PFT satellite algorithm intercomparison + validation plan

T. Hirata (Hokkaido Univ., JPN)
R.J.W. Brewin (Plymouth Marine Lab., GBR)
S. Lavender (Pixalytics Ltd, GBR)
T. Kostadinov (Univ. Richmond, USA)
Acknowledgements

Algorithm developers for PFT data processing

IOCCG for project endorsement

JAXA for PFT workshop support under Global Change Observation Mission (GCOM-C) (1st WS@Sapporo 2011, 2nd WS@Glasgow 2012)

NASA for SeaWiFS data (2003-2007) used in our analysis

ESA for SCHYMACHY data (2003-2007) used in our analysis

Many others for supporting PFT intercomparison activities
To obtain a community consensus as to spatial and temporal variability of PFTs as observed by satellite

→ To deliver our knowledge derived by ocean colour science to other disciplines

→ **Not to reveal which algorithm is good or bad!**

<table>
<thead>
<tr>
<th>Working Group</th>
<th>Leaders</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Guide</td>
<td>N. Hardman-Mountford, C. Mouw</td>
</tr>
<tr>
<td>In situ Data</td>
<td>L. Clementson, R. Barlow, T. Hirawake</td>
</tr>
<tr>
<td>Algorithm Intercomparison</td>
<td>T. Hirata, T. Kostadinov</td>
</tr>
<tr>
<td>Algorithm Validation</td>
<td>R.J.W. Brewin, A. Bracher</td>
</tr>
<tr>
<td>Algorithm providers</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Alvain et al. (CNRS)</td>
<td>Alvan et al., DSR, 2005, Alvain et al., GBC, 2008</td>
</tr>
<tr>
<td>Bracher et al. (U. Bremen/AWI)</td>
<td>Bracher et al., Biogeosci., 2009, Sedaghi et al., Biogeosci., 2012</td>
</tr>
<tr>
<td>Brewin et al. (PML)</td>
<td>Brewin et al., Ecological Modelling</td>
</tr>
<tr>
<td>Bricaud et al. (LOV)</td>
<td>Ciotti &amp; Bricaud, L&amp;O-Meth., 2006, Bricaud et al., GBC, 2012</td>
</tr>
<tr>
<td>Hirata et al. (Hokkaido U)</td>
<td>Hirata et al., Biogeosci., 2011</td>
</tr>
<tr>
<td>Hirawake et al. (Hokkaido U)</td>
<td>Fujiwara et al., Biogeosci., 2011</td>
</tr>
<tr>
<td>Kostadinov et al. (U. of Richmond)</td>
<td>Kostadinov et al., Biogeosci., 2010</td>
</tr>
<tr>
<td>Uitz et al. (LOV)</td>
<td>Uitz et al., JGR-Oceans, 2006</td>
</tr>
</tbody>
</table>

NOTE: MLD required

*Most algorithms analyze Microplankton and Picoplankton (most common among these algors are analyzed)*
Preliminary Results of Spatial Comparison
Opitcs Abundance

Mean derived from monthly fields Micro [mg/m³] from Bracher et al. is calculated using SCHIMACHY Diatom and SeaWiFS Chla.

Micro [%] x SeaWiFS Chla Algorithms using SeaWiFS showed a similar result. 2003 - 2007 Mean Field (Micro [mg/m³])

Bracher et al. (SCHIMACHY Diatoms) Bricaud et al.

Hirawake et al. Roy et al.

Brewin et al. Uitz et al.

Hirata et al. Kostadinov et al.

Ensemble Mean [mg/m³]

Ensemble mean (micro) [mgChla/m^3]
Tends to agree among algorithms, except in higher latitudes
### Microplankton [mg/m³] Basin Average (2003-2007)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>0.319</td>
<td>0.759</td>
<td>0.866</td>
<td>0.359</td>
<td>0.495</td>
<td>0.892</td>
<td>0.399</td>
<td>0.400</td>
<td>0.561</td>
</tr>
<tr>
<td>SOC</td>
<td>0.356</td>
<td>0.243</td>
<td>0.294</td>
<td>0.089</td>
<td>0.117</td>
<td>0.231</td>
<td>0.152</td>
<td>0.231</td>
<td>0.214</td>
</tr>
<tr>
<td>NAT</td>
<td>0.076</td>
<td>0.247</td>
<td>0.252</td>
<td>0.138</td>
<td>0.155</td>
<td>0.229</td>
<td>0.116</td>
<td>0.092</td>
<td>0.163</td>
</tr>
<tr>
<td>SAT</td>
<td>0.143</td>
<td>0.071</td>
<td>0.069</td>
<td>0.051</td>
<td>0.059</td>
<td>0.061</td>
<td>0.039</td>
<td>0.023</td>
<td>0.064</td>
</tr>
<tr>
<td>NPC</td>
<td>0.066</td>
<td>0.190</td>
<td>0.198</td>
<td>0.115</td>
<td>0.131</td>
<td>0.191</td>
<td>0.081</td>
<td>0.081</td>
<td>0.132</td>
</tr>
<tr>
<td>SPC</td>
<td>0.072</td>
<td>0.060</td>
<td>0.064</td>
<td>0.047</td>
<td>0.054</td>
<td>0.069</td>
<td>0.052</td>
<td>0.044</td>
<td>0.058</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Optics</th>
<th>Abundance</th>
<th>Relative diff.[%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>0.639</td>
<td>0.432</td>
</tr>
<tr>
<td>SOC</td>
<td>0.243</td>
<td>0.167</td>
</tr>
<tr>
<td>NAT</td>
<td>0.188</td>
<td>0.121</td>
</tr>
<tr>
<td>SAT</td>
<td>0.079</td>
<td>0.041</td>
</tr>
<tr>
<td>NPC</td>
<td>0.152</td>
<td>0.098</td>
</tr>
<tr>
<td>SPC</td>
<td>0.062</td>
<td>0.050</td>
</tr>
</tbody>
</table>

1. **Optical > Abundance in general**
2. **Difference due to approach is evident**
3. **ARC > SOC > NAT > NPC > SAT > SPC**
Opitcs Abundance 2003-2007 Mean Field (Micro [%])

From PSD Diatom (Alvain et al.) [Probability of “dominance”]

• [%] from Kostadinov et al. is obtained from particle size distribution

Higher latitudes show more discrepancies

Subtropical waters: almost no diatom? Or still some diatoms?

PSD [%] Bricaud et al. Bricaud et al.

Hirawake et al. Roy et al.

Brewin et al. Uitz et al.

Hirata et al. Kostadinov et al.
2003-2007 Bias Field (**Micro [%]-EnMean**)  

More difference is found in [%] than in [mg/m³]

- Bricaud et al.
- Hirawake et al.
- Roy et al.
- Brewin et al.
- Uitz et al.
- Hirata et al.
- Kostadinov et al.
**Microplankton [%] Basin Average (2003-2007)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>N/A</td>
<td>62.67</td>
<td>39.43</td>
<td>3.22</td>
<td>40.73</td>
<td>36.61</td>
<td>31.63</td>
<td>27.09</td>
<td>34.4</td>
</tr>
<tr>
<td>SOC</td>
<td>N/A</td>
<td>27.05</td>
<td>29.46</td>
<td>5.45</td>
<td>20.32</td>
<td>22.80</td>
<td>24.81</td>
<td>16.77</td>
<td>20.95</td>
</tr>
<tr>
<td>NAT</td>
<td>N/A</td>
<td>21.82</td>
<td>23.93</td>
<td>16.89</td>
<td>19.24</td>
<td>21.04</td>
<td>18.59</td>
<td>9.57</td>
<td>18.72</td>
</tr>
<tr>
<td>SAT</td>
<td>N/A</td>
<td>13.91</td>
<td>17.61</td>
<td>17.16</td>
<td>17.60</td>
<td>17.46</td>
<td>17.46</td>
<td>8.17</td>
<td>15.62</td>
</tr>
<tr>
<td>NPC</td>
<td>N/A</td>
<td>19.93</td>
<td>23.49</td>
<td>27.73</td>
<td>21.29</td>
<td>20.59</td>
<td>18.90</td>
<td>10.78</td>
<td>20.38</td>
</tr>
<tr>
<td>SPC</td>
<td>N/A</td>
<td>11.55</td>
<td>16.40</td>
<td>19.02</td>
<td>15.33</td>
<td>16.79</td>
<td>17.23</td>
<td>8.16</td>
<td>14.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Optics</th>
<th>Abundance</th>
<th>Relative diff [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>36.51</td>
<td>31.77</td>
<td>6.93</td>
</tr>
<tr>
<td>SOC</td>
<td>20.57</td>
<td>21.46</td>
<td>-2.11</td>
</tr>
<tr>
<td>NAT</td>
<td>20.47</td>
<td>16.40</td>
<td>11.03</td>
</tr>
<tr>
<td>SAT</td>
<td>16.57</td>
<td>14.36</td>
<td>7.13</td>
</tr>
<tr>
<td>NPC</td>
<td>23.11</td>
<td>16.75</td>
<td>15.93</td>
</tr>
<tr>
<td>SPC</td>
<td>15.57</td>
<td>14.06</td>
<td>5.11</td>
</tr>
</tbody>
</table>

1. Optical > Abundance in general but values are more similar than the comparison in [mg/m³]
2. Difference due to approach is more remarkable in NAT & NPC
3. ARC > SOC > **NPC** > **NAT** > SAT > SPC
Opitcs

Abundance

Bracher et al. (SCHYMACHY Cyanobacteria)

• SCHYMACHY derives cyanobacteria rather than Pico

2003 - 2007 Mean Field (Pico [mg/m^3])

Bricaud et al.

Hirawake et al

Roy et al.

Brewin et al.

Uitz et al.

Hirata et al

Kostadinov et al.
**Picoplankton [mg/m^3] Basin Average (2003-2007)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>0.237</td>
<td>0.192</td>
<td>0.223</td>
<td>0.103</td>
<td>0.102</td>
<td>0.103</td>
<td>0.135</td>
<td>0.161</td>
<td>0.157</td>
</tr>
<tr>
<td>SOC</td>
<td>0.002</td>
<td>0.168</td>
<td>0.114</td>
<td>0.128</td>
<td>0.081</td>
<td>0.082</td>
<td>0.109</td>
<td>0.109</td>
<td>0.096</td>
</tr>
<tr>
<td>NAT</td>
<td>0.010</td>
<td>0.134</td>
<td>0.102</td>
<td>0.053</td>
<td>0.063</td>
<td>0.064</td>
<td>0.065</td>
<td>0.072</td>
<td>0.070</td>
</tr>
<tr>
<td>SAT</td>
<td>0.024</td>
<td>0.150</td>
<td>0.101</td>
<td>0.066</td>
<td>0.065</td>
<td>0.069</td>
<td>0.065</td>
<td>0.069</td>
<td>0.076</td>
</tr>
<tr>
<td>NPC</td>
<td>0.008</td>
<td>0.157</td>
<td>0.108</td>
<td>0.050</td>
<td>0.067</td>
<td>0.072</td>
<td>0.073</td>
<td>0.082</td>
<td>0.077</td>
</tr>
<tr>
<td>SPC</td>
<td>0.015</td>
<td>0.142</td>
<td>0.093</td>
<td>0.062</td>
<td>0.057</td>
<td>0.064</td>
<td>0.063</td>
<td>0.068</td>
<td>0.070</td>
</tr>
</tbody>
</table>

### Optics Abundance Relative diff.[%]

<table>
<thead>
<tr>
<th></th>
<th>Optics</th>
<th>Abundance</th>
<th>Relative diff.[%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>0.171</td>
<td>0.133</td>
<td>12.614</td>
</tr>
<tr>
<td>SOC</td>
<td>0.098</td>
<td>0.092</td>
<td>3.462</td>
</tr>
<tr>
<td>NAT</td>
<td>0.072</td>
<td>0.067</td>
<td>3.873</td>
</tr>
<tr>
<td>SAT</td>
<td>0.081</td>
<td>0.067</td>
<td>9.090</td>
</tr>
<tr>
<td>NPC</td>
<td>0.078</td>
<td>0.075</td>
<td>1.518</td>
</tr>
<tr>
<td>SPC</td>
<td>0.073</td>
<td>0.065</td>
<td>6.340</td>
</tr>
</tbody>
</table>

1. Optical > Abundance (especially in high latitudes = ARC) but the difference between optics- and abundance-based is less than in Micro
2. ARC > SOC > NPC > SAT ≥ NAT ≥ SPC
Cyanobacteria

Optics

Abundance

- SCHYMACHY does not give Chla, so [%] is not calculated.
- Cyanobacteria from Alvain et al.'s algor. seems to give similar distribution to others, in spite that it gives temporal frequency of dominance rather than % in biomass

Clear differences in horizontal extent of subtropical gyres

Ensemble Mean

Bricaud et al., Hirawake et al., Roy et al., Brewin et al., Uitz et al., Hirata et al., Kostadinov et al., 2003-2007 Mean Field (Pico [%])
In general, bias is larger in the subtropical gyres and/or Antarctic polar front, regardless of algorithms.
## Picoplankton [%] Basin Average (2003-2007)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>N/A</td>
<td>37.26</td>
<td>34.80</td>
<td>0.72</td>
<td>13.50</td>
<td>18.39</td>
<td>24.65</td>
<td>29.92</td>
<td>34.48</td>
</tr>
<tr>
<td>SOC</td>
<td>N/A</td>
<td>71.81</td>
<td>43.86</td>
<td>5.22</td>
<td>40.50</td>
<td>34.84</td>
<td>33.00</td>
<td>36.51</td>
<td>20.95</td>
</tr>
<tr>
<td>NAT</td>
<td>N/A</td>
<td>49.67</td>
<td>55.04</td>
<td>32.71</td>
<td>44.66</td>
<td>43.02</td>
<td>42.80</td>
<td>49.30</td>
<td>18.72</td>
</tr>
<tr>
<td>SAT</td>
<td>N/A</td>
<td>80.55</td>
<td>59.12</td>
<td>32.16</td>
<td>42.97</td>
<td>43.18</td>
<td>41.51</td>
<td>43.21</td>
<td>15.62</td>
</tr>
<tr>
<td>NPC</td>
<td>N/A</td>
<td>69.41</td>
<td>54.42</td>
<td>18.46</td>
<td>37.90</td>
<td>40.04</td>
<td>41.06</td>
<td>43.15</td>
<td>20.38</td>
</tr>
<tr>
<td>SPC</td>
<td>N/A</td>
<td>78.56</td>
<td>60.77</td>
<td>31.23</td>
<td>48.03</td>
<td>45.75</td>
<td>42.54</td>
<td>45.45</td>
<td>14.92</td>
</tr>
</tbody>
</table>

### Optics vs Abundance

1. Optical > Abundance in general

1. Difference due to approaches is more remarkable in NAT & SAT

3. ARC > SOC > NPC > NAT > SAT > SPC
Preliminary Results of Temporal Comparison
Seasonality is most dominant in PFT dynamics
Microplankton [mg/m³]: Seasonality (monthly climatology)

Peak timing generally agrees
Picoplankton [mg/m³]: Seasonality (monthly climatology)

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

ARC

SOC

NAT

SAT

NPC

SPC

GLB

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

Mean

Bracher et al.
Briculaud et al.
Fujiwara et al.
Roy et al.
Brewin et al.
Uitz et al.
Hirata et al.
Kostadinov et al.
Picoplankton [%]: Seasonality (monthly climatology)

- **ARC**
- **SOC**
- **NAT**
- **SAT**
- **NPC**
- **SPC**
- **GLB**

January to December

Data sources:
- Bracher et al.
- Bricaud et al.
- Fujiwara et al.
- Roy et al.
- Brewin et al.
- Uitz et al.
- Hirata et al.
- Kostadinov et al.
- Mean

Alvain et al. Synecho-like algor.
<table>
<thead>
<tr>
<th>Micro</th>
<th>Pico</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong>&lt;br&gt;Definition</td>
<td>&gt; 20μm</td>
<td>&lt; 2μm</td>
</tr>
<tr>
<td><strong>HPLC</strong>&lt;br&gt;Definition</td>
<td>Fucoxanthin (DPA, CHEMTAX etc)</td>
<td>Zeaxanthin (DPA, CHEMTAX etc)</td>
</tr>
<tr>
<td>a_{ph} &lt;br&gt;<strong>Definition</strong></td>
<td>a_{ph}^* for &gt; 20μm</td>
<td>a_{ph}^* vs for &lt; 2μm</td>
</tr>
<tr>
<td></td>
<td>a_{ph}^* from dominant Fuco samples</td>
<td>a_{ph}^* from dominant Zea samples</td>
</tr>
</tbody>
</table>

**Comparison between DPA & CHEMTAX using NOMAD data set**<br>(Werdell and Bailey, 2005)

**Ciotti et al., 2002**

**Bracher et al., 2009**
Algorithm Validation Plan
Algorithm Comparison: Guiding Principles

- Approach should be independent from the algorithm developers
- Need to consider what is important for the eventual end users e.g. oceanographic model assimilation
- Need to find parameters for comparison;
  - May not exist within the outputs of all the algorithms or be an indirect output (e.g. dominance frequencies)
  - It’s likely to require an element of modelling and/or making certain assumptions for the in-situ data.
- Where possible, validation in-situ data should not have been used in algorithm development or (if not possible) dependency needs to be quantified
- Should the PFT algorithms be partitioned into common categories? Or perform an overarching comparison incorporating all techniques?
Algorithm Comparison: Proposed Steps

IOCS, May 2013

PFT workshops @ Sapporo 2011, @Glasgow 2012

1) Repeat previous work as conducted by Brewin et al (2011, RSE):
   ▪ To see if there have been improvements since the previous comparison and included further algorithms where available.
   ▪ However, the approach may be deemed biased toward size-models and less suitable for PFT models.
   ▪ Also, fractionation rather than just dominance may now be considered more relevant.
   ▪ Therefore, include further approaches…

2) Independent validation against most-relevant in-situ data

3) Time-series analysis as PFT vs. in-situ and PFT vs. PFT comparisons

IOCS, May 2013
1) Dominant Phytoplankton Size Class Comparison on Pixel Basis i.e. repeat Brewin et al (2011, RSE)

Algorithm output

Satellite pixel dominated by
Pico =< 2µm
Nano = 2-20µm
Micro => 20µm populations

In-situ data

Method 1: Probability of detection

Method 2: Misclassification Matrix
2) Independent Validation Against In-Situ Data

- Independent validation against in-situ data
  - Satellite input is fraction / percentage of each PFT population
  - Estimate of uncertainty for both in-situ and satellite data needed

- Approach:
  - Calculate statistics: Pearson correlation coefficient (r); Type 2 regression (slope and intercept); Root Mean Square Error (RMSE); Bias; Number of samples
  - Score statistics e.g. as undertaken within Ocean Colour CCI intercomparison on basis of mean of all models
  - Plot frequency distribution, Taylor and Target diagrams
2) Time-Series Analysis

- Ability to produce a realistic time-series (phytoplankton seasonality) e.g.
  - plotting of time-series (Hoffmueller) diagrams and climatological plots (see below)
  - calculating time-lagged correlation coefficients
- Potential PFT comparison data sources: BATS, CPR, HOTS & L4

Example for plotting: part of Fig. 8.2 from Lavender et al. (2008, RS of the European Seas) showing climatological monthly plots for CPR cell counts and Chl-a

IOCS, May 2013
Summary

1. Optical- approach tends to show higher values than abundance-based in Microplankton

2. Discrepancy among algorithms using different sensors are obvious (but need an attention to interpret this result)

3. Generally, larger discrepancy among algorithms tends to appear in higher latitudes, and in [%] than in [mg/m³] (more focus should be directed to higher latitudes)

4. Picoplankton [%] showed a relatively larger discrepancies among algorithms (more focus should be directed to Picoplankton than Microplankton)

   • Microplankton abundance:
     ARC > SOC > NAT > NPC > SAT > SPC in [mg/m³]
     ARC > SOC > NPC > NAT > SAT > SPC in [%]

   • Picoplankton adundance:
     ARC > SOC > NPC > SAT ≥ NAT ≥ SPC in [mg/m³]
     ARC > SOC > NPC > NAT > SAT > SPC in [%]

Validation protocol is being constructed in reference to Brewin et al., 2011