Modelling Bio-Optical Properties of Singapore Waters: A Comparison between Field and Spectral Optimization-Retrieved Optical Properties

Ryan Tan[1*], Amihan Yson Manuel[1], Sandric C. Y. Leong[2], and Soo Chin Liew[1]

[1]Centre for Remote Imaging, Sensing and Processing, National University of Singapore, Singapore

[2]Tropical Marine Science Institute, National University of Singapore, Singapore

*tanryan@nus.edu.sg

Abstract : *Aquaculture in Singapore has been rapidly expanding in recent years in order to meet the increasing demand for seafood and to ensure Singapore's food security without heavy reliance on imports. However, variations in the water quality of Singapore waters and occurrences of harmful algae bloom events cause fish kills and hinder Aquaculture production. Therefore, it is imperative to develop ways to monitor such changes in water quality and bloom events. This paper analyses the preliminary development of bio-optical models that are representative of Singapore waters. Measurements of biological and optical properties of water such as absorption, backscattering, chlorophyll, and turbidity are conducted in situ at field stations around Singapore, and analyzed in the laboratory. These measurements are then used to develop preliminary bio-optical models and compared against synthetic data retrieved from Hydrolight. Our study showed that the results from our bio-optical models correspond well with the synthetic data generated from Hydrolight. With more data from in-situ measurements, it is possible to develop machine-learning algorithms to generate water-quality maps and even predict algae blooms in Singapore.*

Keywords: *bio-optical models, hydrolight, water-quality*

Introduction

Bio-optical models are empirical models that relate the intrinsic water optical properties (IOPs) such as absorption and backscattering to the relevant water quality parameters such as chlorophyll, CDOM and turbidity. The IOPs and water quality parameters can be accurately measured via in-situ and laboratory measurements, and used to develop the bio-optical models that are representative of the local coastal waters. However, in-situ measurements have limited spatial and temporal coverage. Satellite-based remote sensing may be used to estimate the IOPs of water over a large spatial extent. Water IOPs can be retrieved from satellite measured remote sensing reflectance of water using conventional spectral fitting techniques. In this paper, we use a radiative transfer code (RTC), Hydrolight (Mobley & Hedley, 2021), to generate

synthetic data of water reflectances. To ensure that the synthetic data is realistic and representative of the water types in our region, the bio-optical models constructed from field measurements are used as inputs to Hydrolight. We then retrieved water IOPs from the Hydrolight simulated synthetic reflectances using a spectral optimization method and compared the retrieved IOPs with the IOPs obtained from in-situ and laboratory measurements.

Study Areas

The study areas are the waters surrounding Singapore. Specially, field measurements have been conducted at 3 fieldwork zones: East Johor Strait (EJS), West Johor Strait (WJS), and Singapore Strait (SgS) as shown in the map in Figure 1. At each zone, there are 5 sampling stations. As of June 2024, we have conducted 24 sampling cruises along the EJS, WJS, and SgS since January 2024.

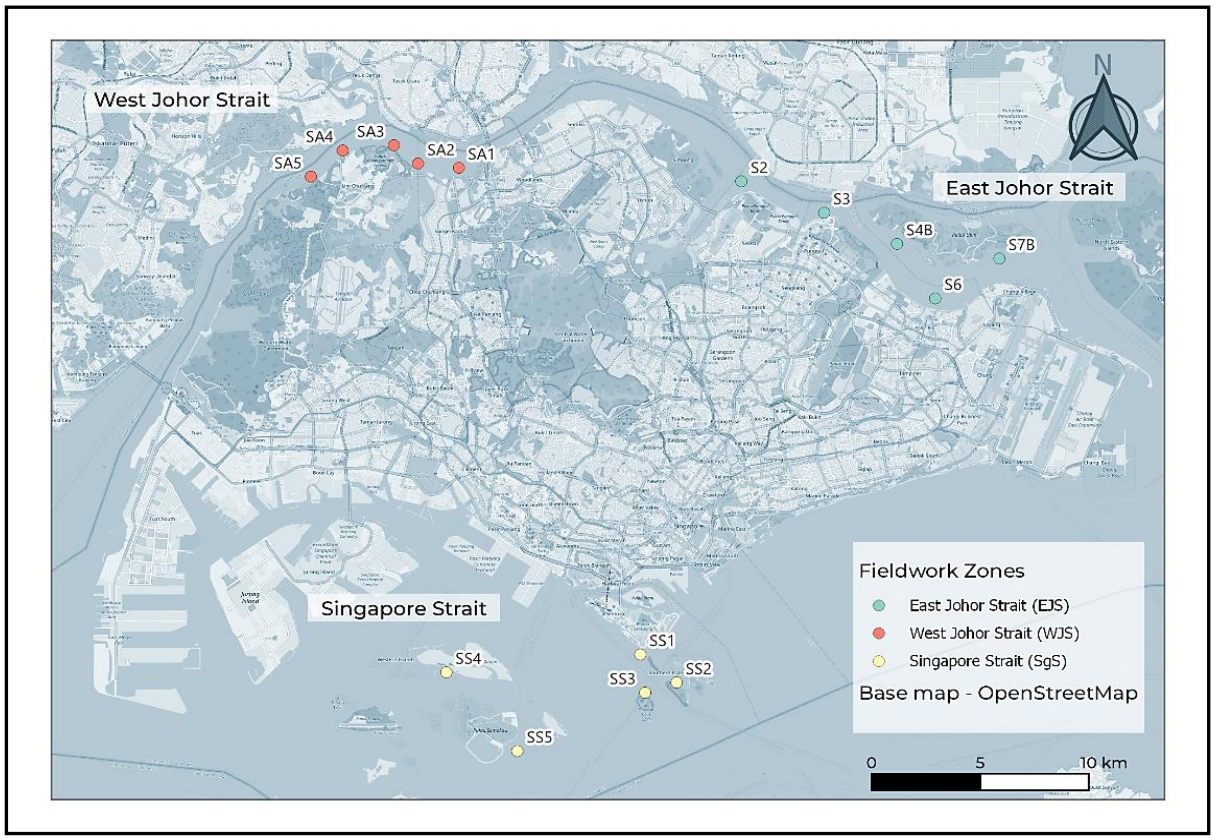

Figure 1: Study areas around Singapore waters

Methodology

a. Instruments and deployment

Water samples for absorption measurements at all sites for the three zones were collected using sampling equipment and kept in the dark until filtration was done within the same day of collection. Samples were frozen until analysis. CDOM samples were filtered using 0.22µm size filter unit and absorption were measured within the same day of collection. Absorption for CDOM, total particulates, and total detritus components were measured and estimated using the Shimadzu UV-2600i spectrophotometer by the quantitative filter technique (QFT) method (Chee Yew Leong & Taguchi, 2005). The analyses of chlorophyll-a (Chl-a) concentrations were measured using the same spectrometer after filtration with solvents (Ritchie, 2008). Backscattering measurements are obtained using WETLABS Eco BB9 at nine discrete wavelengths, namely 412nm, 440nm, 488nm, 510nm, 532nm, 595nm, 650nm, 676nm, and 715nm. BB9 is deployed directly on field sites by lowering it with a rope until the sensor is submerged in water. The instrument illuminates a volume of water using modulated LEDs and detects scattered light at an acceptance angle of 124^{o} from the source beam (Moore et al., 2005). Turbidity is measured using Eutech TN-100 Turbidimeter and the units are in Nephelometric Turbidity Units (NTU). Water samples from each station are collected in a vial for which a 850nm infrared light will then pass through the sample and the detected light scattered at 90^{o} will be internally calculated by the device to get NTU readings of the water samples.

b. Modelling equations

The absorption coefficient of detritus and CDOM is modelled by an exponential equation (Jerlov, 1968):

$$a_{dg}\,(\lambda) = Ge^{-S(\lambda - 440nm)}$$

where $G = a_{dg}\,(440nm)$. 440nm corresponds to approximately the midpoint of the blue waveband peak for which most algae species will have during their photosynthesis process (Kirk, 1994). S represents the spectral shape which is dependent on the type of detritus and CDOM. The absorption of phytoplankton can be modelled using the equation (Lee et al., 1999):

$$a_{ph}\,(\lambda) = P[a_0(\lambda) + \ln(P)a_1(\lambda)]$$

where $P = a_{ph}(440nm)$ is related to the concentration of Chlorophyll. $a_0(\lambda)$ and $a_1(\lambda)$ are basis functions normalized such that $a_0(440nm) = 1$ and $a_1(440nm) = 0$.

The backscattering coefficient of suspended particulate matter is usually modelled by an inverse power law:

$$b_{b_p}(\lambda) = X\left(\frac{532nm}{\lambda}\right)^{\eta}$$

where $X = b_{b_p}(532nm)$ is related to the concentration of particles. η also represents the spectral shape which is dependent on the type and size of suspended particles.

c. Hydrolight Simulations

Hydrolight simulations are performed using the absorption and backscattering measurements obtained from in situ and lab measurements as inputs. The Hydrolight parameters used for pure water IOPs are from Pope and Fry seawater coefficients (Pope & Fry, 1997) and Fournier-forand phase function (Fournier & Forand, 1994) for scattering. A semi-empirical sky model, RADTRAN-X (Gregg & Carder, 1990) is also used to compute the irradiances. Finally, we obtained the remote-sensing reflectances, $r_{rs}(\lambda)$ for all the stations for the wavelength range of 400nm to 720nm.

d. Retrieval of water IOPs from reflectance spectra

In the retrieval procedures, the remote sensing reflectance $r_{rs}(\lambda)$ is calculated using a semi-empirical model (Lee et al., 1999). The total absorption $a(\lambda)$ is the sum of the absorption coefficients of the separate components:

$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_{dg}(\lambda)$$

where $a_w(\lambda)$ is the absorption of pure water, and total backscattering $b_b(\lambda)$ is the sum:

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda)$$

where $b_{bw}(\lambda)$ is the backscattering of pure water. The semi-empirical model for $r_{rs}(\lambda)$ is calculated via the equations:

$$u = b_b/(a + b_b)$$

$$r_{rs} = (g_0 + g_1 u)u$$

where $g_0 = 0.089$ *and* $g_1 = 0.125$ (Lee et al., 2002). The IOPs, *G*, *P*, *X* and *S*, η are then estimated by fitting the Hydrolight generated reflectance to the semi-empirical model using the Sequential Least Squares Programming (SLSQP) (Kraft, 1988) by minimizing the root mean square error (RMSE) between the semi-empirical $r_{rs}(\lambda)$ and Hydrolight-derived $r_{rs}(\lambda)$.

Results and Discussion

a. Reflectance spectra from Hydrolight and semi-empirical model

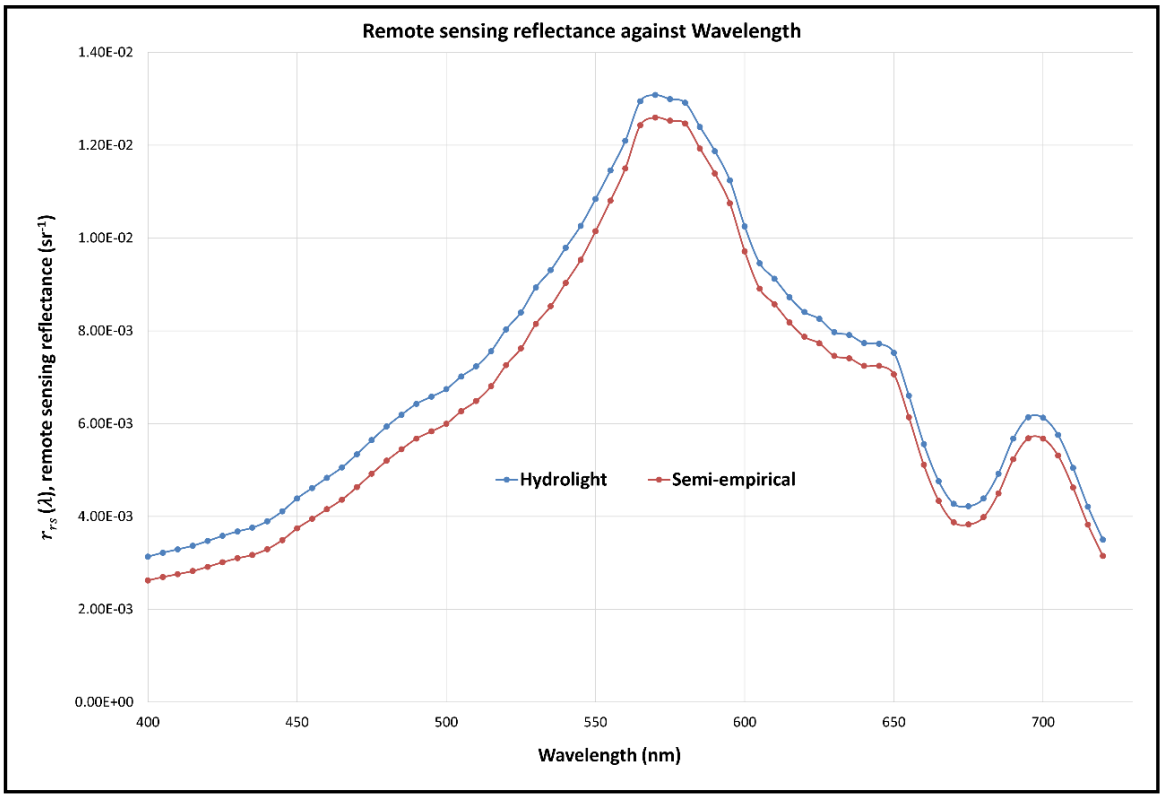

Figure 2: An example of field vs Hydrolight $r_{rs}(\lambda)$ obtained on from station S2

An example comparing the $r_{rs}(\lambda)$ obtained from semi-empirical model and Hydrolight on 4th April from station S2 is shown in Figure 2.

b. Absorption of CDOM and detritus, a_{dg} (λ)

The total data points obtained from the 24 sampling cruises across 6 months are 120 points. Using the field-measured a_{dg} (λ) , the averaged values for $G = a_{dg}$ $(440nm)$ across the 6 months are then obtained for the 15 stations. They are then plotted against the retrieved values of G obtained from spectral optimization as show in Figure 3 below. Based on the scatter plot of Retrieved *G* against Field *G*, we obtained a linear equation of y = 0.845x and a value of R^2 = 0.9711. Based on the scatter plot, we can see that in general, the retrieved values of *G* are close to the field values. However, there is an underestimation of *G* when the values are higher (> 1.0) which occur at stations within WJS where the CDOM and detritus are higher. A similar scatter plot was also obtained for the Retrieved *S* and Field *S* values for all the 15 stations as shown in Figure 4. The linear equation obtained is y = 1.0196x with R^2 = 0.9934. The retrieved *S* values are also compares quite well against field values as most of them lie close to the y = x line, albeit the values of retrieved *S* are slightly overestimated when compared to field *S*.

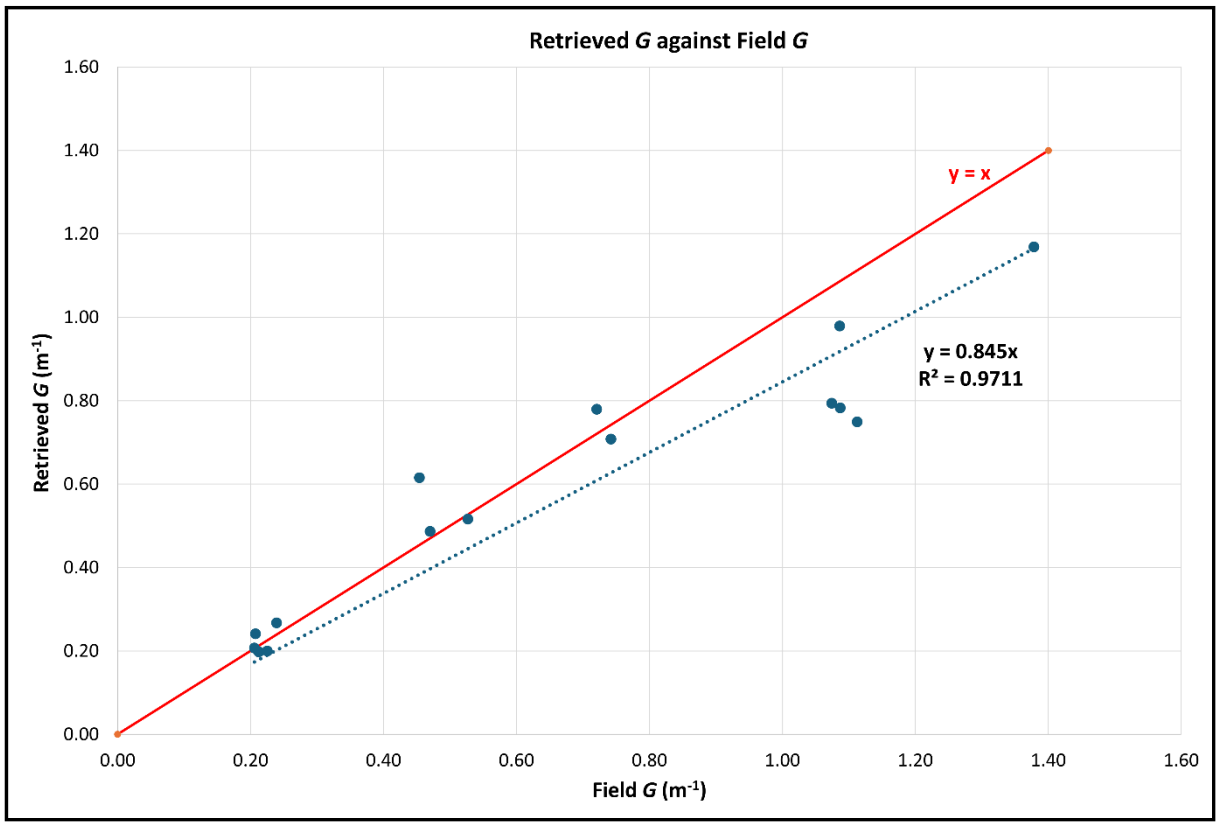

Figure 3: Scatter plot of Retrieved *G* against Field *G*

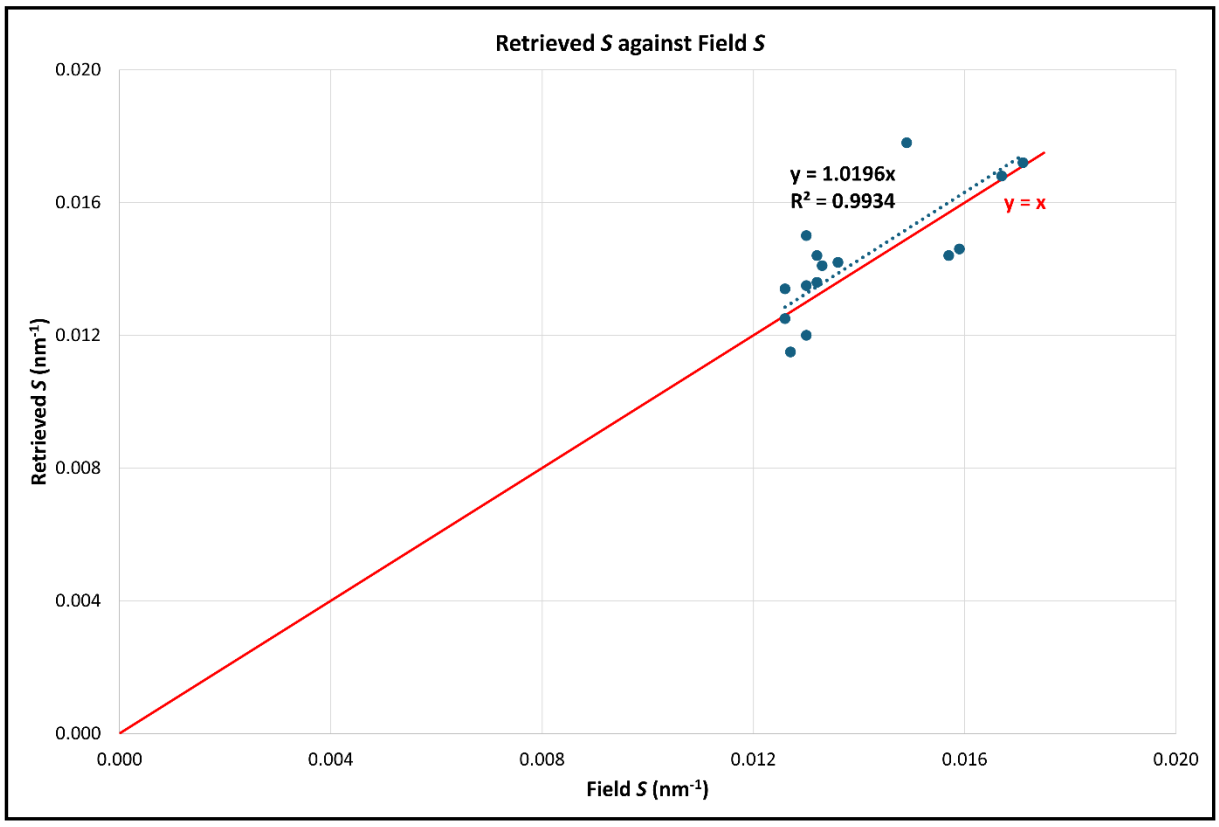

Figure 4: Scatter plot of Retrieved *S* against Field *S*

Parameter	R^2	RMSE	MAPE	MBE
G	0.9711	0.1618	14.171	-0.0698
S	0.9934	0.0012	6.889	0.0003

Table 1: Accuracy assessment of *G* and *S*

We also perform further accuracy assessments using RMSE, mean absolute percentage error (MAPE) and mean bias error (MBE) as shown in Table 1. For the values of *S*, the values are quite accurate with small RMSE (0.012) and MAPE (6.889) and a small overestimation of 0.0003 as we expect based on the plot in Figure 4. For the *G* values, the accuracy results shows larger deviations between the retrieved and field values of *G*. This is attributed to the stations in WJS where the values of CDOM and detritus are higher. The underestimation of *G* is also reflected from the negative MNE value of -0.0698.

c. Absorption of phytoplankton, a_{ph} (λ)

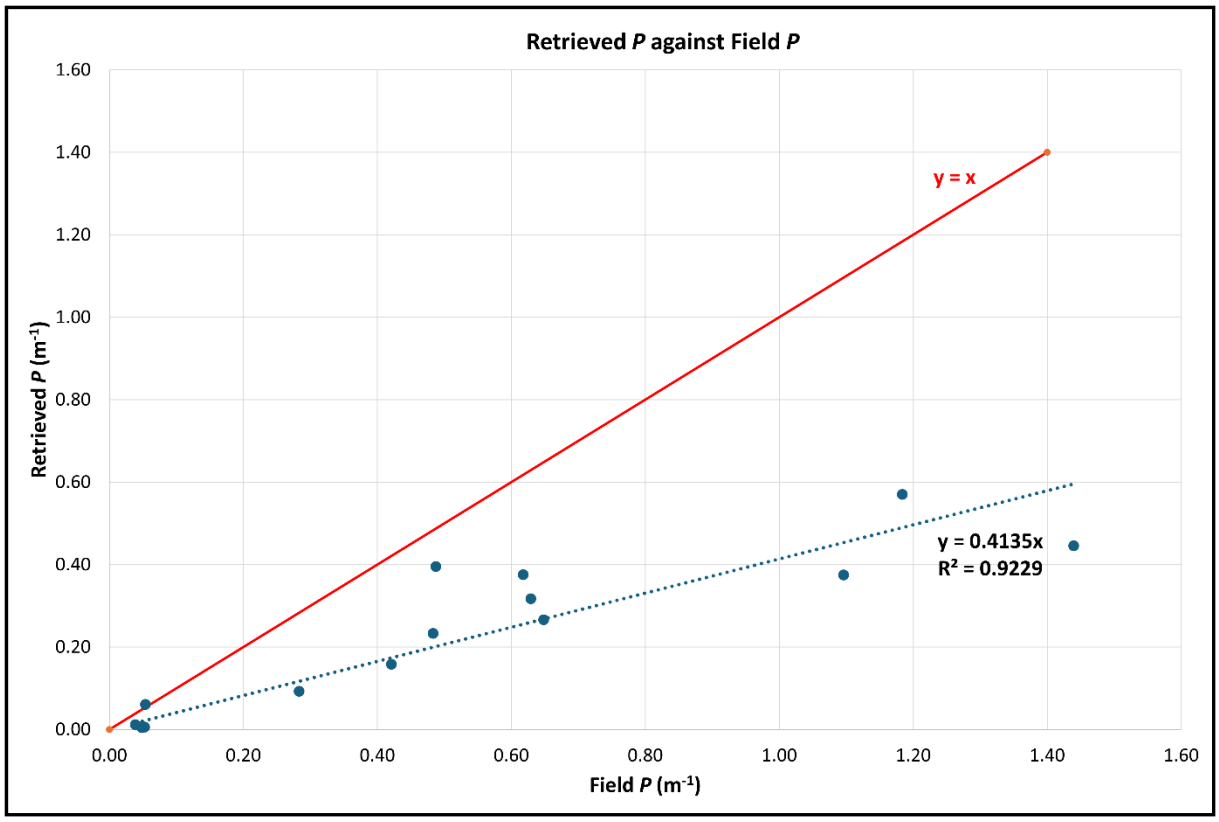

Figure 5: Scatter plot of Retrieved *P* against Field *P*

Parameter	R^2	RMSE	MAPE	MBE
P	0.9229	0.1618	14.171	-0.0698

Table 2: Accuracy assessment of *P*

For the absorption of phytoplankton, $a_{ph}(\lambda)$, we obtained the field values of $P = a_{ph}(440nm)$ for all the 15 stations across the 6 months and plot them against the retrieved values of *P* obtained from spectral optimization as shown in Figure 5 above. Based on the scatter plot, a linear equation of y = 0.4135x and a value of R^2 = 0.9229. This suggest that the values of retrieved *P* are underestimated as they lie below the y = x line. This is also supported by the negative MBE value (-0.0698) as shown in Table 2.

d. Total absorption $a(440nm)$ excluding pure water

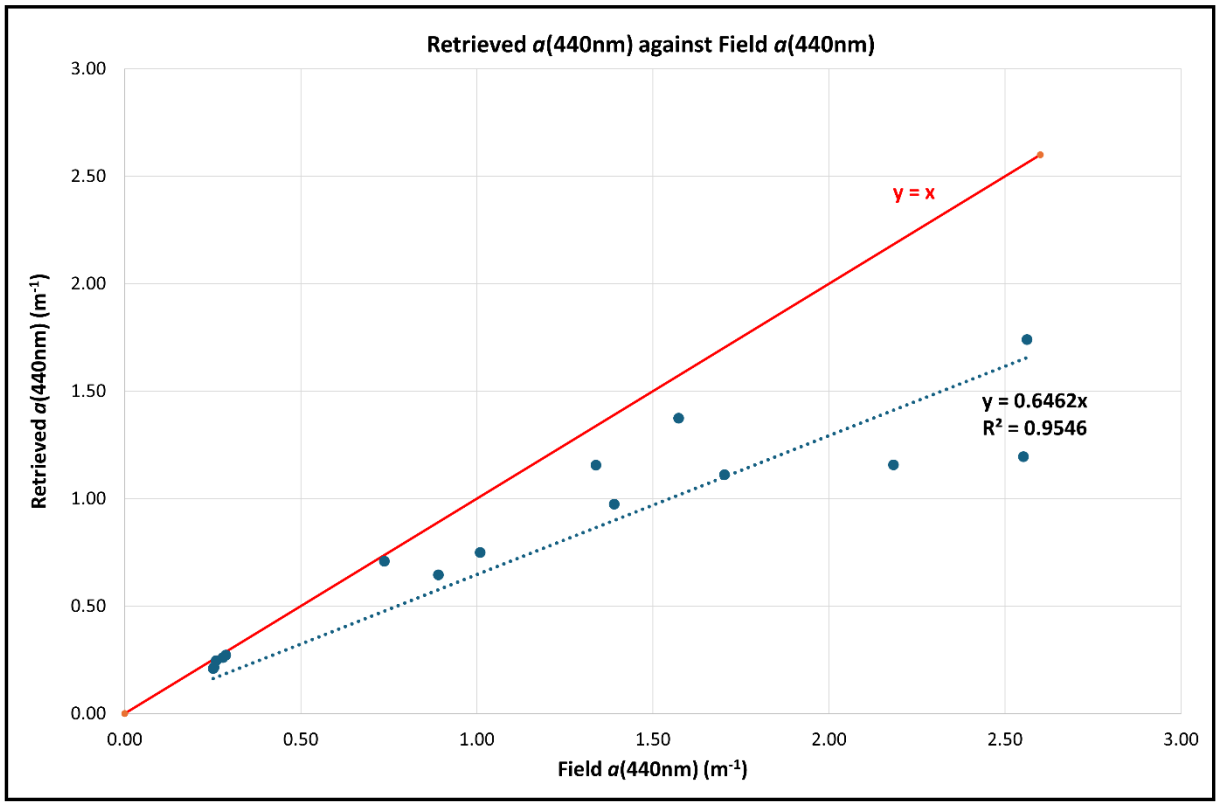

Figure 6: Scatter plot of Retrieved a(440nm) against Field a(440nm)

Parameter	R^2	RMSE	MAPE	MBE
$a(440nm)$	0.9546	0.5355	21.978	-0.3506

Table 3: Accuracy assessment of *a*(440nm)

The total absorption *a*(440nm) excluding pure water is also compared for the total retrieved values of *G* and *P* to the total *a*(440nm) obtained from field measurements. Based on the scatter plot of retrieved *a*(440nm) and field *a*(440nm), the linear equation obtained is y = 0.6462x with R^2 = 0.9546. Overall, the low values of RMSE (0.5355) and MAPE (21.978) suggests that the fitting from optimization can obtain accurate results. However, the value of MBE = -0.3506 also shows that the retrieved *a*(440nm) are understimated. Noticeably, the values of retrieved

a(440nm) are quite accurate for lower values but they become more underestimated at higher values. This is true for both the retrieved *G* and *P* values and they are attributed mainly to the values found in WJS which has the highest field $a_{dg}\ (\lambda)$ and field $a_{ph}\ (\lambda)$ values obtained. This means that at areas where the values of $a_{dg}\ (\lambda)$ and $a_{ph}\ (\lambda)$ are higher, the retrieved values are less accurate. This could be attributed to Lee's model which was specified for the range of $a_{ph}\ (440nm)$ from 0.01 to 1.0 m^{-1} (Lee, 1994). This means that for value beyond this values as with the case for several stations in WJS zone (SA1,SA2 and SA3), it will be an extrapolation and may not be accurate. It is also worth it note that Lee's model are not universally applicable and it is accurate only for water resembling the ones in his original study.

e. Backscattering of suspended particulate matter, $b_{b_p}(\lambda)$

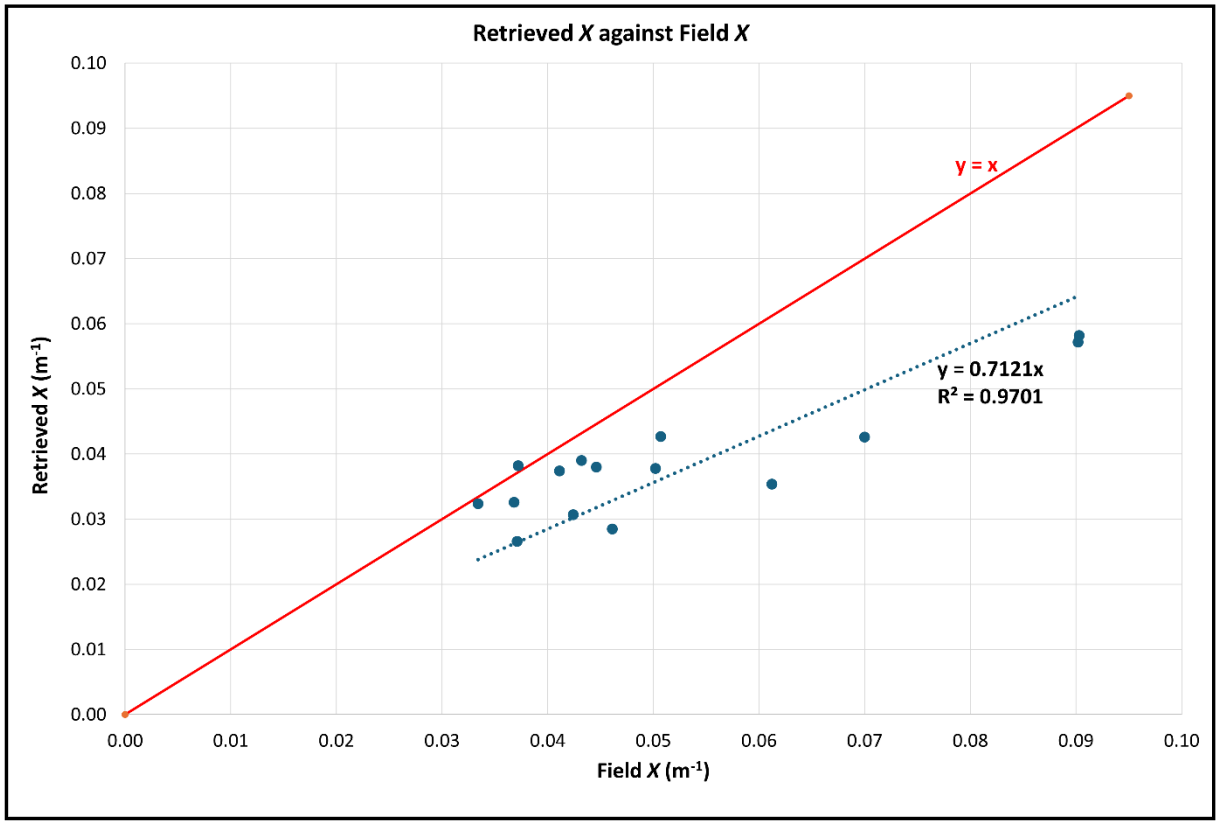

Figure 7: Scatter plot of Retrieved *X* against Field *X*

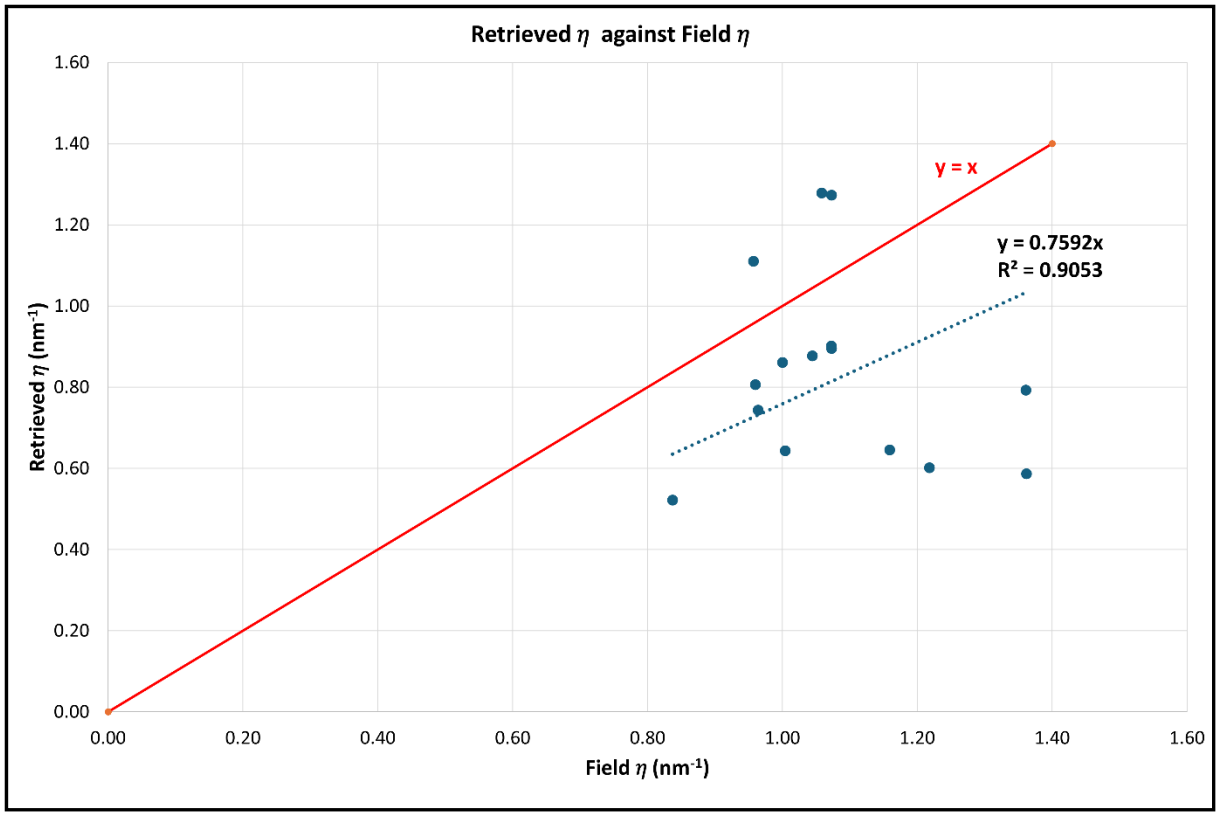

Figure 8: Linear plot of Retrieved η against Field η

Parameter	R^2	RMSE	MAPE	MBE
X	0.9701	0.0171	22.573	-0.0131
η	0.9053	0.3733	28.261	-0.2404

Table 4: Accuracy assessment of X and η

For backscattering of suspended particulate matter $b_{b_p}(\lambda)$, a scatter plot for $X = b_{b_p}(532nm)$ is obtained for the retrieved X and field X as shown in Figure 7. A linear equation is obtained as y = 0.7121x with R^2 = 0.9701. Similarly, for the scatter plot of retrieved η and field η, a linear equation of y = 0.7592x with R^2 = 0.9053 is obtained. An accuracy assessment using RMSE, MAPE, and MBE was also made to compare X and η retrieved to the field values. We can see that overall, X and η generated from spectral fitting have low RMSE and MAPE values albeit the values are underestimated (negative MNE values) when compared to field values. The RMSE values for η are also higher compared to X. The values of η are known to have a wide variation and it is often challenging to accurately estimate values of η (Yu et al., 2023). Historically, η is assumed to be close to 1 (Gordon et al., 1988) for oceanic waters and 0 for turbid coastal waters (Morel & Maritorena, 2001) while some studies reported higher values.

Conclusion

Using field measurements as inputs into Hydrolight, we simulated the synthetic remote sensing reflectance, $r_{rs}(\lambda)$ from Hydrolight. Then, we used the synthetic $r_{rs}(\lambda)$ and semi-empirical $r_{rs}(\lambda)$ to retrieved IOPs, *G, P, X* and *S, η* through spectral optimization. These parameters are then compared to the IOPs from actual field data. Overall, all of the retrieved values are close to the actual field values with $R^2 > 0.9$. Though the R^2 values are high, there are systematic biases. Most of the retrieved values are also underestimated when compared to our field results. A more robust retrieval method such as a Machine learning (ML) based retrieval algorithm could be used in future work to better capture the complex and subtle relations that affect the spectral characteristics of reflectance values.

Acknowledgments

This project is partly funded by the Office for Space Technology and Industry (OSTIn) under the Space Technology Development Programme (Grant No. S23-020016-STDP).

References

Chee Yew Leong, S., & Taguchi, S. (2005). Optical characteristics of the harmful dinoflagellate Alexandrium tamarense in response to different nitrogen sources. *Harmful Algae*, *4*(2), 211–219. https://doi.org/10.1016/j.hal.2004.02.004

Fournier, G. R., & Forand, J. L. (1994). Analytic phase function for ocean water. *Ocean Optics XII*, *2258*, 194–201.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., & Clark, D. K. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research: Atmospheres*, *93*(D9), 10909–10924.

Gregg, W. W., & Carder, K. L. (1990). A simple spectral solar irradiance model for cloudless maritime atmospheres. *Limnology and Oceanography*, *35*(8), 1657–1675.

Jerlov, N. G. (1968). *Optical Oceanography Elsevier*. Amsterdam.

Kirk, J. T. O. (1994). *Light and photosynthesis in aquatic ecosystems*. Cambridge university press.

Kraft, D. (1988). A software package for sequential quadratic programming. *Forschungsbericht- Deutsche Forschungs- Und Versuchsanstalt Fur Luft- Und Raumfahrt*.

Lee, Z. (1994). *Visible-infrared remote sensing model and applications for ocean waters*. University of South Florida.

Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G., & Patch, J. S. (1999). Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Applied Optics*, *38*(18), 3831–3843. https://doi.org/10.1364/AO.38.003831

Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied optics*, *41*(27), 5755-5772.

Mobley, C. D. (2001). Radiative transfer in the ocean. *Encyclopedia of Ocean Sciences*, *4*, 2321–2330.

Mobley, C. D., & Hedley, J. D. (2021). *Hydrolight 6.0 Ecolight 6.0 Technical Documentation: Numerical Optics Ltd*.

Moore, C., Barnard, A., & Hankins, D. (2005). Scattering Meter (BB9) User's Guide, Revision A. *WET Labs Inc., America*, *213*.

Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research: Oceans*, *106*(C4), 7163–7180.

Pope, R. M., & Fry, E. S. (1997). Absorption spectrum (380–700 nm) of pure water II Integrating cavity measurements. *Applied Optics*, *36*(33), 8710. https://doi.org/10.1364/AO.36.008710

Ritchie, R. J. (2008). Universal chlorophyll equations for estimating chlorophylls a, b, c, and d and total chlorophylls in natural assemblages of photosynthetic organisms using acetone, methanol, or ethanol solvents. *Photosynthetica*, *46*(1), 115–126. https://doi.org/10.1007/s11099-008-0019-7

Yu, X., Lee, Z., & Lai, W. (2023). Global distribution of the spectral power coefficient of particulate backscattering coefficient obtained by a neural network scheme. *Remote Sensing of Environment*, *296*, 113750. https://doi.org/https://doi.org/10.1016/j.rse.2023.113750