VALIDATING OCEAN COLOR PRODUCTS

Thanks to Chris Proctor and SeaBASS team
What is SeaBASS?

- Data Archive
- Data Format (.sb)
- Services & Tools
SeaBASS Core Functions

Public-Facing Deliverables

- Bio-Optical Archive
- Validation Match-up Archive
- NOMAD Dataset

1. Data submitted

2a. Format verification & archive
   2b. Data QA/QC; assess for validation

3. Validation match-ups (SIPS interface)

4. NOMAD data curation
Ocean Optics & Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation


Ocean Optics & Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation

IOCCG Protocol Series Volume 7.0, 2021

Aquatic Primary Productivity Field Protocols for Satellite Validation and Model Synthesis (DRAFT)

Report of a NASA-sponsored workshop with contributions (alphabetical) from:
Ocean Optics & Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation


Authors:
Joaquín E. Chávez, Ivona Cerini, Giorgio Dall’Olmo, Meg Estapa, Wilford Gardner, Miguel Góth, Jason R. Graff, Peter Herms, Phoebe J. Lum, Zhanfen Lu, Michael W. Lomas, Antonio Mannino, Michael G. Novak, Robert Turnerwitsch, P., Jeremy Werdell, Toby K. Westberry
Level-2 match-ups – extra notes
General processing flow and exclusion criteria

In situ data
Select L1A files to be processed to L2 (Extract 101x101 pixels)
Process to L2 using MSI12 (L2gen)
Find closest pixel to station and extract 5x5 pixel box centered on it
Exclude pixels flagged by any of the following L2 flags: LAND, HILINT, HILT, STRAYLIGHT, CLDICE, ATMFAIL, LOWLW, FILTER, NAVFAIL, or NAVWARN; Count Number Good Pixels (NGP)

Is time difference < 180 minutes?

YES

Is NGP > NTP/2 + 1?
NTP = Number Total Pixels, excluding land pixels

NO

Fails validation criteria

Select filtered mean per product (unflagged pixels within +/- 1.5° of mean)

Is median of CV calculated just for [Rrs 405-570 nm & AOT ~0.6 nm] (excluding any land bands) < 0.15?

NO

Sensor: Zenith < 60° and Solar: Zenith < 75°?

NO

Evaluate Sensor Zenith range. If < 10°, pick smallest time difference. Else, pick smallest Sensor Zenith.
Also, if applicable, eliminate any redundant L2 files with non-optimal resolution (e.g., SeaWiFS MLAC > GAC)

YES

If multiple L2 files for the same in situ data:
Eliminate stations from same L2 file that are too close together (overlapping of 5x5 box)

Passes validation criteria

SeaBASS web-based Level-2 Validation Search (https://seabass.gsfc.nasa.gov/search#val)

<table>
<thead>
<tr>
<th>Product Name</th>
<th>#</th>
<th>Mean Bias</th>
<th>Mean Absolute Error (MAE)</th>
<th>MODIS-Aqua Range</th>
<th>In situ Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>poc</td>
<td>210</td>
<td>0.86273*</td>
<td>1.42003*</td>
<td>21.44696 - 1013.09992</td>
<td>21.30000 - 1168.77698</td>
</tr>
</tbody>
</table>

* statistical calculations based on log_{10} (implies ignoring values equal to or less than zero)

Seegers et al, 2018
Upgrade to validation (graphics and approach)!

- McKinna et al. 2021
- Similar metrics to what current SeaBASS is using with addition of uncertainties
- Uncertainties are present as a correction factor (~degree of overlap)
- Keep the x/y linear plots but with Bland–Altman-type scatter plots
- Zeta scores with color coding (indicating satisfactory (<|2|), questionable(|2-3|), and unsatisfactory (>|4|) zeta values)

<table>
<thead>
<tr>
<th>Product name</th>
<th>n</th>
<th>bias</th>
<th>Bias’</th>
<th>MAE</th>
<th>MAE’</th>
<th>Mean zeta</th>
<th>Mean zeta’</th>
<th>Range OC</th>
<th>Range in situ</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH</td>
<td>326</td>
<td>1.21</td>
<td>1.12</td>
<td>1.33</td>
<td>1.16</td>
<td>0.888</td>
<td>0.561</td>
<td>???</td>
<td>???</td>
</tr>
</tbody>
</table>
References


• HyperInSPACE github - [https://github.com/nasa/HyperInSPACE](https://github.com/nasa/HyperInSPACE)

• SeaBASS validation - [https://seabass.gsfc.nasa.gov/search#val](https://seabass.gsfc.nasa.gov/search#val)
Designation of data maturity level

- Following the parameter specific validation procedure, data maturity level is assigned to each of the products following NASA's Data Maturity Levels designation.

- **Stage 1 Validation**: Product accuracy is estimated using a small number of independent measurements obtained from selected locations and time periods and ground-truth/field program efforts.

- **Stage 2 Validation**: Product accuracy is estimated over a significant set of locations and time periods by comparison with reference in situ or other suitable reference data. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.

- **Stage 3 Validation**: Product accuracy has been assessed. Uncertainties in the product and its associated structure are well quantified from comparison with reference in situ or other suitable reference data. Uncertainties are characterized in a statistically robust way over multiple locations and time periods representing global conditions. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.

- **Stage 4 Validation**: Validation results for stage 3 are systematically updated when new product versions are released and as the time-series expands.

- Stolen from: [https://science.nasa.gov/earth-science/earth-science-data/data-maturity-levels](https://science.nasa.gov/earth-science/earth-science-data/data-maturity-levels)