Statistical (empirical) ocean color algorithms

Jeremy Werdell
NASA Goddard Space Flight Center

Acknowledgements: Curtis Mobley

2023 Ocean Optics Summer Course
Satellite instruments measure the spectral radiant flux leaving the top of Earth’s atmosphere

This includes contributions from everything directly in or directed into its instantaneous field of view

Ideally, forward models representing the combined ocean-atmosphere system (COAS) could be repeatedly run to find the combination that best reproduces the measured top-of-atmosphere radiance

Historically, this has been (and in many ways still is) too computationally expensive

Blum et al. 2012, Advances in Space Research
IOP(\(\lambda\))[chl, whatever] \(\rightarrow\) **forward model** \(\rightarrow\) Rrs(\(\lambda\))

Rrs(\(\lambda\)) \(\rightarrow\) **inverse model** \(\rightarrow\) IOP(\(\lambda\)), chl, whatever
empirical (adjective): based on, concerned with, or verifiable by observation or experience rather than theory or pure logic

These models are essentially just correlational models obtained from inspection of data sets containing both the inputs inputs ($R_{rs}$) and outputs ($Chl, a_{CDOM},$ water depth, etc). The models are not necessarily based on any physical insight as to why the correlation exists.
Spectra of Backscattered Light from the Sea Obtained from Aircraft as a Measure of Chlorophyll Concentration

George L. Clarke
Gifford C. Ewing
Carl J. Lorenzen

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 20543

“ocean color”

Fig. 3. Data from the high and low chlorophyll curves plotted as percentage of the incident light and compared with data taken on the same day from an area with very low chlorophyll concentration south of the Gulf Stream.
Clear water radiances for atmospheric correction of coastal zone color scanner imagery

Howard R. Gordon and Dennis K. Clark

15 December 1981 / Vol. 20, No. 24 / APPLIED OPTICS 4175

Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates

Howard R. Gordon, Dennis K. Clark, James W. Brown, Otis B. Brown, Robert H. Evans, and William W. Broenkow

20 APPLIED OPTICS / Vol. 22, No. 1 / 1 January 1983
Coastal Zone Color Scanner
1978-1986

<table>
<thead>
<tr>
<th>metric</th>
<th>CZCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>primary ocean bands (nm)</td>
<td>443, 520, 550, 670</td>
</tr>
<tr>
<td></td>
<td>○ chl-a + phaeopigments</td>
</tr>
<tr>
<td></td>
<td>○ diffuse attenuation at 490 nm</td>
</tr>
<tr>
<td>nadir res.</td>
<td>825 m</td>
</tr>
<tr>
<td>nadir swath</td>
<td>1636 km</td>
</tr>
<tr>
<td></td>
<td>discontinuous operation</td>
</tr>
<tr>
<td>tilt</td>
<td>none</td>
</tr>
<tr>
<td>det. per band</td>
<td>1</td>
</tr>
<tr>
<td>digitization</td>
<td>8 bits</td>
</tr>
</tbody>
</table>
The seminal idea of ocean color remote sensing: Chl concentration and water-leaving radiance are correlated.

Fig. 10.1. Water-leaving radiances $L_w$ as a function of wavelength for four chlorophyll concentrations $C$, in case 1 waters. The shaded regions labeled 1-4 indicate the detector bandwidths of the CZCS sensor. [redrawn from Gordon, et al., (1985), by permission]
revolutionized perceptions of the ocean

first ever look at global distribution of marine phytoplankton, ocean productivity

Feldman et al. 1984 and 1989
since the CZCS-era, 3 general flavors* of empirical Chl algorithms have emerged:

1. Rrs band ratios
2. Rrs line heights
3. neural networks
4. blended band ratios & line heights

* this, of course, is a grand oversimplification of progress and algorithm development, as many other approaches exist and many have been trained to retrieve IOPs, Kd, water depth, and other aquatic parameters
Rrs maximum band ratio

\[ X = \log_{10} \left[ \frac{\text{Rrs}(443 > 490 > 510)}{\text{Rrs}(555)} \right] \]

\[ \log_{10}(\text{chl}) = a_0 + a_1 X + a_2 X^2 + a_3 X^3 + a_4 X^4 \]


developed using a “global” dataset of in situ Rrs(\(\lambda\)) and chl
Rrs line height (baseline subtraction)


\[ CI = R_{rs,555} - \left[ R_{rs,443} + \frac{(555-443)}{(670-443)} \times \left( R_{rs,670} - R_{rs,443} \right) \right] \]
blended band ratio & line height


\[
\text{Chl}_{\text{OCI}} = \begin{cases} 
\text{Chl}_{\text{C1}} & \text{for } \text{Chl}_{\text{C1}} \leq 0.25 \text{ mg m}^{-3} \\
\alpha \times \text{Chl}_{\text{OC4}} + \beta \times \text{Chl}_{\text{C1}} & \text{for } 0.25 < \text{Chl}_{\text{C1}} \leq 0.3 \text{ mg m}^{-3},
\end{cases}
\]

where \(\alpha = (\text{Chl}_{\text{C1}} - 0.25)/(0.3 - 0.25)\) and \(\beta = (0.3 - \text{Chl}_{\text{C1}})/(0.3 - 0.25)\).
cannot emphasize enough how many empirical (statistical) approaches exist and for how many varied aquatic retrievals (IOPs, Kd, water depth, others …)

some are even nested (e.g., $K_d, bb = f(Chl = f(Rrs))$)
SeaBASS holdings by year: 2006-2009
Advantages:
• fast (computationally inexpensive)
• easily portable & reconfigurable
• often immune to some atmospheric correction issues
• sometimes insensitive to Rrs uncertainties
• easily expanded to consider other info (PCs, SST; e.g., Lange et al. 2020, Optics Express)
• often simply good enough

Disadvantages:
• non-unique (pervasive in inverse modeling)
• geographically biased (when using global products)
• Chl approaches considered “too Case-1”
• sometimes highly sensitive to Rrs uncertainties
• can easily prohibit insightful, conscientious interpretation of results
• not predictive

Dirty secret:
• embedded into nearly 100% of all ocean color

Your opinions and thoughts?
Band-ratio algorithms can be less sensitive to bad atmospheric correction than some other techniques such as spectrum matching.
While errors in the same direction for both Rrs(blue) and Rrs(green) can cancel out … … opposite direction errors in Rrs(blue) and Rrs(green) amplify errors in many retrievals
Band-ratio algorithms are vulnerable to non-uniqueness problems because the $R_{rs}$ ratioing throws out magnitude information that makes spectra unique. Every unique spectrum below has $R_{rs}(490)/R_{rs}(555) = 1.71 \pm 0.01$, which gives $Chl = 0.59$ mg/m$^3$ by the SeaWiFS OC2 algorithm; all of these spectra had $Chl < 0.2$ mg/m$^3$ (these spectra are influenced by bottom reflectance).
analytical methods to calculate uncertainties exist …

… but they don’t encompass all sources (yet), such as those assigned to the in situ measurements and fit parameters
Model Selection

In some situations, you can figure out (from intuition, theoretical guidance, or data analysis) the general mathematical form of the model that links the input and output (e.g., the polynomial functions that relate the band ratios to Chl). You can then use the available data (e.g., simultaneous measurements of $R_{rs}(\lambda)$ and Chl) to get best-fit coefficients in the model via least-squares fitting.

But what if you don’t have any idea what the mathematical form of the model is?