A class-based approach for mapping the uncertainty of empirical chlorophyll algorithms

Timothy S. Moore
University of New Hampshire
NASA MODIS Meeting
January 26, 2010

...in collaboration with...

Mark Dowell, JRC
Janet Campbell, UNH
Updates since OCRT (May 2009)

- Fix to NOMAD screening (more oligotrophic points).
- Fix to membership function (increase in class memberships).
- Generalized table for SeaWiFS, MODIS, MERIS.
- Migrated to a developmental l2gen.
- Updates to empirical chl uncertainties from v6 reprocessing.
What’s the problem?

- Current single, bulk estimates of chlorophyll error (50-78%) for the empirical algorithms exceed the desired goal of 35%.

- This is *misleading*, as algorithms do not perform to the same level of accuracy in different optical environments.

- Product error is relevant to higher-order algorithms that use OC products, and understanding changes in CDRs.

- Question: How can we more accurately assess OC product ‘error’ and geographically map them?
Average absolute error: 50% based on NOMAD V2
NOMAD V2
Approach

• Previously, we have implemented a fuzzy logic methodology for distinguishing different optical water types based on remote sensing reflectance.

• The same techniques can be adapted for characterizing chlorophyll uncertainty (or more accurately called discrepancy) for empirical algorithms.

• The advantage gained is that different regions of the empirical algorithm can be 1) discretely characterized and 2) individually mapped using satellite reflectance data.
- NOMAD V2
- SeaWiFS Validation Set
- Aqua Validation Set
- Rrs
- In situ Chl
- Algorithm Chl
In-situ Database (NOMAD V2)

Rrs(\lambda)

OC3/4 Rel. Error

Cluster analysis

station data sorted by class

8 classes

class-based average relative error

Satellite Measurements

Merged Image Product

Calculate membership

Individual class error

\text{Class } M_i, \Sigma_i
NOMAD V2 Clustering Results

- Cluster analysis on SeaWiFS Rrs bands
- 8 clusters optimal based on cluster validity functions

N~2400
Class Mean Reflectance Spectra

- Rrs mean spectra behave as endmembers
- Rrs class statistics form the *fuzzy membership function*. 
Characterizing class uncertainty

NOMAD V2

N=1543

Aqua validation set

N=541

SeaWiFS validation set

N=1576

Log10(max(Rrs443,Rrs488)/Rrs551)

chlorophyll mg/m³

chlor a uncertainty

Type 1 2 3 4 5 6 7 8
## Relative Error - %

<table>
<thead>
<tr>
<th>Class</th>
<th>NOMAD (OC3)</th>
<th>OC3 (v5)</th>
<th>OC3 (v6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
<td>62</td>
<td>58</td>
</tr>
<tr>
<td>5</td>
<td>77</td>
<td>62</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>94</td>
<td>79</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>86</td>
<td>60</td>
</tr>
<tr>
<td>8</td>
<td>55</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Avg.</td>
<td>53</td>
<td>78</td>
<td>74</td>
</tr>
</tbody>
</table>
Producing the Discrepency Map

For each pixel,

\[ \sum f_i \ast \]

\[ i = 1 \ldots 8 \]

<table>
<thead>
<tr>
<th>Relative Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
</tr>
<tr>
<td>33</td>
</tr>
<tr>
<td>42</td>
</tr>
<tr>
<td>58</td>
</tr>
<tr>
<td>59</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>N/A</td>
</tr>
</tbody>
</table>
Frequency of low membership sum
Conclusions

• Single, bulk estimates of algorithm performance do not realistically describe the spatial distribution of error.

• Basing OC3/OC4 error statistics with the Aqua and SeaWiFS validation data set is recommended because it reflects product discrepancy.

• The class-based method is a way to characterize product discrepancy for different optical environments and to dynamically map them pixel by pixel.

• Class-based approach provides a common framework that can be applied to different satellites and different algorithms at multiple spatial scales.

• We envision the error maps as separate, companion products to the existing suite of NASA OC products (currently in developmental l2gen).
MERIS image - Aug. 22, 2008
What is fuzzy logic?

- Designed to handle data imprecision and ambiguity
- Allows for multiple outcomes using a fuzzy membership

Hard:
- Traditional minimum-distance criteria

Fuzzy:
- Fuzzy graded membership
  - Water = 0.05
  - Wetland = 0.65
  - Forest = 0.30
The Membership Function

\[ Z^2 = (V_{rs} - y_j)^\top \Sigma_j^{-1} (V_{rs} - y_j) \]

\( V_{rs} \) – satellite pixel vector
\( y_j \) – \( j \)th class mean vector
\( \Sigma_j \) – \( j \)th class covariance matrix

Result: A number between 0 and 1 that is a measure of the vector’s membership to that class.
May 2005

SeaWiFS OC4

Aqua OC3

Relative Error (%)
SeaWiFS OC4 Error
Aqua OC3 Error

Jan 2005

Apr 2005

Jul 2005

Oct 2005

Relative Error (%)