P. Gernez, and E. Knaeps, "Improved correction methods for field measurements of particulate light backscattering in turbid waters," Opt. express **24**, 3615–3637 (2016).
25. O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," Heliyon **4**, e00938 (2018).
26. K.-S. Jeong, G.-J. Joo, H.-W. Kim, K. Ha, and F. Recknagel, "Prediction and elucidation of phytoplankton dynamics in the nakdong river (korea) by means of a recurrent artificial neural network," Ecol. Model. **146**, 115–129 (2001).
27. Y. Chebud, G. M. Naja, R. G. Rivero, and A. M. Melesse, "Water quality monitoring using remote sensing and an artificial neural network," Water, Air, & Soil Pollut. **223**, 4875–4887 (2012).
28. H. G. Kim, S. Hong, K.-S. Jeong, D.-K. Kim, and G.-J. Joo, "Determination of sensitive variables regardless of hydrological alteration in artificial neural network model of chlorophyll a: case study of nakdong river," Ecol. Model. **398**, 67–76 (2019).
29. J. Anmala and T. Venkateshwarlu, "Statistical assessment and neural network modeling of stream water quality observations of green river watershed, ky, usa," Water Supply **19**, 1831–1840 (2019).
30. H.-S. Yi, S. Park, K.-G. An, and K.-C. Kwak, "Algal bloom prediction using extreme learning machine models at artificial weirs in the nakdong river, korea," Int. journal environmental research public health **15**, 2078 (2018).
31. E. M. Alves, R. J. Rodrigues, C. dos Santos Corrêa, T. Fidemann, J. C. Rocha, J. L. L. Buzzo, P. de Oliva Neto, and E. G. F. Núñez, "Use of ultraviolet–visible spectrophotometry associated with artificial neural networks as an alternative for determining the water quality index," Environ. monitoring assessment **190**, 1–15 (2018).
32. D. Gebler, G. Wiegleb, and K. Szoszkiewicz, "Integrating river hydromorphology and water quality into ecological status modelling by artificial neural networks," Water research **139**, 395–405 (2018).
33. T. M. Tung, Z. M. Yaseen *et al.*, "A survey on river water quality modelling using artificial intelligence models: 2000–2020," J. Hydrol. **585**, 124670 (2020).
34. M. Deo and C. S. Naidu, "Real time wave forecasting using neural networks," Ocean. engineering **26**, 191–203 (1998).
35. O. Makarynskyy, "Improving wave predictions with artificial neural networks," Ocean. Eng. **31**, 709–724 (2004).
36. R. Kalra, M. Deo, R. Kumar, and V. K. Agarwal, "Rbf network for spatial mapping of wave heights," Mar. Struct. **18**, 289–300 (2005).
37. N. P. Juan and V. N. Valdecantos, "Review of the application of artificial neural networks in ocean engineering," Ocean. Eng. **259**, 111947 (2022).
38. D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," Chemom. intelligent laboratory systems **39**, 43–62 (1997).
39. F. Chollet *et al.*, "Keras," https://keras.io (2015).
40. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015). Software available from tensorflow.org.
41. G. J. Bowden, G. C. Dandy, and H. R. Maier, "Data transformation for neural network models in water resources applications," J. Hydroinformatics **5**, 245–258 (2003).
42. I. Ioannou, A. Gilerson, B. Gross, F. Moshary, and S. Ahmed, "Deriving ocean color products using neural networks," Remote. Sens. Environ. **134**, 78–91 (2013).
43. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).
44. H. Sandven, T. Petit, A. Stallemo, E. B. V. den Bergh, F. Klockmann, H. R. Økland, J. Knutsen, K. T. Galtung, M. H. Malin Lunde, N. Eryilmaz, and T. S. Sagen, "Optical property measurements collected in storfjorden, svalbard, during the uak 2020 cruise," https://doi.org/doi:10.21335/NMDC-NERSC-1741962272 (2020).

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230862421 (print)
9788230865088 (PDF)