Introductory guidance to bio-optical algorithms for use in coastal and inland waters

Satellite remote sensing is an increasingly-used and cost-efficient complementary approach to traditional environmental monitoring and helps to address environmental issues such as eutrophication and climate change (El Serafy et al., 2021). In order to estimate ecological response variables (Adrian et al., 2009; Dörnhöfer and Oppelt, 2016) linked to the biogeochemical state of aquatic ecosystems from remote-sensed images, we consider two types of algorithms: 1/ Atmospheric correction algorithms to remove the effect of the atmosphere from the top-of-the-atmosphere (TOA) signal recorded by the remote sensor, and 2/ Bio-optical algorithms to retrieve information about the aquatic ecosystem from such data. This section focuses on bio-optical algorithms and is directed towards an audience new to the use of satellite data products and product development. More in-depth guidance on algorithm selection, use of these products, and product development, intended for a more experienced audience, can be found in our companion document.

The development of new satellite data products commonly concentrates on the biophysical, biological, and biogeochemical state of aquatic ecosystems as remote sensing can provide information on the water level, surface water temperature, ice cover, water colour, and morphometry of coastal and inland waters, while more complex models are needed to derive information on additional desirable variables (Tab. 1). Bio-optical algorithms are used to translate the optical signal recorded by the remote sensor to the bio-optical and biogeochemical variables of interest.

The optical signal recorded by the remote sensor is the result of complex interactions in the atmosphere, at the water surface, and in the water column (Fig. 1). Bio-optical algorithms explore these interactions and use measurable changes in the colour of the water caused by variations in the concentrations of key water quality constituents for their retrievals (IOCCG, 2018). Some of these algorithms might translate the optical signal recorded by the remote sensor without the removal of the contributions of the atmosphere and air-water interface, while most of them presume these effects were at least partially accounted for through the process of atmospheric correction. Partial atmospheric correction approaches are exemplified in some commonly used algorithms for the detection of cyanobacterial blooms (Binding et al., 2021; Clark et al., 2017; Lunetta et al., 2015; Stumpf et al., 2016). Details on atmospheric correction can be found in a future document.

Table 1. Remote sensing of aquatic ecosystem properties (aquatic ecosystem properties and response variables modified from Adrian et al., 2009 to add sediment loads, remotely sensed response variables modified from Dörnhöfer and Oppelt, 2016). **Bold printed properties and constituent concentrations determine the underwater light field.** DOC - Dissolved organic carbon, K_d - Diffuse attenuation coefficient for downwelled light (light attenuation), TSM - Total suspended matter concentration, a_{cdom} - Coloured dissolved organic matter absorption coefficient, S_{cdom} - Slope of the CDOM absorption coefficient, Chl*a* - Chlorophyll-*a* concentration (most abundant algal pigment), PC and PE - Phycocyanin and phycoerythrin concentrations (cyanobacterial pigments indicative of harmful algal blooms)

Aquatic ecosystem properties	Response variables	Remotely sensed response variables	Example satellite data products
Hydrology	Water level	Water level	Lake water level (ESA Lakes CCI project ECV product, CGLS water product)
Temperature	Epilimnetic temperature	Surface water temperature	Lake surface water temperature (ESA Lakes CCI project ECV product, CGLS water product)
Ice phenology	Ice-out, ice duration	Ice-out, ice-on and ice-off, ice cover	Lake ice cover (ESA Lakes CCI project)
Transparency and sediment loads	DOC, K_d, Secchi depth phenology, euphotic depth, TSM, turbidity	$\boldsymbol{a_{cdom}}$, S_{cdom}, K_d, Secchi depth, euphotic depth, **TSM**, turbidity	Turbidity derived from TSM (CGLS water product)
Community structure	Algal blooms, changes in relative species composition, primary productivity, invaders	**Chl*a*, PC**, **PE**, **algal biomass**, trophic state index, primary productivity, presence of aquatic invasive species	Trophic state index derived from Chl*a* (CGLS water product)
Habitat structure	Habitat refuge	Lake morphometry, bathymetry, Secchi depth, emerged, submerged, and floating aquatic vegetation, substrate	Lake water extent (ESA Lakes CCI project ECV product), water bodies (CGLS water product)

ESA Lakes CCI project ECV product: European Space Agency Lakes Climate Change Initiative project Essential Climate Variable product (https://climate.esa.int/en/projects/lakes/), CGLS water product: Copernicus Global Land Services water product (https://land.copernicus.eu/global/themes/water)

Constituent retrieval approaches are typically divided into data-based empirical and semi-empirical (Tab. 2) or physics-based semi-analytical and analytical methods (IOCCG, 2018). Empirical algorithms may have a bio-optical basis, such as the blue-to-green ratio adopted for open ocean Chla retrievals (most algae absorb blue light more strongly than green), or may use implicit approaches based on machine learning techniques such as neural network approaches, support vector machines and hybrid active learning models (IOCCG, 2018). Machine learning models trained by a radiative transfer model might represent intermediate types of models as they provide physics-based output while the ones trained with field measurements would represent truly empirical models. Semi-analytical algorithms always have a bio-optical basis as they are built on radiative transfer and bio-optical models. Optical water type classifications (Moore et al., 2014, Spyrakos et al., 2017) can support algorithm choice in water quality studies as some algorithms are more adapted to certain optical conditions and constituent concentration ranges (Neil et al., 2019).

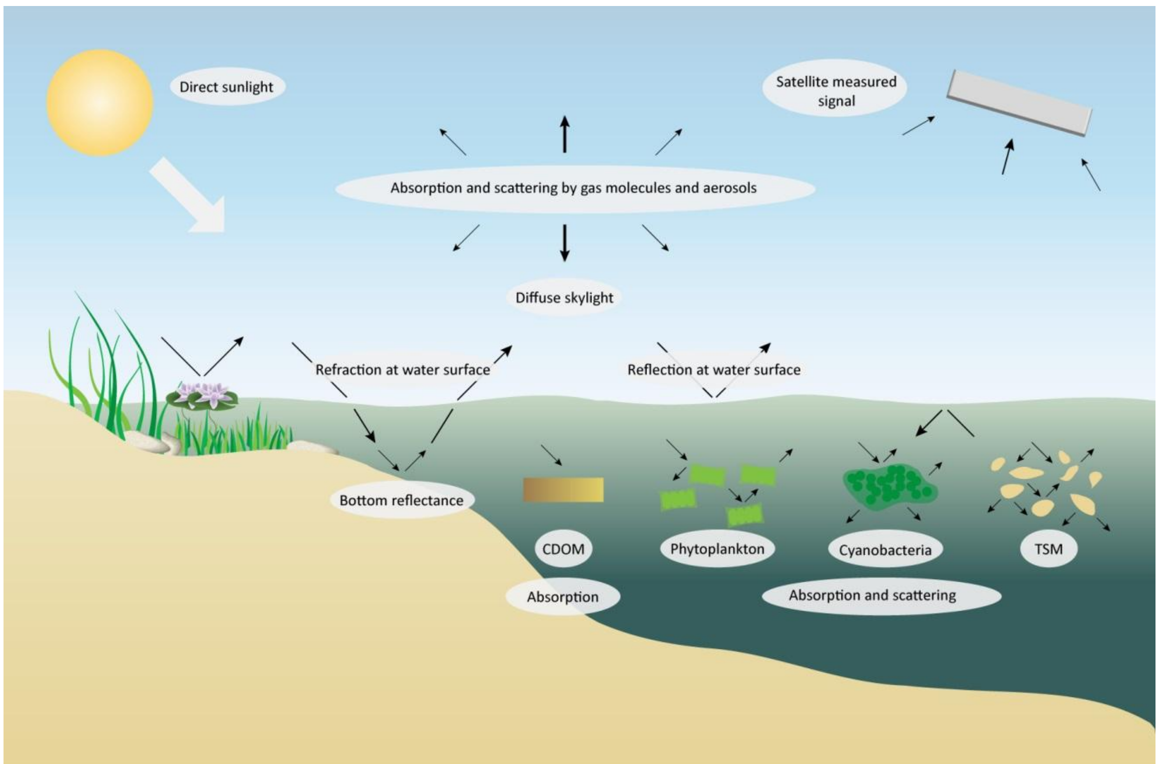

Figure 1. Conceptual figure of the remote sensing of aquatic ecosystems. The optical signal recorded by the remote sensor is the result of complex interactions in the atmosphere, at the water surface, and in the water column. Bio-optical algorithms aim to characterize key water quality constituents through their absorption and scattering properties.

Constituent retrieval strategies can range from the implementation of independent algorithms to series of algorithms with substantial computational effort. The computational effort associated with a presumably simple index for trophic state assessments is exemplified in the CGLS trophic state index (TSI) data products (https://land.copernicus.eu/global/products/lwq). These data products are the result of a complex image processing chain with an atmospheric correction of the pre-processed satellite data, optical water type classification to select the most appropriate algorithms for constituent retrievals, retrieval of Chla with the algorithm selected dependent on the water type class, and the association of Chla with the TSI in an additional calculation step. The choice of appropriate algorithms for environmental monitoring applications requires the consideration of their needs for field measurements for calibration and validation efforts, as well as their accuracy, reliability, maturity, and complexity.

Table 2. Advantages and disadvantages of different families of algorithms (description, advantages, and disadvantages modified from CEOS, 2018). SIOPs - Specific inherent optical properties (concentration dependent properties of key water quality constituents).

Family of algorithms	Description	Advantages	Disadvantages
Empirical and semi-empirical	**Use statistical relationships** of the variables of interest to the optical signal recorded by the remote sensor, **empirical algorithms can take many inversion approaches** from univariate or multivariate linear regressions to spectral decomposition methods to estimate these variables, **semi-empirical algorithms use knowledge of the underlying physics** to select the most appropriate single bands and/or band combinations	**Method easily interpretable** without the need to understand the underlying physical relationships, **computationally less expensive** than other methods, **semi-empirical algorithms can partly annul** some of the contributions of the atmosphere and air-water interface	**Coincident field measurements needed** to calibrate and validate the algorithms either for global-scale applications or for specific locations and timeframes, **struggle when measurements lie outside the range** upon which the pertinent statistical relationship was built, **difficult to adapt** to new locations and sensors, **less reliability in retrospective data analysis** compared to other methods, **untraceable uncertainties**
Machine learning (computationally more complex subset of empirical methods with physics-based and/or empirical elements)	**Use the same mechanism** as empirical algorithms, **can use a radiative transfer model**, bio-optical model, semi-analytical algorithm, or field measurements to build statistical relationships of the variables of interest to the optical signal recorded by the remote sensor	**Implementable without the need for a priori assumptions** if trained with field measurements, **very computationally efficient** in the execution of data processing tasks, **improved accuracy** dependent on the range and distribution of provided training data and accounted environmental conditions	**Coincident field measurements potentially needed** to calibrate and validate the algorithms, **less easily interpretable**, **struggle when measurements lie outside the range** of provided training data, **no mechanistic transparency**
Semi-analytical and analytical	**Use knowledge of the underlying physics** of light transfer in waters and analytical inversions to simultaneously estimate the variables of interest	**Easy to adapt** to new locations and sensors, **more reliability in retrospective data analysis and under varying recording conditions** compared to other methods, **improved accuracy** for the range of environmental impacts accounted**, traceable uncertainties, mechanistic transparency**	**Representative SIOPs and their ranges needed** to develop and run most algorithms, **less easily interpretable**, **computationally more expensive** than other methods

Authorship information

This user's guide was produced on behalf of the AquaWatch Algorithm Focus Group. We acknowledge Daniela Gurlin (Wisconsin DNR), Anders Knudby (University of Ottawa), Cédric Jamet (Université du Littoral Côte d'Opale), Arnold G. Dekker (SatDek), Maria Tzortziou (City University of New York), and Daniel Odermatt (Swiss Federal Institute of Aquatic Science and Technology) for their contributions. We encourage interested or critical members of the public to provide feedback via email with the subject line: **Introductory Algorithm Feedback** to: Mbneely@geoaquawatch.org.

Date

2021-09-27

References

Adrian, R., O'Reilly, C.M., Zagarese, H., Baines, S.B., Hessen, D.O., Keller, W., Living-stone, D.M., Sommaruga, R., Straile, D., van Donk, E., Weyhenmeyer, G.A., and Winder, M., 2009. Lakes as sentinels of climate change. Limnol. Oceanogr. 54(6.2), 2283-2297.

Binding, C.E., Pizzolato, L., and Zeng, C., 2021. EOLakeWatch; delivering a comprehensive suite of remote sensing algal bloom indices for enhanced monitoring of Canadian eutrophic lakes. Ecol. Indic. 121, 106999.

CEOS, 2018. Feasibility study for an aquatic ecosystem Earth observing system. Dekker, A.G. and Pinnel, N. (eds.), CSIRO, Canberra, Australia.

Clark, J.M., Schaeffer, B.A. Darling, J.A., Urquhart, E.A., Johnston, J.M., Ignatius, A. R., Myer, M.H., Loftin, K.A., Werdell, P.J., and Stumpf, R. P., 2017. Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. Ecol. Indic. 80, 84-95.

Dörnhöfer, K. and Oppelt, N., 2016. Remote sensing for lake research and monitoring - Recent advances. Ecol. Ind. 64, 105-122.

El Serafy, G.Y.H., Schaeffer, B.A., Neely, M.-B., Spinosa, A., Odermatt, D., Weathers, K.C., Baracchini, T., Bouffard, D., Carvalho, L., Conmy, R.N., Keukelaere, L.D., Hunter, P.D., Jamet, C., Joehnk, K.D., Johnston, J.M., Knudby, A., Minaudo, C., Pahlevan, N., Reusen, I., Rose, K.C., Schalles, J., and Tzortziou, M., 2021. Integrating Inland and Coastal Water Quality Data for Actionable Knowledge. Remote Sens. 13, 2899.

IOCCG, 2018. Earth Observations in Support of Global Water Quality Monitoring. Greb, S., Dekker, A. and Binding, C. (eds.), IOCCG Report Series, No. 17, International Ocean Colour Coordinating Group, Dartmouth, Canada.

Lunetta, R.S., Schaeffer, B.A., Stumpf, R.P., Keith, D., Jacobs, S.A., Murphy, M.S., 2015. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. Remote Sens. Environ. 157, 24-34.

Moore, T. S., Dowell, M.D, Bradt, S., and Ruiz Verdu, A., 2014. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. Remote Sens. Environ. 143, 97-111.

Neil, C., Spyrakos, E., Hunter, P.D., and Tyler, A.N., 2019. A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. Remote Sens. Environ. 229, 159-178.

Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.H., Neil, C., Barbosa, C.C.F., Binding, C., Bradt, S., Bresciani, M., Dall'Olmo, G., Giardino, C., Gitelson, A.A., Kutser, T., Li, L., Matsushita, B., Martinez-Vicente, V., Matthews, M.W., Ogashawara, I., Ruiz-Verdu, A., Schalles, J., Tebbs, E., Zhang, Y., and Tyler, A.N., 2017. Optical types of inland and coastal waters. Limnol. Oceanogr. 63, 846-870.

Stumpf, R.P., Davis, T.W., Wynne, T.T., Graham, J.L., Loftin, K.A., Johengen, T.H., Gossiaux, D., Palladino, D., Burtner, A., 2016. Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. Harmful Algae 54, 160-173.